Part I
Introductory and General
The study of semantic alignment: retrospect and state of the art

SØREN WICHMANN

1.1 Background and definitions*

This volume grew out of the conference ‘The Typology of Stative-Active Languages’, which took place at the Max Planck Institute for Evolutionary Anthropology in Leipzig, 20–22 May 2005, and which was organized by the editors of the present volume. It contains many of the papers presented at that conference, in addition to a few that were not presented there. The aim of the conference was to explore similarities and differences among languages of the ‘stative-active’ (or ‘split intransitive’, ‘agent-patient’, ‘agentive’, etc.) marking type, in particular similarities and differences regarding issues of argument structure and valency-related phenomena. The original call for papers invited a focus on ‘classic cases where an agentive S is encoded, through case marking, verbal agreement, or both, in the same way as A and non-agentive S in the same way as P’ and ‘where the agentive vs. non-agentive distinction is a pervasive feature of the grammar.’ In addition, we made an effort to ensure that different relevant geographical regions were all well represented. One does not organize a conference in the expectation of simply getting premeditated results, however, so we were only too pleased to also see papers addressing ‘non-classic’ instances of the language type under consideration as well as papers considering cases where the ‘agentive vs. non-agentive distinction’ is not necessarily pervasive, but does intrude on morphosyntactic alignment. We were particularly pleased and surprised that many papers addressed the topic suggested from a diachronic perspective; and these papers, as a whole, have greatly improved our understanding of the mechanisms by which semantic alignment may develop or disappear, as summarized towards the end of this introduction.

Given the confusion in the literature over terminology related to the topic under consideration, and the proliferation of terms, none of which is quite adequate, we proposed the label ‘semantic alignment’ for the language type under

* I would like to acknowledge comments from my co-editor as well as from Andrej Malchukov.
consideration. The term 'Split-S' (Dixon 1979) only makes sense when one views semantically aligned languages as somehow derivative of accusative or ergative languages; both accusative and ergative languages have a morphosyntactically relevant 'S' category, but since semantically aligned languages do not, it makes little sense to posit such a category, only to have it split up into the subcategories such as Sa and Sp. The term 'split intransitive' (Merlan 1985) similarly implies an S category, although less directly. 'Stative-active' or simply 'active' are also unfortunate terms, since the type of language to which they have been applied include languages for which the dichotomy of stativity vs. activity does not necessarily constitute a major organizing principle. Such terms should only be used to denote the subtype of semantically aligned languages where the argument of stative verb is treated like the patient of transitive verb, and the argument of an active verb like the agent of a transitive one. A similar problem attaches to the terms 'agent-patient' or 'agentive', since they carry the assumption that agentivity is the factor underlying the differential treatment of arguments of intransitive verbs across languages of the type in question. Again, such terms could be used for the subtype where the implied assumption holds true, but as general terms they are misleading. The advantage of the term 'semantic alignment' is that it does not carry any particular assumption about factors affecting the differential treatment of subjects of intransitives, except that they are semantic (rather than syntactic) in nature. Since the broad term 'semantics' subsumes both semantic roles, aspect, and Aktionsart (lexical aspect), the term covers different subtypes and also accommodates different ways of looking at one and the same phenomenon. It has been debated whether Aktionsart or semantic roles are more fundamental to the phenomenon at hand; using the term 'semantic alignment' avoids the introduction of this controversy into one's terminology. A potential negative reaction to the proposed term—the only one that I can think of—is that it is not always straightforward to predict, from semantic principles, how the arguments of all intransitive predicates in a given language belonging to the type under consideration are going to be treated; there may be—and, indeed, usually are—predicates that behave exceptionally. The existence of such exceptions, which need to be specified lexically, do not preclude the fact that the differential treatment of intransitive arguments is on the whole semantically motivated, which is what matters. In addition to being broad and theory-neutral, the term 'semantic alignment' has the further advantage that it contrasts naturally with the term 'syntactic alignment', which, if felt necessary, may be used to subsume ergative and accusative languages under one label (tripartite languages would belong here as well). Common to syntactically aligned languages is that the encoding of verbal arguments depends on the valency, and not the meaning, of predicates. Moreover, most syntactically aligned languages have constructions such as passives that require a dissociation between semantic and syntactic roles, whereas semantically aligned languages tend not to have such constructions (Klimov 1977, Foley and Van Valin 1977, Wichmann, to appear).
1.2 A brief retrospective overview of the study of semantic alignment

The field of language typology was not designed according to a master plan but rather represents a pool of research results and ideas which have accumulated in no small measure as a result of historical accidents. Thus, in order to characterize the motivation behind the present collection of papers it is necessary to trace the histories of some of the ideas around which they revolve.

Two of the foundational impetuses for the study of semantic alignment and related issues come from Sapir (1917) and Perlmutter (1978). Sapir’s achievement was to set up a clear typology of what is now known as ergative, ‘active’, accusative, and tripartite languages; he also made the suggestion, reconsidered from a diachronic perspective by some authors in this volume (see below), that inactive predicates in ‘active’ languages could be interpreted as having unexpressed impersonal subjects, e.g. that ‘I sleep’ could be interpreted as ‘it sleeps me’. Perlmutter explained the distinction between unergative and unaccusative subjects in terms of the Relational Grammar notions of initial and final strata; an unaccusative subject starts out at a deeper level (the initial stratum) as a 2 and an unergative subject starts out as a 1 (an analysis which actually bears some resemblance to Sapir’s, although the framework is of course radically different). The traditional approach in the Chomskyan literature is similar inasmuch as the unergative/unaccusative distinction is seen as arising through underlying syntactic structures where the surface subject of unaccusatives is internal to the VP, i.e. underlingly an object, while the surface subject of unergatives is also underlingly a subject (Burzio 1981, 1986).

Rather different traditions have arisen from these two stimuli, one being a functionally oriented whole-language typologizing approach, the other being a formally oriented approach to individual constructions (most commonly directed at Romance or Germanic languages). Because of the different traditions, the literature has created a sense that ‘active languages’ and ‘unaccusativity’ are somehow different, albeit related, phenomena. To what extent are they similar or different? It seems that the major difference between privileged whole-language typological features, such as alignment, word order, and morphological type, and features that are often considered less important when languages are typologized, such as ‘unaccusativity effects’, is the number of constructions for which the typological feature is relevant. Once this is realized, it becomes clear that the discrepancy between whole-language typology and construction-specific typology may only be superficial: whole-language typology is construction-specific too, but it happens that the types of construction which it is concerned with includes the widest possible range of constructions in a given language.

It could be claimed, then, that there is no fundamental empirical difference between the observation made by Sapir (1917) and by Perlmutter (1978) or between the two traditions ensuing from these contributions: both address differences in
constructions, and any explanation for different alignment patterns should also be applicable to unaccusativity effects and the other way around. Still, more work needs to be done to reconcile the traditions.

There is a third tradition which is fundamentally different from the two just discussed. This is the 'contentive typology' permeating the works of Klimov and defined in Klimov (1983). Klimov (1974, 1977) sees an inherent organic unity between alignment and other features such as the treatment of property concepts, possession, agreement, etc. For Klimov, an 'active' language is one which fulfills several typological criteria. The idea that there is an overall design uniting apparently disparate features of languages goes back to the Humboldtian notion of the 'genius' of a language, and is very different from (any version of) a constructional approach. Klimov's characterization of 'active' languages, however, has been severely criticized. Lazard (1986) has pointed out that all the traits that Klimov considers characteristic of the 'active' type are common in other languages as well, and in a similar but more explicitly quantitative vein Nichols (1990) showed that three of these 'active' traits, namely inalienable possession, inclusive/exclusive pronouns, and grammatical gender, are either not or not straightforwardly predicted by the presence of semantic alignment. Thus, Nichols finds that having a contrast between alienable and inalienable possession is correlated with head-marking and only indirectly with semantic alignment; that the presence of an inclusive/inclusive distinction is not predicted by the alignment type; and that gender only correlates with alignment type insofar as the accusative type disfavors gender. As Klimov suggested, the absence of canonical passives does indeed correlate, at least statistically, with semantic alignment (Wichmann, to appear); other possible correlates may be a preference for aspect rather than tense systems and the verbal encoding of property concepts (Wichmann 2005a), but these two additional possible correlations require further investigation.


Another controversial position is that of Dixon (1979), who treated semantically aligned languages as a subtype of ergative languages and saw case alternations in intransitives as being similar to split ergativity. Against this, Dahlstrom (1983) argued that semantically aligned languages are fundamentally different in having case marking alternations in intransitive clauses determined by lexical properties of the verb and, perhaps most importantly, that semantically aligned languages simply do not have a unified category such as ‘S’ (cf. also Mithun 1991: 542). Although Dixon has continued to push his position (Dixon 1994: 77–8), it is rare nowadays to see semantically aligned languages treated as a subtype of ergative ones.

The introduction of the notion of semantic roles (Fillmore 1968) allowed Van Valin (1977) and Foley and Van Valin (1977) to characterize semantically aligned
The study of semantic alignment

languages as languages lacking syntactic roles and thus having an orientation towards (semantic) roles rather than reference. Once it was realized that semantic roles are crucial for the description of semantically aligned languages, a new problem arose, namely the fact that languages often exhibit mismatches between the semantic status of an argument and its grammatical treatment. This problem has often been noted, e.g. for Georgian (Hewitt 1987), Svan (Sumbatova 1993: 265–267), and various North American languages (Mithun 1991), and it also turns up in studies of individual construction types. In the latter context it has been noted that ‘unaccusativity diagnostics do not uniformly pick up the same class of verbs, both within and across languages’ (Alexiadou et al. 2004: 8–9). There is a growing literature on such ‘unaccusativity mismatches’, e.g., Rosen (1984), McClure (1990), Dowty (1991), Gerds (1991), Zaenen (1993), and Legendre et al. (2006). Peter Arkadiev’s contribution to this volume (summarized below) addresses the problem of apparent mismatches by proposing a fine-grained typology of semantic roles which need not be instantiated in exactly the same way in all languages.

At present it is clear that the systematic study of semantic alignment is still in its infancy, despite the long history of attention to the phenomenon. More descriptive data needs to be brought to the fore, theories need to be tested on a wider range of data, scenarios for the diachronic understanding of the phenomenon must be developed further, and the discrepancy between the whole-language typologizing approach and the construction-specific one should be addressed. The present volume takes important steps towards accomplishing some of these tasks. The following summaries of individual chapters explain these steps; towards the end of this chapter the most outstanding overall contribution of the volume, namely new approaches to the diachrony of semantic alignment, is synthesized.

1.3 Summaries of chapters

Donohue’s chapter on the what-are-whats and what-are-nots of semantic alignment represents a tour de force through cases usually not taken into consideration in discussions of semantic alignment (although the author thinks they should be), as well as cases that have been, or conceivably could be, considered as instances of semantic alignment (although the author wants to exclude some of these). Donohue limits semantic alignment, properly speaking, to cases where verbal lexical semantics affects word order, case marking, or agreement. Phenomena such as Dutch auxiliary choice (familiar from the literature on unaccusativity) or accompaniment applicatives in Tukang Besi (perhaps less familiar) thus do not comply with the definition, even though verbal lexical semantics is involved in both cases. Similar exclusions apply to cases in which a core syntactic phenomenon such as agreement is involved in alternations, but is not governed by lexical verbal semantics, e.g. different strategies in Chamorro realis vs. irrealis clauses. Donohue contrasts these more spurious phenomena with true semantic alignment in languages such as Galela (West Papuan), Waris (Border), and Ambonese
Malay (Austronesian). Examination of patterns of alignment in the domains of agreement morphology, case marking, and word order show that in any of these both nominative–accusative, ergative–absolutive, and semantic alignment can be found. Alignment is also relevant for adjunct clauses, relative clauses, conjoined clauses, floating quantifiers, imperatives, secondary predication, and presumably other syntactic phenomena.

Other kinds of split, although they are not spurious but permeate languages, are to be considered ‘nots’ since they are not semantically governed, e.g. word order alternations conditioned by definiteness of objects in Puare (Macro-Skou), case marking conditioned by subordination in Mamean (Mayan), by tense-aspect in Sindhi (Indo-European), or by animacy (many languages).

Early on in the discussion of ergative languages, scholars set out to investigate whether accusative or ergative morphologies imply similar kinds of alignment in syntax. Donohue continues this line of investigation, concluding, like others before him, that ergative morphology usually combines with accusative syntax; but he also presents the case of Oirata (Trans New Guinea), which has accusative morphosyntactic alignment although it shows some evidence for an ergative syntactic pivot. A related but less often addressed issue is whether languages that show semantic alignment in their morphology also exhibit semantic alignment in syntax. This is probably seldom the case, but Donohue does cite the case of Eastern Pomo, for which it has been claimed that agentivity is involved in switch-reference.

The author then goes on to discuss splits in case marking involving more than two categories. Muskogean and Yapen (West Papuan) languages have three different agreement paradigms by which an S can be marked (nominative, accusative, dative), as well as two different cases for the A (nominative, accusative) and the P (accusative, dative). Icelandic shows as many as four different cases for Ss (nominative, accusative, dative, genitive) and three different ones for As (nominative, accusative, dative). The question is raised whether such phenomena should be considered under the label of semantic alignment, and if not, why not.

Like Malchukov, Mithun, and Nichols, Donohue also discusses ‘trans impersonal’ constructions, where the single argument is not in the case normally associated with monovalent subjects (as German mir ist kalt ‘I am cold’). Variations on this theme are illustrated, for example, by the use of the inverse with verbs such as ‘to fall’ in Tanglapui (Trans New Guinea), and by the treatment of experiencers as objects in Warendjor (Austronesian).

Received wisdom has it that semantic alignment is a phenomenon whereby S can align with either an A or a P. Donohue expands this narrow view, exemplifying cases where there are two ways of expressing an S, aligning with two ways of expressing an A (Haida, Isolate) or where more than one way of expressing a P aligns with more than one way of expressing an S (Kolana, Timor-Alor-Pantar). The first phenomenon, exemplified by Haida, recurs in Pilagá, a language to which the last chapter of this volume is devoted (see below).
The study of semantic alignment

Donohue’s contribution raises more questions than it answers, but nevertheless does offer some conclusions. One is that it is futile to analyse alternations in case marking along the lines of the ‘unaccusativity hypothesis’, as an alternation between surface and underlying subjects. This analysis is belied by cases such as Haida and Kolana. Another conclusion is that it makes sense to distinguish between truly semantically aligned languages and more spurious cases of ‘unaccusativity effects’, although it also makes sense to study the semantic conditions for both in unison. Some questions are answered more hypothetically. One such question is whether we can expect semantically aligned languages to also exhibit semantically conditioned syntax. The hypothesis is that we sometimes can, but in restricted domains. Another question is what we are to do with more than two different case markers for Ss. Is semantic alignment a phenomenon restricted to just two semantic roles—agent vs. patient (or Actor vs. Undergoer)—and why should it be? The hypothesis is that such cases are not necessarily to be excluded. Deconstructing the legacy of Sapirian typology, Donohue invites us to approach these questions, directing our attention towards individual constructions.

Malchukov explores a particular diachronic path by which semantic alignment (the author chooses the label ‘split intransitivity’) can arise, namely through constructions in which a P is reanalysed as an Sp under pressure to treat a prominent object as a ‘subject’. As already mentioned, this idea goes back to Sapir (1917: 85), who suggested that object-inflecting verbs in some Native American languages could be understood as having an impersonal agent and are therefore to be equated with what Haas (1941) termed ‘trans impersonal’ constructions. Following Haas’s treatment of Tunica, Malchukov turns Sapir’s idea into a diachronic explanatory framework which may be applied, if not to all semantically aligned languages, at least to some of those for which semantic roles rather than verbal aspect determine the differential treatment of arguments of intransitives and where patientive intransitives make up the minority of intransitives (i.e. semantically aligned languages with an ‘accusative base’ in the terminology of Nichols 1992). The author then goes on to present evidence, first from Native American languages and then from Papuan ones.

For Native American languages, Malchukov shows how some have trans impersonal constructions that clearly remain transitive because the impersonal subject is expressed overtly (e.g. West Greenlandic ‘it stormed us’ meaning ‘we were caught by the storm’), whereas in others (e.g. Lakota) there is no overt subject marking on what may have been erstwhile trans impersonals, implying the complete convergence of the two constructions and therefore the existence of a semantic alignment system. Two conditions required for the latter situation to arise seem to be the non-overt (or zero-) marking of third person subjects and the availability of verbal indexing for objects. Given these conditions, a trans impersonal construction may become formally identical to an intransitive construction where the single argument is marked as a patient. Because of the absence of documentation for sufficiently early language stages, the evidence must necessarily remain indirect. There
is no proof available to show that a situation such as the one in West Greenlandic has, in fact, given rise to a situation such as the one in Lakota, but Malchukov does identify cases where different dialects or related languages show different stages of reanalysis towards the final stage represented by languages such as Lakota. For instance, in West Greenlandic transimpersonal inflection is restricted to weather expressions, but in Siberian Yupik it extends to any intransitive predicate and conveys a sense of suddenness or non-volitionality, achieving the ability to make semantic distinctions reminiscent of fluid semantic alignment systems.

For Papuan languages, Malchukov shows how object experiencers may develop into non-canonical subjects and eventually into Sps. Examples of different stages along this trajectory are presented. For instance, in Amele, in a sentence such as ‘I am hungry’, the experiencer is coded as an object, but also displays a couple of subject properties, namely the ability to appear in topic position and to control switch-reference. In Hua, certain experiencer objects are treated as such in the indicative mood, but are reanalysed as intransitive subjects in the imperative. Finally, in Galela the dummy subject agreement prefix in transimpersonal predicates has become optional on certain verbs and has completely vanished from others, leaving a predicate where the argument is an object intransitive (see also Holton’s chapter in this volume).

In a final section Malchukov shows different examples where the path towards semantic alignment is blocked, for different reasons. The Iwaidjan (non-Pama-Nyungan) language Ilgar of Australia has alternative constructions for ‘he is sneezing’, one following the transimpersonal pattern and one leaving out the agent affix from the verb, effectively leaving an object-marked intransitive. However, even if this pattern were to generalize, semantic alignment would not arise because person marking operates on an ergative basis. Thus, in an ergative language an object-marked intransitive would be indistinguishable from a subject-marked intransitive. Another type of development of transimpersonals not leading to semantic alignment is found in Germanic language, e.g. the German construction Mich hungert ‘I am hungry’, which is sometimes replaced by Ich hungere. Here a transimpersonal is simply replaced by a subject-marked intransitive.

Arkadiev addresses the issue of how best to characterize the semantic parameters that determine argument encoding in semantically aligned languages. Arguing against a recent proposal by Primus (1999), he suggests that it is not so much the sheer number of Proto-Agent or Proto-Patient properties which determines whether the single argument of an intransitive verb will be encoded as Agent-like or Patient-like, but rather the presence of a particular distinctive property (or combination of different properties) in a given language. Through different case studies it is demonstrated that different languages may select different properties as being distinctive. Thus, the Southwestern Mande language Loma is sensitive to the aspectual distinction stative vs. dynamic, the Kartvelian language Georgian is sensitive to telicity, the Nakh-Dagestanian languages Bats and Tabassaran to volitionality, and the Pomoan language Central Pomo to
affectedness. The fact that individual languages select certain properties does not mean that there is no structured relationship among these different properties from a typological point of view. Thus, \([\pm\text{volitionality}]\) and \([\pm\text{change of state}]\) correlate inasmuch as \([+\text{volitional}, -\text{change of state}]\) normally imply agentive marking and \([-\text{volitional}, +\text{change of state}]\) normally imply patientive marking. The marked combination \([+\text{volitional}, +\text{change of state}]\) will result in either agentive or patientive marking, depending on whether \([\pm\text{volitionality}]\) is more prominent than \([\pm\text{change of state}]\) or the other way around in a given language.

Arkadiev goes on to show how the approach of Levin and Rappaport Hovav (1995, 2000) to unaccusativity in English can achieve the same kind of cross-linguistic explanatory adequacy as the specification and ranking of proto-properties, provided that their ‘linking rules’ for the assignment of deep subject (roughly = actor) and deep object (roughly = undergoer) status are ranked and their default rule is made to assign either deep subject or deep object status (and not just the latter) to arguments of intransitive predicates depending on the language in question. Arkadiev thus paves the way for a refined theoretical approach to the semantics of semantic alignment systems, an approach which uses the solid foundation of the ‘proto-properties’ of Dowty (1991) and which, as the author shows, may be recast along the lines of Levin and Rappaport Hovav (1995, 2000). Although this is not made explicit, the approach is perhaps even closer in spirit to Optimality Theory with its notions of constraint ranking (Prince and Smolensky 2004). In short, Arkadiev’s paper should have an appeal to theorists of a variety of persuasions.

From a whole-language typologizing approach to construction-specific approaches, a logical next step is lexical typology. In what, in this regard, is a unique contribution to this volume Nichols takes such an approach, surveying argument encoding for verbs corresponding to 20 verb glosses across 41 languages of Eurasia, the Pacific, and the Americas. The glosses were chosen so as to have a mixture of agent, patient, and experiencer subjects, semantically speaking. Plots showing the percentages of A- vs. O-coded Ss for the sample of verb meanings in the various languages show a continuum where ergative and accusative languages cluster towards opposite extremes and languages normally classified as semantically aligned fall in between. A similar result is obtained from plotting the number of A vs. non-A codings of S and A. From such plots Nichols concludes that ‘from a lexical-typological perspective, “split subject” covers a diffuse and fairly diverse range of languages, without a very clear distinction from ergative at one end of the range to accusative at the other.’ It is perhaps not surprising that, as one moves from a categorical whole-language approach to a statistical approach, the results will begin to show clines rather than clear-cut distinctions, but, as Nichols admits, the sample can also be manipulated so as to bring out more vs. less clear clusters. Insofar as the results partly follow from the approach, using the new approach will be motivated by the kinds of result that can be obtained. One such result could, for instance, be to reveal more subtle areal distributions than will emerge from the whole-language approach.
Other than developing a new type of methodology, Nichols also achieves an important insight into the question of why, in the traditional understanding, semantic alignment is rare in Eurasia. Whether semantic alignment is indeed rare here, she argues, depends on one’s point of view. If verbs that take dative subjects in Eurasia are treated on a par with verbs that take O-marked subjects in semantically aligned languages in Americas and the Pacific, then the two groups of languages begin to look much more similar. As Nichols points out, there is a good reason to equate the two kinds of argument encoding: a direct object language will typically use dative marking for non-A coded subjects, whereas a primary object language will use O coding; and direct object languages dominate in Eurasia, whereas primary object languages are somewhat more frequent than direct object languages in the Pacific and the Americas. The insight that ‘the uneven geographical distribution of classic stative-active languages worldwide is connected to, and probably a consequence of, large areal differences in object alignment’ of course begs the question of why there are distributional differences in object alignment. Still, the hypothesis may turn out to be crucial for the diachronic understanding of semantic alignment insofar as it identifies a possible necessary-though-not-sufficient condition for the development of semantic alignment. Thus Nichols’s paper provides a corollary to the hypothesis of other contributors to this volume (Malchukov, Mithun, Donohue, Holton) that transimpersonal constructions have paved the way to the development of semantic alignment by suggesting, indirectly, that such constructions need to occur in the context of a primary object language in order to develop into semantic alignment systems.

Ket, which belongs to the small Yeniseic language family all of whose other members are extinct, has sometimes been described as semantically aligned. Vajda disputes this, and instead describes the language as one which possesses a number of lexically selected conjugation classes that are not easily fitted into a single overall alignment type; according to Vajda (2004: 49) one of the classes, which he calls the ‘Active Conjugation’, is all that remains of an agentive-patientive contrast. In his contribution to the present volume Vajda deals in greater detail than in earlier works with diachronic issues, pursuing the hypothesis—accepted by most specialists—that Proto-Yeniseic had a semantic alignment system, and tracing the changes to Modern Ket through comparisons with data from other Yeniseic languages. A model of Proto-Yeniseic verbal inflection is offered as a point of departure for understanding the situations in the daughter languages. In this model, only 1st and 2nd person patientive subjects are indexed on the verb by means of pronominal prefixes. Agreement with 3rd persons was expressed by means of animacy classifiers, also prefixed. Free pronouns (perhaps sometimes cliticized) later became part of the verbal morphology either as prefixes (e.g. in Ket) or

1 The model seems to have been reached only partly by systematic reconstruction of individual morphemes using the comparative method, and is thus somewhat more abstract and hypothetical than normal reconstructions.
The study of semantic alignment

suffixes (e.g. in Kott). Contact with suffixing Samoyedic, Turkic, and Tungusic languages appears to have been responsible for much of today’s complexity by causing affix positions to get shuffled.

Traces of an erstwhile semantic alignment system are said to have been best preserved in Kott, which has an intransitive conjugation used with ‘active’ verbs involving a ‘sentient or at least an active subject’ and two intransitive conjugations for ‘inactive’ verbs. A fourth intransitive conjugation is reserved for change-of-state or resultative verbs. There are no direct identities of expression such that A = Sa and P = Sp, so ‘semantic alignment’ in Vajda’s usage does not refer to the canonical type of system. Rather—and this seems to hold for the reconstructed system as well—the term simply refers to a differential treatment of subjects of intransitives, in this case through different conjugations separate from transitive conjugations. That is, we are witnessing a four-way distinction between A, Sa, Sp, and P, where each can be encoded in multiple ways. Since such a system could potentially have developed into true semantic alignment, the title of the paper should perhaps have referred to the failure to acquire semantic alignment rather than to the ‘loss’ of such a system. In any case, the chapter demonstrates how formal expressions of grammatical relations can become highly idiosyncratic when different factors such as semantics, language contact, preservation of inherited structure, prosody, economy of expression, and the lexicon exert their various influences.

The chapter by Khanina deals with Tundra Nenets, a language belonging to the Samoyedic group of Uralic. Samoyedic languages are spoken to the north of the former Yeniseic area, and, as mentioned by Vajda, languages of the two groups are likely to have been in contact. Like proto-Yeniseic (and Amis, see below), Tundra Nenets is an example of a language which, even if it does not meet the definition of canonical semantic alignment, exhibits a semantically conditioned argument encoding. It exhibits accusative alignment (at least in part\(^2\)) and a mixture of head- and dependent-marking. The focus of Khanina’s description is the differential treatment of argument encoding for intransitives. There are two inflectional paradigms, A and B, and a given verb will either take one or the other or may alternate between the two, with consequences for the semantic interpretation. Of importance for the selection of paradigm is whether an action is ‘homogeneous’, i.e. not decomposable into different phases or ‘non-homogeneous’, i.e. spanning an inception phase. Paradigm A is used for the former and B for the latter type of action. ‘Fluid’ verbs describe the inception (‘non-homogeneous’) phase of a given action when taking paradigm B and the stable (‘homogeneous’) phase when taking paradigm A. The contrast of ‘homogeneity’ is related but not identical to that of telicity inasmuch as homogeneous actions are telic, whereas non-homogeneous actions may be either telic or atelic. The

\(^2\) A, S, and P are expressed in distinct ways (typically) when a referential P is involved, following a tripartite pattern, whereas (typically) non-referential Ps are indexed by means of the set of pronominals that also express an S involved in a ‘homogeneous’ action (see below), following an accusative pattern.
North Samoyedic languages are the only ones in the Uralic family that exhibit the B paradigm. Khanina therefore assumes that the differential S marking is an innovation in North Samoyedic. Given the affinity between the kinds of semantic distinctions involved in differential S marking in Tundra Nenets and the telic vs. non-telic distinction underlying at least some systems of semantic alignment (Van Valin 1990, Arkadiev, this volume: section 4.2.2), Khanina draws a parallel between Tundra Nenets and semantic alignment.

In his contribution on Basque, Aldai describes the historical evolution from what he describes as a ‘loose ergative alignment’ system in 15th–17th-century Basque (‘Old Basque’) to a ‘patientive semantic’ case-marking system in Western Basque. Data are also presented from Eastern and Central Basque, the former of which remains more conservative, while the latter represents an evolutionary stage intermediate between Old Basque and Western Basque. ‘Strict’ ergative alignment, known from many languages in the world, is a situation where all subjects of intransitives are marked as absolutive and where non-canonical transitive sentences, including sentences involving non-definite or oblique objects, similarly require subjects to take absolutive case. In a ‘loose’ ergative system such as that of Old Basque, there is no requirement that non-canonical transitive sentences have absolutive subjects. Finally, a ‘patient semantic alignment system’, such as that of modern Western Basque, extends the range of intransitive constructions allowing for ergative marking further to agentive intransitives. In this system, subjects of intransitives which combine features of agentivity and patientivity (typically motion verbs) are treated like subjects of intransitives which are purely patientive (‘die’, ‘fall’, etc.), taking absolutive marking. In Western Basque, then, agentive verbs of manner of motion (‘dance’, ‘jump’, etc.) typically take ergative subjects, whereas they take absolutive subjects in Central and Eastern Basque, and the same is true of agentive activity verbs such as ‘fight’ or ‘play’. This paper breaks new empirical ground by describing differences across constructions and across dialects. No attempt is made here to inquire into the motivations for the changes that led to the current Western Basque situation, but the observations made should feed into future comparative work on other languages exhibiting similar changes and thereby help to disclose the ultimate mechanisms behind the particular pathway towards semantic alignment that led to the modern Western Basque system.

Klamer presents an overview of semantic alignment systems in the Austronesian languages Acehnese, Kambera, Kedang, Taba, Larike, Selaru, and Dobel and the Timor-Alor-Pantar languages Klon, Abui, and Tanglapui. All but one of these languages are spoken in eastern Indonesia. In most of these languages the properties of arguments rather than properties of inherent verbal aspect determine the encoding of arguments of intransitives. Klamer finds that in eight of them, Dowty’s (1991) agent proto-role [±volitional] adequately accounts for the differential marking, whereas in the two remaining languages (Klon and Tanglapui) the patient proto-role [±undergoer of change of state] is distinctive. In the two
The study of semantic alignment

languages where verbal aspect does play a role, Taba and Dobel, semantic features of the arguments nevertheless co-determine argument encoding, so for all the languages studied by Klamer the generalization holds that such features are crucial. Another similarity is that all the languages surveyed use dependent pronouns (affixes or clitics) to encode S and to encode either A or P or both A and P. This is not just a peculiarity of these eastern Indonesian languages; as noted by several scholars (Nichols 1990, Mithun 1991, Siewierska 2004), semantic alignment is normally associated with head-marking morphology. There are naturally also differences among the languages observed (and they furthermore as a set contrast with numerous other languages in Indonesia which do not exhibit semantic alignment). The literature on semantic alignment as concerns the Indonesian area has been dominated by references to Acehnese (though see Donohue 2004b). Klamer’s chapter is therefore a welcome step towards widening the horizon and bringing more of the existing diversity to the fore.

While Klamer took an areal and synchronic approach, Holton’s approach is genealogical and diachronic. He provides an overview of alignment patterns in the small North Halmahera family of non-Austronesian languages (North Maluku, eastern Indonesia). These languages present what may be interpreted as stations along a diachronic development from syntactic (accusative) alignment (Tobelo), over canonical semantic alignment (Galela), back to syntactic alignment (W. Makian). Holton shows that these differences are actually rather superficial, inasmuch as they relate to formal constraints on the occurrence of pronominal markers on verbs rather than to deeper differences in semantic distinctions among predicate classes. The starting point of the proposed development is a type where stative predicates carry an impersonal agent, as in the ‘it-sleeps-me’ of Sapir (1917). This is attested in Tobelo. From this type there is only a small step to canonical semantic alignment, since all it takes to develop such a system is for the impersonal marker to be lost. How small this step is can be demonstrated by contrasting 19th-century Galela, where the impersonal marker is often still retained, with modern Galela, where it is lost. Like 19th-century Galela, Tabaru, in which the impersonal marker of the trans impersonal construction only shows up in the 1st person plural, represents an intermediate type. If pronominal P markers on predicates become optional and eventually get lost, as happens in some North Halmahera languages, a movement away from semantic alignment sets in. Now it is only the presence vs. absence of A markers on intransitives which may signal a semantic distinction; but since arguments of stative intransitives are not exactly expressed like patients of transitives, a language of this type can hardly be described as being standardly semantically aligned. This exact intermediate stage seems not to be represented within the family, but the final stage, where all intransitives carry A markers, is. Holton hypothesizes that this stage may be arrived at by an extension of the use of A markers in analogy with active intransitives.

The chapter by Tsukida rounds off the section discussing languages of the Pacific, moving the perspective to Taiwan, the homeland of the widespread
Austronesian family. The focus is on Amis, an Austronesian language of eastern Formosa. In the context of this book Amis is interesting even if it does not exhibit semantic alignment in any canonical sense: given its relatedness to and similarities with languages that do, it contributes a comparative perspective on semantic alignment. Amis exhibits a variant of the well-known Philippine-type ‘focus system’ where one argument is selected as subject (marked by nominative case) and where the semantic role of this subject is marked on the verbs by means of either an Agent Voice pre- or infix (for As of transitives or Ss of intransitives) or a Goal Voice suffix (signalling other semantic roles). Case marking is carried by free pronouns and determiners. Agent Voice marking is distributed over four conjugations which specify various combinations of values of the features [±state], [±affected], and [±control] for the subject. Many verbs can alternate between different conjugations with concomitant semantic changes relating to causation, control, volitionality, affectedness, effectness, etc. Tsukida briefly discusses similarities and differences between the four conjugation classes of Amis and the system that has been reconstructed for Proto-Austronesian, which also exhibits multiple verb classes. She also discusses similarities with canonical semantic alignment systems, pointing out: ‘The ma conjugation is used when the notional Patient is the subject and mi conjugation when the notional Agent is the subject. Most of the intransitive stems take either mi conjugation or ma conjugation in Amis. This is strikingly analogous to P-marking and A-marking in ordinary semantic alignment systems.’ She furthermore notes that the same semantic features that had served to describe the functions of the conjugation classes in Amis are also customarily used to describe semantic alignment systems.

If, indeed, Amis is relatively conservative with regard to its conjugation classes and the absence of canonical semantic alignment, then languages such as those discussed by Klamer would represent innovations. Much in the spirit of Holton, one could then infer that the existence of semantic distinctions among predicate classes is a more stable feature than the formal encoding of arguments. Indeed, an important difference between Amis and the semantically aligned Austronesian languages is that the latter normally exhibit pronominal markers bound to the predicates, following a head-marking pattern, whereas the former has free, case-marked pronouns. Given the already-mentioned correlation between head-marking and semantic alignment the hypothesis might be ventured that the general conditions for the development of semantic alignment in many Austronesian languages was a combination of shifts from dependent-marking to head-marking morphology and the prior existence of semantically defined predicate classes.

In her overview of semantic alignment systems in selected areas of North America, i.e. Northern California, the northwest coast, and the southeast, Mithun claims that such systems are not necessarily as stable as has been assumed by some scholars (E. Sapir in Golla 1984: 349, Nichols 1992), and that they may
arise under the stimulus of diffusion, and suggests various mechanism by which this may happen (see also Donohue 2004b on the existence of an areally defined semantic alignment zone in eastern Indonesia). The case studies presented include the following:

(i) Yuki (Northern California) is shown to have pronouns similar in shape to related Wappo, but their functions are more similar to unrelated Pomoan languages. Thus both Yuki and Pomoan mark S as either A or P, depending on the parameters of affectedness and control. An erstwhile accusative alignment system in Yuki could have been reanalysed as a pati entive system if transitive clauses where 3rd person subjects are omitted were reanalysed as intransitive, former objects now being treated as grammatical patients. (ii) Remaining in Northern California, we encounter Karuk, which exhibits features of semantic alignment and may have been influenced in this regard by Chimariko. This could have happened through the same type of reanalysis as posited for Yuki. (iii) Wiyot (also Northern California) does not have a fully developed semantic alignment system, but still offers some insight into yet another way in which such a system could develop. It appears that an erstwhile passive suffix followed by pronominal subjects markers have been reinterpreted as markers of a patientive argument. (iv) On the Northwest Coast of the Americas we encounter Tlingit, which is the only member of the Tlingit-Eyak-Athabaskan family to show semantic alignment. Mithun suggests that Tlingit subjected its nominative-accusative system to reanalysis under contact with Haida, and that the mechanism whereby this happened was to reinterpret indefinite subject verbs as intransitives without specified agents. (v) In the southeast, semantic alignment permeates both the Muskogean family of languages and the isolates Atakapa, Chitimacha, Tunica, and Natchez. The author considers it likely that semantic alignment is an innovation in at least some of the languages, and that areal influence is involved, either from within or from outside the southwest.

While Mithun admits that the types of reanalysis discussed in her chapter could happen spontaneously in any language given the right conditions, she considers it likely that language contact is an additional factor which facilitated the development and spread of semantic alignment in Native North American languages.

The absence of devices such as passives and antipassives, which serve to foreground an argument at the expense of another, have been claimed to be characteristic of semantically aligned languages (e.g. Foley and Van Valin 1984: 155–9). The contribution of Pustet and Rood presents a detailed analysis of the exceptional case of Lakota (Siouan). As with some other semantically aligned languages, Lakota may employ a 3rd person plural actor in an impersonal sense to convey agent backgrounding (as in ‘they shot him’), Lakota has taken this type of construction one step further, allowing for the agent to be expressed. A construction such as (1) literally means ‘they, the bear, chased that man’, but is reanalysed as a passive construction—the -pi third person plural marker now functioning as a valency-decreasing derivational affix.
Søren Wichmann

(1) Wicháša hé mathó Ø-khuwá-pi.
man that bear 3SG.PAT-chase-PASS
'That man was chased by bears/a bear/the bear(s).' (= ex. 67)

The authors suggest that the development of a passive could have taken place under the influence of English, but also discuss the possibility that -pi already in early Siouan times indicated 'marked focus', i.e. a shift from default attention focused on the actor to attention on the undergoer, and that the plural meaning in fact is a later development. An argument for this is that the cognate of -pi in Omaha includes 'marked focus' as one of its functions and that the Caddoan language Wichita, which may be a distant relative, has a suffix with a similar range of meanings. Other affixes in the language serve to dereferentialize patients: wa- 'patient dereferentializer (usually of inanimates)', taku- 'inanimate patient dereferentializer', and wicha- 'animate patient dereferentializer'. Just as -pi derives from a 3rd person plural agent marker, wicha- derives from a 3rd person plural patient marker.

In his contribution, Palancar examines the coexistence of the two major subtypes of semantic alignment—agent/patient as well as active/stative patterns—in a single language, namely (the San Ildefonso Tultepec variant of) Otomi (Otomanguean, Mexico). The agent/patient pattern concerns a distinction between a small set of intransitive verbs whose single arguments are expressed by means of the same verbal suffixes as objects of transitives and which are opposed to the rest of the intransitives, which take subject markers. In the author’s materials only 10 patientive intransitives have been identified, including 'be/get tired', 'be/get burnt', and 'be/get pale'. The active/stative pattern is constituted by a class of semantically stative verbs which describe property concepts and take patientive inflection. They are moreover recognizable by certain morphological peculiarities, including a special paradigm of proclitics, a nasal prefix, and the lack of a contrast between forms used before a prosodic and/or a clausal boundary ('free forms') and forms used intracausally ('bound forms'). The author proposes that the agent/patient pattern is old, whereas the category of stative verbs is a more recent innovation, which adopted some characteristics of what Palancar calls 'stative forms' of agentive verbs (these have also been designated 'participles' by others)—a category already existing in the language. Stative forms are impersonal forms that describe the resulting state of an action corresponding to a transitive verb. These verbs take a dummy 3rd person proclitic to express tense-aspect-mood categories. Thus, morphologically a stative form such as 'I am comforted' inflects like a transitive ('s/he comforted me') but has an impersonal reading. The same transitive morphology is carried over to stative verbs. So a stative verb such as 'I am tall' also takes the dummy 3rd person proclitic and an object marker for the single argument ('me'). Since active/stative distinctions are rarely reported in Otomanguean languages, and since little is known about how such distinctions emerge in general, Palancar’s chapter is an interesting contribution to description and diachrony.
In her study of Guaraní (Tupi-Guaraní, Paraguay), Velázquez-Castillo investigates the closest equivalents of voice categories in this language, i.e. the categories labelled ‘inverse’ and ‘reflexive/passive’ by some other linguists. She argues that these labels are inadequate for the description of Guaraní, as are analytic concepts such as ‘subject’ and ‘object’. Instead, she operates with the grammatical relations coined ‘event source’ and ‘event site’ in Velázquez-Castillo (2002). In her own words, ‘the event source is a participant construed as the initiator or origin of a dynamic event, which may or may not have control over the situation. The event site is a participant construed as containing the situation denoted by the predicate.’

The ‘inverse’ does function as a normal inverse in transitive contexts. The highest-ranking participant on the hierarchy $1 > 2 > 3$ is indexed on the verb, and inactive pronominal markers are selected when 1st person is acted upon by a 2nd or 3rd person or when a 2nd person is acted upon by a 3rd person. When a speech act participant acts upon a 3rd person, or when a 1st person acts upon a 2nd person, an active pronominal marker is selected. An analysis in terms of direct-inverse, however, has limited applicability. Guaraní (like some other Tupi-Guaraní languages) exhibits a prefix $r-$, which, when brought into the purview, requires the analysis to be modified. This prefix turns up between the person marker and the stem in inverse constructions (when the verb is vowel-initial), and could thus be interpreted as an inverse marker (e.g. Payne 1994b). A problem for this analysis, however, is that $r-$ also turns up when inalienable nouns are possessed, and on monovalent inactive predicates designating physical or emotional conditions. Therefore Velázquez-Castillo argues that what $r-$ marks is proximal argument-predicate connections in inactive situations.

The second problem discussed by Velázquez-Castillo relates to constructions involving a prefix $je$-. It is exemplified by a sentence such as Toma o-je-japi (Toma in active-middle-shoot), which can mean either ‘Tomas was shot’ or ‘Tomas shot himself’. This accounts for the traditional analysis invoking the notions of passive or reflexive. The author, however, argues that the function of $je$- is better described as signalling ‘the cancellation of the default outward projection (i.e. outward directionality) inherent in an active predicate’, and should be analysed as a kind of middle voice. Various pieces of evidence are mustered, the most compelling perhaps being the occurrence of $je$- with monovalent predicates, e.g. Kañada-pe o-je-jeroky ...(Kañada in active-middle-dance) ‘in Kañada there is dancing…’

The description of what would be the closest equivalents to voice in Guaraní, then, shows that they are insensitive to transitivity, and that what matters is rather semantic relations. New theoretical notions must be developed to adequately describe these relations, and Velázquez-Castillo’s contribution does precisely this.

Danielsen and Granadillo introduce the agreement systems of two Arawak languages, Kurripako of the North Arawak subgroup (spoken in the border zone of Venezuela, Columbia, and Brazil) and Baure from the South Arawak
Søren Wichmann

Kurripako verbs exhibit subject agreement prefixes (A markers) and object agreement suffixes (P markers), the latter for 3rd person objects only. Some intransitive verbs, described by the authors as 'static', take agreement markers that are identical with the P set, others—the 'active' ones—take A markers. Thus, Kurripako behaves like a prototypical semantically aligned language—more specifically, it is argued, one which is sensitive to the parameter of eventhood, rather than agentivity. In contrast, Baure behaves more like an accusative language, in that nearly all intransitive verbal predicates take the same nominative agreement markers which also express subjects of transitives. Here, the differential marking of subjects of intransitives mainly involves verbal predicates as opposed to predicates derived from nouns and adjectives, the latter taking accusative agreement markers. There are, however, some verbs which fail to conform to the general accusative system, the clearest case of an intransitive being kwo- ‘exist, be’, which takes accusative agreement markers. Moreover, the verbal base -moro’in(o)- has alternate usages, one of which involves the meaning ‘to be thirsty’ and triggers accusative agreement for its experiencer argument. On the basis of these pieces of evidence, the authors hypothesize that earlier stages of Baure may have had a semantic alignment system involving some fluid marking similar to that of Kurripako, but that it now shows only a few traces. This case study appears to confirm Mithun’s suggestion that semantic alignment is not as stable a genetic marker as has been claimed.

Vidal presents the system of semantic alignment in Pilagá (Guaykuruan, Argentina). This is highly unusual typologically. It is not the case that S is expressed either like A or like P depending on some semantic parameter, as in canonical cases of semantic alignment. Rather, there are two different markers, called set A and set B, which can both be used to express A or S. P is marked by a third set, different from sets A and B. While markers from either set A or set B are always present, the P markers are mutually exclusive with an overt (nominal or pronominal) expression of the corresponding argument. Vidal does not treat differential marking of the agent of transitive in any degree of detail, but concentrates on the distribution of set A and B markers over different classes of intransitive. The kinds of event that take set A include most physical activities where the single argument is a performer of the action, but also some events of perception and involuntary bodily processes as well as states (e.g. ‘be sick’, ‘have a headache’). Verbs that exclusively take set B are a small minority and include certain state verbs (‘be happy’, ‘be sincere’), verbs of bodily postures or spontaneous bodily processes, and certain activity verbs (‘to drink’, ‘to run’). The difficulty in establishing clear differences between set A- vs. set B-taking verbs suggests that much lexicalization has been going on. On the other hand, the vast majority of intransitive stems can alternate between the two sets, and these alternating verbs are more clearly circumscribed semantically. When events are induced (externally caused), non-reflexive, or involve intentionality, set A is employed, whereas when events are spontaneous (inchoative), reflexive, or non-intentional, set B is used.
The study of semantic alignment

With motion verbs and other verbs that may be said to imply motion in some way (e.g. ‘to give’, ‘to learn’), the alternation between person marker paradigms acquires deictic implications, such that set A generally implies motion away from the speaker’s vantage point (the deictic centre) and set B motion towards it. In Velazquez-Castillo’s treatment of the Guaraní construction which some have labelled ‘reflexive’ (see above), it was argued that default outward projection or direction of actions described by active verbs is a grammatically relevant notion in Guaraní. The behaviour of the Pilagá motion verbs demonstrates in a very concrete way that such a notion may, indeed, sometimes be relevant for the understanding of case marking. Applying Velazquez-Castillo’s notions to Pilagá, forms taking set A may be said to imply default, outward projection, while those taking set B ‘cancel’ this default direction, implying an inward projection. Vidal herself analyses the case alternation in terms of the notions ‘viewpoint’ and ‘affectedness’.

1.4 The diachrony of semantic alignment

Since the literature on semantic alignment systems has not hitherto addressed questions regarding how such systems develop, one of the outstanding new features of the present collection of papers is a set of approaches to this problem.

At the beginning of this introduction I defined a ‘classic’ semantically aligned language as one where an agentive S is encoded—through case marking, verbal agreement, or both—in the same way as A, and non-agentive S in the same way as P. As several contributions to this book show, the definition at best applies to a prototype. Languages can differ in many ways from the prototype and yet exhibit a case-marking or agreement system which is organized according to semantic principles. Nevertheless, the prototype remains a useful point of view from which the various deviant cases may be studied, since it can be argued that there are natural pathways by which the prototype may develop, and the failure of all the conditions for its development to take place, or of certain further developments, may then explain the less prototypical cases.

Which, then, are the natural pathways and conditions for the development of semantic alignment? Malchukov and Mithun, in particular, argued that transpersonal constructions may lead to the reanalysis of an erstwhile transitive predicate with an experiencer undergoer as an intransitive carrying object person markers. By this mechanism syntactic objects become semantic patients. This diachronic hypothesis finds support in the fact that semantically aligned languages tend to be head-marked. This fact can be explained if a requirement for the development of semantic alignment is zero-marking of 3rd person subjects. In languages having free pronouns, 3rd persons will tend to not be zero-marked. In her contribution Nichols suggested a further correlate, noting that most semantically aligned languages have primary/secondary objects rather than direct/indirect ones. In languages having indirect objects, experiencers will tend to treated as datives, which is why such languages (e.g. Icelandic, discussed by Donohue) do
not develop canonical semantic alignment. Finally, Malchukov made the observation that if a language changes a syntactically aligned system to a semantically aligned one through the reanalysis of transimpersonal constructions, the point of departure has to be an accusative system—in an ergative system there would be no difference between the erstwhile experiencer object and arguments of intransitives (incidentally this provides a good diachronic reason for not viewing semantically aligned languages as subtypes of ergative ones, as Dixon 1979 did). This last observation means that we should look for different kinds of mechanism when dealing with semantically aligned languages that have developed out of ergative ancestors, such as Western Basque (discussed by Aldai) or the Mayan languages Chol and Chontal (Gutiérrez and Zavala 2005). Indeed, Aldai showed that Western Basque has extended the marking of subjects as ergative to intransitive contexts, which is in a sense the opposite development. To sum up: the pathway from accusative to semantic alignment goes through the extension of the domain of the accusative argument to intransitive contexts, whereas the pathway from ergative to semantic alignment goes through the extension of the domain of the ergative argument to intransitive contexts.

The existence of the various highly specific conditions on the rise of semantic alignment helps us understand why there should be certain areas in the world from which such systems are absent or rare (Africa, Southeast Asia, Eurasia). On the other hand, Mithun’s contribution, which demonstrates that semantic alignment is prone to diffusion, helps us to understand why semantic alignment can be very widespread in other areas (the Americas, insular Southeast Asia).

Holton’s small-scale but highly revealing case study also suggests that the presence of semantic alignment is prone to change, given that it is no more stable than the phonological shapes of its exponents. Holton suggested that the semantic verbal distinctions that underlie the differential treatment of pronominal arguments of intransitives are more stable and fundamental than the pronominal markers themselves. Tsukida’s study points in a similar direction: Proto-Austronesian seems to have been somewhat similar to the Formosan language Amis with respect to the existence of semantically distinct verb classes; it would not have exhibited the prototypical semantic alignment that is found in some daughter languages (cf. Klamer’s chapter), then, but it would have had the semantic ‘ingredients’. Similarly to Holton’s study of Halmaheran, Vajda’s study of Yeniseic shows how formal changes affecting verbal person markers can lead to the loss of semantic alignment.

The above observations on the rise of semantic alignment apply to the patietive subtype. The active-stative subtype is rarer and perhaps somewhat less well understood. Palancar suggested that the stative verbs in Otomi were modelled on certain participle-like forms, and a comparison of the agreement systems of the Arawak languages Kurripako and Baura described by Danielsen and Granadillo suggest that a differential treatment of the marking of subjects of verbal and nominal or adjectival predicates (as in Baure) could be the basis from which a
The study of semantic alignment

A semantic alignment system based on stativity (as in Kurripako) might develop. Both the Otomí and the Arawak case studies, then, suggest a scenario whereby the argument coding associated with an inherently stative part of speech (noun, participle, adjective) may spread to some intransitive verbs that have a stative meaning. This scenario may well apply to other cases of stative-active languages, although this remains to be seen.

We hope that, partly as a result of the materials and hypotheses presented in this volume and partly as a result of the dissemination of new data from less well-known parts of the world, we shall soon see the development of more complete theories of semantic alignment.

Alternatively, Arawak may historically have had a more widespread split based on eventhood which is reflected in a reduced way in Baure.
2

Semantic alignment systems:
what’s what, and what’s not

MARK DONOHUE

2.1 Introduction

This chapter has a single aim that is both modest and overly ambitious. I wish to
address two (related) questions: what, in terms of participant-coding strategies,
should be considered in a discussion of ‘semantic alignment’, and what we can say
about languages with ‘semantic alignment’ languages in terms of the two related,
but distinct, dimensions of morphology and syntax.

Semantic alignment is, as stated in the introduction to this volume, the phe-
nomenon whereby basic alignment properties of a language can best be described
by appealing to semantic factors, rather than syntactic ones (typically syntactic
factors involving argument structure, a simplification of semantics). This has been
the claim, whether overt or otherwise, since Boas and Deloria (1941) described
the grammar of Dakota (Lakota); the kinds of factor that can contribute to
these morphosyntactically coded distinctions in semantics are well described in
Merlan (1985) and Mithun (1991). If we wish to claim that semantic alignment is
a linguistic type, and not simply an epiphenomenal feature that emerges in some
languages as the result of the ‘direct’ coding of semantic roles in the morphosyntax
of a language, we need to carefully examine both parts of the term: ‘semantic’
and ‘alignment’. In this chapter I address the issue of what ‘counts’ for alignment
purposes. This is not completely straightforward, since many of the tests that are
normally employed in evaluating alignment are syntactic in nature, and we are
explicitly interested in investigating alignment where syntax is not a direct factor
in determining alignment.

I shall not in the main be making any controversial claims, but simply following
the logical implications of a number of accepted assumptions about linguistic
structure. Apart from examining splits in the encoding of the S or a monovalent
clause—something that is essential (though not sufficient) to the consideration
of semantic alignment in a language—I shall also examine bivalent clauses to
find parallel instances of semantic information appearing in otherwise syntacti-
cally aligned coding strategies (see Beavers 2006 for a recent synthesis of these
kinds of alternations, which include the well-described conative alternations of English).

2.1.1 Splits in alignment

In the occasional rare language, all morphological and syntactic (constructional?) tests show the same grouping of eligible and ineligible, privileged and non-privileged arguments. While this is an ideal, one that is often propagated in undergraduate introductory classes, it is not the usual case for all languages, or even across the whole of one language. English, a well-known and very strongly and consistently nominative-accusative language, shows numerous instances of the S and A being treated similarly, and differently to the P; in (1)–(3) some of the most obvious and non-controversial of these are presented, being the data basic to a determination of alignment: case marking on NPs, agreement on the verb, and the position of the NPs with respect to each other and the verb. In all three cases the S patterns with the A, differently to the P, and so the language can be described as showing nominative-accusative alignment.¹

Alignment in English: nominative-accusative patterns

(1) pronominal case

A I saw them.
S I walked.
P They saw me.

(2) verbal agreement

A She sees them.
S She walks.
P They see her.

(3) position with respect to the verb

A I saw them.
S I walked.
P They saw me.

Even English, however, presents conflicting evidence: there is some data on minor construction types that does not show the near-universal nominative-accusative pattern in pivot choice. In (4) we can see that the preposition choice for the post-head appearance of arguments in nominalizations depends on an ergative/absolutive split, with As being marked by by, and Ss and Ps marked by of. In (5) the restrictions of prenominal past participle modification are shown. As are ineligible to appear in the construction, while Ps are eligible (the eaten glutton could be construed grammatically, but only if the glutton is the argument eaten, not the argument which carries out the eating). Examining the behaviour of Ss reveals a split along semantic grounds: an agentive S may not appear in this construction, while more patientive Ss do participate. Finally the phenomenon of Right-node raising, illustrated in (6), inevitably allows for gapping of a non-nominative entity, due to the preverbal position of any As and Ss in English.

¹ I use A, S, and P, following Comrie (1978), to refer to the arguments of monovalent and bivalent verbs.
(4) Preposition choice when the head precedes in a nominalized clause

A the investigation by the police
S the arrival of the guests
P the consumption of alcohol

(5) Eligibility for a nominal to be modified by a prenominal past participle verb

A the eaten glutton
Sa the sung choir
Sp the fallen leaves
P the beaten protesters

(6) Ellipsis in gapping constructions

A * __ washed the plates and I dried the dishes.
S * __ ran and I fell over.
P I washed __ and you dried the dishes.

Do the syntactic facts in (4)–(6) ‘count’ for the purposes of determining the alignment of the language? Given that linguists have discussed syntactic, as opposed to morphological, ‘ergativity’, we need to develop a theory of what ‘counts’, and what doesn’t, when evaluating pivots and alignment. Given the data in (4)–(6), can English be said to have ergative dependent marking, or ‘stative-active’ syntax? Why not? The answer lies in an examination of the ways in which we determine alignment, and an examination of the factors that separate semantic alignment from the more familiar syntactic alignment patterns such as ergative-absolutive and nominative-accusative.

2.1.2 Discerning alignment

There are a number of ways to define alignment in the grammar of a language, or in any sub-part of that grammar (see Siewierska 2003 for discussion of some of the complications arising when determining alignment). Foley (1993) shows that even in a single language, verbs may show different alignment in different parts of their paradigms, for the coding of the same arguments, and Siewierska (2003) follows this with a detailed study of a range of languages with complications in this area. Similarly, in different constructions we might find different patterns of alignment. In addition to verbal agreement, however, we also find case (or adposition) marking frequently used as an indicator or alignment, and, more rarely, the position of the arguments in the clause.

Equally importantly, and well-acknowledged in the literature, is the fact that morphological behaviour does not necessarily correlate with syntactic behaviour (see e.g. Chung 1978, Simpson 1991). In Table 2.1 I list a number of different ‘tests’ for morphological alignment, and a number of the more commonly applied tests used in determining the grammatical status of the arguments of a clause as subject, object, etc. I shall illustrate, and discuss, these different constructions below.

2 Such a theory is not wholly undeveloped; the work of Falk (2000, 2006) in particular addresses the question of evaluating different tests for ‘subjecthood’, and determining which component of grammar hosts them, and thus what they really test.
I shall introduce a cover term for the use of head-marking (agreement), dependent-marking (case or adposition use), and position (order of elements in the clause) when used to determine alignment: these are primary morphosyntactic coding devices. Their primacy is established by the fact that these three devices are used to some extent in all clause types. Examples of ‘morphological’ alignment determined by these different mechanisms are not hard to assemble. In (7)–(9) the zero (or optional) category is shown in brackets; thus in Russian there is agreement that marks the A or the S of the clause; P is not indicated on the verb. In Kabardian, on the other hand, all of A, S, and P are cross-referenced on the verb, with the S using the same set of prefixes as the P. Much has been written about the markedness relations that hold between different case or agreement patterns; all generalizations appear to be statistical, not universal. Thus, for instance, while an overt ergative case is overwhelmingly common in languages with ergative-absolutive alignment marked on the dependants of the clause, there are counterexamples in which absolutive is marked and ergative is unmarked (two such examples are Nias (Donohue and Brown 1999) and Tlapanec (Wichmann 2005b)).

Head marking

(7) Russian: A+S vs (P) agreement, showing nominative-accusative alignment; Enga: A+S vs P agreement; Bunak: (A+S) vs P agreement; Kabardian: A vs S+P agreement, showing ergative-absolutive alignment; Burmeso: (A) vs S+P agreement; Acehnese: A+Sa vs (P+Sp) agreement, showing semantic alignment; (numerous): A+Sa vs P+Sp agreement (see this volume).
Table 2.2. Alignment determined by different morphosyntactic coding devices

<table>
<thead>
<tr>
<th></th>
<th>Nom-Acc</th>
<th>Erg-Abs</th>
<th>Semantic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agreement</td>
<td>Swahili</td>
<td>Tz’utujil</td>
<td>Acchnese</td>
</tr>
<tr>
<td>Case</td>
<td>Yindjiharday</td>
<td>Jiwarli</td>
<td>Tsowa-Tush</td>
</tr>
<tr>
<td>Position</td>
<td>Thai</td>
<td>Paumari</td>
<td>Ambonese Malay</td>
</tr>
</tbody>
</table>

Dependent marking

(8) Yindjiharday: (A+S) vs P case marking, showing nominative-accusative alignment;
Japanese: A+S vs P case marking; Oirata or Oromo: A+S vs (P) case marking;
Jiwarli: A vs (S+P) case marking, showing ergative-absolutive alignment;
Nias or Tlapanec: (A) vs S+P case marking;
Tsowa-Tush: A+Sa vs P+Sp case marking, showing semantic alignment.

Word order is less frequently cited as a determiner of alignment, but it is also relevant. The examples in (9) show simple

Word order

(9) Thai: AVP/SV order of NPs and V, showing nominative-accusative alignment;
Paumari: AVP/VS order of NPs and V, showing ergative-absolutive alignment (also Tepehua, and PVA/SV for Mangarrayi);
Ambonese Malay: AVP/SaV & VSp order of NPs and V, showing semantic alignment.

The information in (7)–(9) is summarized in Table 2.2. Here, unsurprisingly, we see that any of the morphosyntactic strategies may be used to mark any of the main alignment types.

Both nominative-accusative and ergative-absolutive alignments share the fact that they are syntactically determined: specifically, the determination of an ergative or accusative alignment depends on examining all, and only, A, S and P. In other words, the semantic differences between different kinds of Ss is neutralized in these languages. The label ‘syntactic alignment’, when used in opposition to ‘semantic alignment’, is intended to represent the fact that the morphosyntactic groupings are based on syntactic (specifically, argument structure) positions.

3 To a lesser important degree the semantic differences between different A and Ps are also neutralized when determining syntactic alignment systems, since by definition an A and a P represent a basic, idealized notion. As we shall see in sections 2.6 and 2.7, even languages with predominantly syntactic alignment show degrees of semantic ‘interference’ in their morphosyntactic choices.
When we examine syntactic, rather than morphological, alignment, examples of grammatical constructions having 'pivots' that are more syntactically privileged than non-pivots are similarly easy to find in the literature. Some examples are shown below.

2.1.2.1 Adjunct clauses In English a [([While/After/Before] V-ing . . .)] clause is restricted to having an ellipsed S or A which must be coreferential with the S or A of the main clause. In (10) we can see that the ellipsed argument of the adjunct clause is an A, while in (11) it is an S. (12) is an attempt to construct an adjunct clause with an ellipsed P that is coreferent to an argument in the main clause, and it is not grammatical.

(10) *She tickled her niece [ while Ö1 preparing her summary ].
(11) She tickled her niece [ while Ö1 relaxing ].
(12) *She tickled her niece, [ while she was amusing Ö1 ].

Other languages show different restrictions on coreference into adjunct clauses, but the point is that this is a syntactic construction that can have restrictions which can be interpreted as syntactic pivots.

2.1.2.2 Relative clauses In Dyirbal (Dixon 1972), a [ . . . V-ngu . . .] clause, which can roughly be described as marking non-complement subordination (including functions similar to relativization), is restricted to having an ellipsed S or P. This restriction can be illustrated with the textual examples shown as (13)–(15). The example in (13) shows a -ngu clause headed by an S, kuku. In (14) the verb wukal ‘give’ is affixed with the valency-decreasing -ngay to function as an antipassive, leaving the agent of the verb as the S of a monovalent predicate, and so eligible to head a -ngu clause. Finally in (15) we can see that a relative clauses headed by wandal ‘hang up (tr)’ allows the P to head the relative clause. (Examples from Dixon 1972: 386, 388, 371.)

Dyirbal

(13) bura-n kinyanbayji kuku [ nyina-ngu ],
see-NFUT DEM-III-DOWN.HILL-SHORT.WAY owl sit-REL
[ mulku-mba-ngu ].
NOISE-INTR.VRBLZR-REL
‘I saw a mopoke owl sitting down there [in the grass] making a noise.’

(14) ngali, kajin-du baka balakarranya, [ kudu wuka-l-nga-ngu ],
IDU stick-INTR dig TWO-ACC NEG give-ANTIPASS-REL
‘We’ll dig with a yamstick around our two brothers [in the tree], who wouldn’t give [us any big starlings].’

Dyirbal examples have been retranscribed in a version of a general Australianist orthography. The velar stop is shown as <k>, to avoid confusion between sequences of /n/+/g/ and the velar nasal /ŋ/, shown as <ng> (voicing is not contrastive in Dyirbal).
Looking up, [they] saw a baby hanging there in a cradle; the two of them took [the baby and cradle] down. (literally, ‘a baby that (someone) had hung up there in a cradle’)

Examples of S,A restrictions in relative clauses are easy to find in many other languages, and will not be illustrated here (see Keenan and Comrie 1977 and much subsequent work).

2.1.2.3 Conjoined clauses If there are restrictions on gapping across coordinated clause boundaries, it seems to be universal that the pivot is, or at least allows, S or A gapping. English is one language that follows this restriction, as shown in (16).

(16) She, tickled her niece, and then Øi, prepared her summary.

Even languages with predominantly ergative morphology can show a preference for S,A pivots in coordinate structures. Citing Dyirbal again, we can note the following example (Dixon 1972: 73), showing that the A of the first clause is coreference with the A of the second clause (most cases of coordinate gapping in texts follow this pattern).

(17) a. Ngaja bala yuku yuba-n,
    1sg:erg iii:abs stick put.down-nfut
    ‘I put down the stick,…

b. balan jukumbil jilwa-n.
    ii:abs woman kick-nfut
    [and] kicked the woman.’

Other languages may monitor the conditions on coreference across the clause boundary, as in the Skou (Skou family, north-central New Guinea) examples in (18) and (19). In (18) the A of the first clause is coreferential with the S of the second clause, and the instrumental =pa is used at the end of the first clause to signal this relationship. In (19), in which the A and the S are not coreferential, the obviative is used. Importantly, the distinction monitored is whether the S or A in each of the clauses is the same or different vis-à-vis the S or A in the following clause.

5 Claims about S,P privileges in languages such as Dyirbal do not stand up to an examination of textual, rather than elicited, material, where instances of pivots involving a P almost exclusively involve resultative clauses. (17) is one example that defies characterization as being a clause in a language with an S,P pivot for coordination.
2.1.2.4 Floating quantifiers The reference of a V-adjoined quantifier in Japanese is restricted to the S or P of the clause (Donohue 2004a). (20) shows that an S may be the referent of a V-adjoined floating quantifier, while (21) shows that, given a clause with an A and a P, an A cannot be interpreted as modified by a floating quantifier, while a P can.

Japanese
(20) Gakusei ga niwa de san-nin asob-ta.
\text{student nom garden instr three-clf play-past}
‘Three students played in the garden.’

(21) Gakusei ga kodomo o niwa de san-nin mi-ta.
\text{student nom child acc garden instr three-clf see-past}
‘The students saw three children in the garden.’
\ast \ ‘Three students saw the children in the garden.’

Floating quantifiers with different restrictions can be found in English (in which the restriction is to an S,A argument) and many other languages (see Sportiche 1988, Shlonsky 1991, Aoun and Li 1993, Kuno et al. 1999).

2.1.2.5 Imperatives The addressee of an imperative in English can be any S or A. In (22) and (23) we see examples of and S and an A being omitted when coreferential with the addressee of an imperative construction. (24) shows that a P may not be so omitted.

(22) \_ \_ Go to the shops!
(23) \_ \_ Finish your homework!
(24) \ast \text{Daddy tickle \_\_!}

Most language have an [agent] restriction as well, or instead, but English is surprisingly free in terms of causative construal. The requirement on agentivity in many languages makes this construction sound a lot like a semantically restricted construction; but now and then, as in English, we find clear evidence that the construction has been syntactically restructured. In (25) the S of a verb that is clearly lexically
specified as non-agentive is gapped in the imperative construction, and in (26) the non-agentive S of a passive construction is gapped. Such imperatives are not possible in languages with a more semantically determined imperative construction, such as Tukang Besi (Donohue 1999a).

(25) When you get the cue, __ fall over in a heap!
(26) __ Be seen at Saturday’s big opening spectacular!

2.1.2.6 Secondary predication Secondary predication is often restricted to core arguments, or a particular set of core arguments. Typically, resultatives are restricted to predicating Ps, while depictives can have a wider scope. In (27) the secondary predicate senseless is predicated off the P of knock, not of the A, such that the only interpretation possible for the sentence is that the idiot, not she, becomes senseless as a result of the event in the main clause.

(27) She knocked the idiot senseless.

2.1.3 Splits in subject coding that are not semantic alignment

In many cases a language can exhibit a split in the behaviour of its subjects without us wanting to consider this to be an instance of semantic alignment. The lexical semantics of particular verbs is often the characteristic that allows them, or disallows them, to appear in particular constructions. In English we have already seen that prenominal modification by a past participle verb is only grammatical if the verb is non-agentive (see (5) and the discussion preceding it). This is a clear indication of the lexical semantics of the verb affecting the grammaticality of a construction; but it does not play a role in alignment, since the word order, case marking, and (lack of) agreement do not monitor this semantically regular split.

In Tukang Besi (Donohue 1996a, 1999a), only agentive verbs may occur with the accompaniment applicative -ngkene, as can be seen in (28) and (29). In (28) the accompaniers are marked as obliques in PPs headed by kene, and both agentive and non-agentive predicates are possible. In (29) the accompaniers are introduced as the objects of complex predicates composed of a verb and the accompaniment applicative (which is almost certainly related to the preposition kene historically, but which has a distinct set of synchronic restrictions). In (29) we see that the clause with an agentive base verb, tinti, is grammatical, but one with a non-agentive base verb such as buti is not.

Tukang Besi

(28) a. No-tinti kene iai=su.
   3REAL-run with younger.sibling=1SG.GEN
   ’He ran with my little sister.’
b. No-buti kene iai=su.
   3real-fall with younger.sibling=1sg.gen
   ‘He fell with my little sister.’

(29) a. No-tinti-ngkene te iai=su.
   3real-run-accom core younger.sibling=1sg.gen
   ‘He ran with my little sister.’

b. *No-buti-ngkene te iai=su.
   3real-fall-accom core younger.sibling=1sg.gen
   ‘He fell with my little sister.’

Agreement on the verb is identical in the sentences in (28) and (29), as is word order and case marking. We must conclude that, while the lexical semantics of the verbs certainly do intrude into the grammar of the language, they do not intrude into the alignment system of the language. On the other hand, the fact that only experiential verbs allow for the genitive encoding of their subjects is evidence that there is a dative-marking pattern in Tukang Besi (see Nichols, this volume), albeit a minority pattern.

Tukang Besi

(30) a. No-tinti.
   3real-run
   ‘He ran.’

b. No-buti.
   3real-run
   ‘He fell.’

c. No-mo’aro.
   3real-hungry
   ‘He’s hungry.’

(31) a. *Tinti=no.
   run=3gen
   ‘He ran.’

b. *buti=no.
   run=3gen
   ‘He fell.’

c. Mo’aro=no.
   hungry=3gen
   ‘He’s hungry.’

By this restriction to the definition of semantic alignment, with the emphasis being on the aligning of the arguments of the verb in terms of head-marking, dependent-marking, or the relative order of these elements, we exclude the ‘classical’ unergative/unaccusative splits that are found in many western European languages, such as Dutch, in which auxiliary choice is the only overt marker of class membership.

Dutch

(32) a. Ik heb haar gezien.
   1sg.nom have 3sg.f.acc see.ptcpl
   ‘I saw her.’

b. Ik heb daar gesprongen.
   1sg.nom have there jump.ptcpl
   ‘I jumped there.’

c. Ik ben daar naar toe gegaan.
   1sg.nom be there to go.ptcpl
   ‘I went there.’
Splits in alignment which do not involve conditioning by the lexical semantics of the predicate are also outside the purview of the label 'semantic alignment.' As an example of this, consider the following data from Chamorro (Austronesian, Guam; Topping 1973, Gibson 1992, Chung 1998, and Chung 2003 for discussion of the absolutive enclitics). Here, prefixal agreement in realis clauses is found only for ergative arguments, while in irrealis clauses both ergative and nominative arguments trigger agreement. Since all verbs participate in this alternation, the split in marking of any given S is not dependent on the lexical semantics of the verb, but on the independently imposed TAM category of the clause. (Examples of languages in which, for example, verbs that are inherently specified as being stative are marked differently to those that are inherently eventive would be treated as instances of semantic alignment. The crucial difference is that it is lexical aspect that governs the split.)

Chamorro

(33) a. Un-tugi’i i kätta.
   2sg-write 1 letter
   ‘You wrote the letter.’

b. Pära un-tugi’i kätta.
   will 2sg-write 1 letter
   ‘You will write the letter.’

(34) a. H<un>_anao-kao pära si Rita.
   go<si>-2sg.abs to si Rita
   ‘You went to Rita.’

b. Un-hanao pära si Rita.
   2sg-go to si Rita
   ‘You will go to Rita.’

Semantic alignment, then, should be seen as a subset of all kinds of split in the marking of Ss, rather than simply being a replacement label for the same phenomenon.

Another thing that is not here considered to be an instance of semantic alignment is the contrast in the morphosyntactic encoding of arguments of predicates of different syntactic categories. This is most commonly found in the contrast between subjects of (monovalent) verbal predicates and the 'subjects' of non-verbal predicates. This topic is pursued at length in Danielsen and Granadillo (this volume), but a few examples of languages which show differences along this parameter are given below. In Tukang Besi the elements of a clause follow the verb; postverbal objects and adjuncts have been seen in (28) and (29), and (35) shows a postverbal subject. In a nonverbal clause, however, the subject precedes the predicate, (36) (subjects are shown in bold in (35) and (36)). There is clearly a split, based on the syntactic category of the predicate, in terms of where the subject of the clause appears. Since the split depends on the syntactic category of the predicate, we cannot consider it to be an instance of semantic alignment.

Tukang Besi

(35) No-initi na iai=su.
    3real-tun nom younger.sibling=1sg.gen
    ‘My younger sister ran.’
Many languages which allow nonverbal clauses to appear without copular verbs show an epiphenomenal distinction between predicates: verbal predicates may take (verbal) agreement morphology, while nonverbal clauses may not. While this does distinguish two types of clause, each of which consists only of a subject and a predicate, it does so on the basis of the the morphological possibilities that are available to one syntactic category (verbs can take agreement) and which are not, in this language, available to other syntactic categories (nouns and adjectives do not take agreement). We can see this in many languages, and it is illustrated here with data from Skou. Examples of monovalent and bivalent verbal clauses have already been seen in (18) and (19); importantly, we noted that the S of A of a clause shows agreement by means of nominative prefixes (and clitics), while the P does not. In (37) we can see a simple (monovalent) verbal clause, in which the person and number features of the S are multiply indexed on the verb. In (38) the predicate is not verbal, but nominal, and there is no agreement. In other words, unlike the S of (37), which aligns with the A for the purposes of agreement coding, the ‘S’ of (38) shows the same (lack of) marking on the verb that characterises a P. This might be thought of as being an instance of a split in the coding of the S, but the fact that we are examining different clause types, split along syntactic category lines, means that we cannot consider this to be an instance of semantic alignment. Further, rather than being interpreted as a difference in the morphosyntax of verbal and nonverbal clauses, this should be interpreted as one of the defining contrasts that can be found between verbs and nouns.

**Skou**

(37) \( Pe=\text{ueme}=\text{a} \quad \text{pe}=\text{ta}<w>/\text{ing}. \)

\( 3\text{sg.f}=\text{woman}=\text{the} \quad 3\text{sg.f}=\text{sit}<3\text{sg.gf}> \)

‘The woman sat down.’

(38) \( Pe=\text{ueme}=\text{ing a} \quad \text{kurù}. \)

\( 3\text{sg.f}=\text{woman}=\text{the teacher} \)

‘The woman is a teacher.’

In One (Toricelli, northern New Guinea) we can see a difference between verbal and nonverbal clauses in terms of case marking. The case \( sa \) is optional in both verbal and nonverbal clauses, and marks a contrastive topic when it appears, as in (39) and (40). In negated clauses \( sa \) remains optional in verbal clauses, but is obligatory in nonverbal clauses, as shown in (41) and (42). In addition to the polarity distinctions, there is a difference in grammaticality based on the syntactic category of the predicate.
This section has presented a number of ways in which there can be a split in the marking of arguments that does not correspond to what we are calling evidence of semantic alignment: the split can be based on the behaviour of the verbs with respect to affixation, reflecting the lexical semantics of the predicate, but not reflecting any agreement or case, or word order differences. The split might be based on factors that are external to the verb, such as TAM values for the clause, in which case this is again not semantic alignment. Finally, and perhaps most straightforwardly, when the split makes reference to syntactic categories, such as word class, then we do not consider the split to be an instance of semantic alignment.

2.2 Semantic alignment: morphological and ordered

Examples of semantic alignment systems are not hard to find, though their reporting has largely concentrated on America, isolated pockets of Eurasia (see Mithun, this volume, and Nichols, this volume). Notable exceptions to this trend in reporting can be seen in Donohue (2004b) and Foley (2005), Tsukida (this volume), and the well-documented case of Acehnese (Durie 1985, 1988, Foley 2005, Klamer, this volume).

For Acehnese, an Austronesian language from far western Indonesia, Durie (1988) reports that there are no reasons to group Sa and Sp together in terms of their behaviour in discourse. Morphologically, we can distinguish these two monovalent arguments in terms of their agreement: and Sa will always show agreement by proclitic, while with an Sp agreement is optional, and by enclitic (Durie 1985: 181); the semantic factor that distinguishes Sa from Sp is volitionality.6

Acehnese

(43) Jih ka=ji=jak.
he INCHOATIVE=3=go
'He has gone.'

(44) Gopnyan ka=sakct=geuh.
he INCHOATIVE=sick=3
'He is sick.'

6 It is likely that the alignment system in Acehnese was diffused from the Austro-Asiatic languages of the area (see e.g. Semelai (Kruse 2004)), which probably (given their distribution in peninsular Malaysia and the Nicobar islands, as well as the unusual, and very Austro-Asiatic-like, phonologies of the Sumatran Austronesian languages, and the considerable number of non-Austronesian toponyms in the area) represent a pre-Austronesian substrate across Sumatra.
To provide some further examples from the southwest Pacific, we can see that in these languages semantic alignment can be coded through agreement, through case marking, or through NP position in the clause, just as can the more common syntactic alignments. In (45)–(47) we see that in Galela verbal prefixes split along semantic grounds (Shelden 1991; see Holton, this volume, for more discussion of the North Halmahera group of languages to which Galela belongs), with wo-marking an agentive argument, whether it is the A of a bivalent clause or the S of a monovalent clause. Mi-, indexing a 3rd person feminine argument, indicates non-agentivity, regardless of whether that non-agentivity is assigned to an S or a P.

Galela

(45) **Wo-mi-sasano.**
   3SG.M.A-3SG.F.P-ask
   '(Aweng) questioned her.'

(46) **Wo-mau....**
   3SG.M.A-want
   'He wants (to go).'  

(47) **Mi-sirangu.**
   3SG.F.P-nose.runs
   'She has a runny nose.'

In Waris (Brown 1988), from New Guinea, the case marking found on As is also used for agentive Ss. The case marker that marks recipients of verbs of transfer is also used for some Ps, if they are animate or relatively unaffected inanimate Ps. It is also used for human affected Ss (see Nichols, this volume, for discussion of this type of pattern).

Waris

(48) **K̲a̲-va ye-m hévakomandha-v.**
   1-TOP 2-DAT KILL-PRES
   'I kill you.'
   (−m only appears on animate or less-affected inanimate Ps)

(49) **K̲a̲-va mongla-na pró-na.** (50) **He-m daha-v.**
   1-TOP foot-gen come-PAST 3-DAT die-PRES
   'I came by foot.'
   'He is dying.'

Ambonese Malay, an eastern variety of Malay spoken in Maluku province, Indonesia, shows a split based on agentivity and coded by the position of the NP relative to the verb: in bivalent clauses AVP order is found, and in monovalent clauses an agentive S precedes the verb, while a non-agentive S follows (these ideals are frequently complicated by arguments appearing in sentence-initial position when topical).
Ambonese Malay

(51) Dorang cari betang konco.
3pl. search.for my friend
‘They’re looking for my friend.’

(52) Betang konco su-bajaang.
my friend perf-walk
‘My friend walked away.’

(53) Su-jato betang konco.
perf-fall my friend
‘My friend has fallen over.’

In Palu’e (Austronesian, southern Indonesia) we find a variant of the Ambonese Malay system seen in (9) and (51)–(53). Monovalent clauses always allow an SV order, as in (54); bivalent clauses shown an AVP order. In addition to the subject-initial order, non-agentive predicates also allow their S to follow the verb, as in (56); this is not possible for agentive predicates (see Donohue 2005a for arguments concerning clause structure which imply that we should not treat (54b) as an instance of topicalization). The Sa is always coded preverbally, in the same position as an A, while the Sp may appear postverbally, mimicking a P, or preverbally.

Palu’e

(54) a. Ia phana-’u.
3sg go-perf
‘She’s gone.’
b. Ia molu-’u.
3sg fall-perf
‘She’s fallen over.’

(55) Kami phote nio.
1pl.ex pick.coconut coconut
‘We picked some coconuts.’

(56) a. ∗Phana ia-’u.
go 3sg-perf
‘She’s gone.’
b. Molu ia-’u.
fell 3sg-perf
‘She’s fallen over.’

The languages presented in this section have all shown semantic alignment in one way or another, via agreement, via case marking, or via ordering in different positions with respect to the verb. In the following section we shall examine ways in which alignment might be split and yet not be considered to be an instance of semantic alignment.

2.3 Morphological splits in marked alignment

We have seen that different subsystems of the grammar can take different alignment, in terms of the morphology used. The appearance of such splits is not uniformly distributed, however: examining languages with splits, we find that
while morphological splits are common for grammatical systems with overtly ergative components in their grammar, splits involving an S,A component tend to be restricted to agreement being realized with a nominative grouping, and not having dependent-marked nominative categories.

There are no reported cases of word order varying to show a split (though see e.g. Payne 1994a), but variation in word order correlating with degrees of transitivity is a common enough phenomenon. In Puare, a Macro-Skou language from north-central New Guinea, SOV is the dominant clausal order, but there are systematic differences between preverbal and postverbal objects: the postverbal position is only used with indefinite or nonspecific objects, while the preverbal position is unrestricted. This can be seen in (57). A bare NP object may appear either preverbally or postverbally, with a corresponding difference in interpretation. When the object is modified with a demonstrative, and thus is overtly specified as being definite, the postverbal coding option is ungrammatical.7

Puare

(57) a. N-ae[e n-ua]a |ku.
   1sg-go 1sg-search.for egg
   'I went to look for eggs.'

b. N-ae[e ku n-ua]a
   1sg-go egg 1sg-search.for
   'I went to look for the egg.'

c. "N-ae[e n-ua]a |ku pende.  
   1sg-go 1sg-search.for egg that
   'I went to look for that egg.'

d. N-ae[e ku pende n-ua]a.
   1sg-go egg that 1sg-search.for
   'I went to look for that egg.'

A frequently noted characteristic of alignments other than the simple (at least, statistically more frequent) nominative-accusative one is the presence of a split in the morphosyntax of the language, such that one part of the language shows one alignment, while another area of morphology or syntax has a different pattern. Much has been written about the areas of morphosyntax in which these splits can occur, and the motivations for these splits. Here I shall present some examples of alignment splits in general, and of those involving semantic alignment in particular.

2.3.1 Head vs. dependent

In Warlpiri (Pama-Nyungan, Australia) the morphology on the head of the clause, the auxiliary, shown a nominative-accusative alignment, while dependents are

7 Objects which are inherently specific, such as pronouns, are similarly restricted to preverbal position.
all marked on an ergative-absolutive basis. We can see that the presence of the ergative -rlu on ngarrkajarra in the bivalent clause shown in (58) does not affect the ability of the nominative enclitic pala to appear on the auxiliary showing agreement for this argument, just as is found with ngarrkajarra in the monovalent clause seen in (59).

Warlpiri

Man-DU-ERG PRES-3DU.NOM-3PL.ACC kangaroo-PL see-NONPAST
‘The two men see the several kangaroos.’

(59) Ngarrka-jarra ka-pala parnka-mi.
man-DU PRES-3DU.NOM run-NONPAST
‘The two men are running.’

The syntax of Warlpiri depends on the S,A vs. P opposition much more than it does on the case-marked S,P vs. A distinction. In addition to the verbal agreement data seen above, data from controlled clauses indicates that the syntactic pivots in this language ignore the ergative case marking (see 2.4.1).

2.3.2 Main versus subordinate

Mam (Mayan, Guatemala; England 1983) has an ergative-absolutive system of agreement in main clauses, and a neutral system in subordinate clauses. Other related languages use the agreement morphemes in a nominative-accusative pattern in subordinate clauses (England 1983: 262). Table 2.3 shows the functions of verbal agreement markers in main and dependent clauses in three languages. In all three the main clause shows an uncomplicated ergative-absolutive alignment. In Ixil, dependent clauses show a nominative-accusative alignment, with
the ‘ergative’ prefixes being used to cross-reference both Ss and As. In Mam the ‘ergative’ prefixes are used for any core arguments, regardless of their syntactic role, thus instantiating a neutral alignment pattern. Aguacatec has two different types of dependent clause (determined by alignment on the verb), one which patterns like Ixil dependent clause, and one which patterns like Mam dependent clauses.

Examples from Mam illustrating this split can be seen in (60)–(63). In (60) we can see the absolutive and ergative agreement prefixes for a bivalent clause. In (61) the main clause is monovalent, and the S of that clause is indicated with absolutive agreement. The subordinate clause is also monovalent but, by contrast, shows agreement by use of the same ergative prefix that was seen in (60) cross-referencing the A.

Mam

(60) Ma ch-ok t-b’iyo-7n Cheep kab’ xiinaq.
REC 3PL.ABS-DIR 3SG.ERG-HIT-DIRECTION.SUFFIX José two man
‘José hit two men.’

(61) N-chi ooq [ t-poon ky-txuu ].
PROG-3PL.ABS CRY 3SG.ERG-ARRIVE 3PL-mother
‘They were crying when their mother arrived.’

In (62) the relative clause contains a passive verb, and the single argument of that verb is indexed with the ergative prefixes. (63) shows the use of two sets of ergative prefixes when there are two arguments in the dependent clause, making it clear that there is complete neutralization in terms of head-marking morphology.

(62) O tzaalaj xjaal t-i7j t-paa [ aj
PAST 3SG.ABS:be.happy person 3SG-about 3SG-bag when
 t-kan-eet prim-x ].
3SG.ERG-find-PASS early-ENCL
‘The person was happy about his bag when it was found early.’

(63) Ok go tzaalaj-al [ ok t-q-il u7j t-e
POT 1PL.ABS be.happy-POT when 3SG.ERG-1PL.ERG-see book 3SG-POS
word 3SG-POS Ixtahuacán
yool t-e lytzal].
‘We will be happy when we see the Ixtahuacán dictionary.’

2.3.3 TAM-dependent
Sindhi (Indo-European, Pakistan and India) (together with many other languages of the area) shows an ergative case-marking pattern only in perfective or past clauses. In (64) we see that the subject of the unrealized clause takes the regular nominative form, while in (65) the subject appears in the oblique form, which is available only for As (amongst the core arguments).
2.3.4 Person-dependent

Since Silverstein (1976), many authors have commented on the tendency for more highly animate arguments to show a nominative-accusative alignment, while ergative alignment is associated with less animate arguments. This can be manifested in many ways: the appearance of (nominative-accusative) case marking only on (highly animate) pronouns in English, compared to the unmarked NPs elsewhere. Table 2.4 shows the case marking for core arguments found in Wik-Mungkan (Pama-Nyungan, northeast Australia) on singular pronouns, and on a representative common noun. While the pronouns show a nominative-accusative distinction, the nouns are inflected with an ergative-absolutive alignment (using the same case suffix, -ang, in each case; similar patterns are reported for Kala Lagaw Ya in Comrie 1981). This pattern of having a split in case marking according to animacy is common across a wide selection of Pama-Nyungan languages of Australia.

The animacy hierarchy is also realized, indirectly, in the tendential appearance of ergative marking on lower-animate or pragmatically focused subjects in many languages of the New Guinea highlands (Donohue and Donohue 1997, Donohue 2005b).

2.3.5 Person- (and tense-) dependent

Iha (West Bomberai (non-Austronesian), Indonesia; Donohue and Brown 1999) shows a split in case marking on local pronouns. In past tenses the local singular...
Semantic alignment systems: what’s what, and what’s not

Table 2.5. Case distinctions in Iha past tenses

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>Sa</th>
<th>Sp</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>on</td>
<td>on</td>
<td>on/ni</td>
<td>ni</td>
</tr>
<tr>
<td>2sg</td>
<td>ko</td>
<td>ko</td>
<td>ko/ki</td>
<td>ki</td>
</tr>
<tr>
<td>3sg</td>
<td>mi</td>
<td>mi</td>
<td>mi</td>
<td>ndo</td>
</tr>
<tr>
<td>1pl ex: mbi, 1pl in: in, 2pl: ki, 3pl: mi (invariant)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Common nouns kabágat kabágat kabágat kabágat

* Bold line indicates boundary of optionality.

pronouns, shown in Table 2.5, can be used in a way that is compatible with a semantic alignment analysis. The basic divisions are shown in Table 2.5; note that common nouns, illustrated here with kabágat ‘tree kangaroo’, as well as the plural pronouns, are invariant in form. Of the singular pronouns, the 3rd person has an accusative form distinct from the nominative. The local persons have a distinct accusative, and this accusative may be used for non-agentive Ss in non-future tenses. Unlike the use of the accusative to mark a P, the use of the accusative for an Sp is always optional.

In (66)–(69) we can see that the choice of mi vs. ndo for 3rd person singular pronominals is not random but reflects a nominative-accusative alignment (mi is also used for 3rd person plural reference, in which situation there are no case distinctions).

Iha: 3rd person A and P coding

(66)  
Mi kalípan ngbréhe-bidya.  
3 mat weave.mat-PRES,3  
‘He/She is weaving a mat.’

(67)  
∗ Ndo kalípan ngbréhe-bidya.  
3SG,P mat weave.mat-PRES,3  
‘He/She is weaving a mat.’

(68)  
Kpyémbot on ndo kpáke-bon.  
yesterday 1SG 3SG,P fight-PAST,1SG  
‘I fought her/him yesterday.’

(69)  
Kpyémbot on mi kpáke-bon.  
yesterday 1SG 3(PL only; "3SG") fight-PAST,1SG  
‘I fought them(‘her/him) yesterday.’

When we examine monovalent verbs in detail, we see an interesting pattern. With agentive verbs the nominative set of pronouns must be used, as expected. With non-agentive verbs either the nominative or the accusative pronouns may be used,
when the subject is 1st or 2nd person. Third person subjects may only use the nominative pronouns.

Agentive verb, past

(70) Kebér on on-ma kpéh néγak ha-wahá-γe.
    just.then 1SG 1SG-POSS village towards climb-go-IRR
    ‘I wanted to go back up to my village just then.’

(71) * Kebér ni onma kpéh néγak havaháγe.

Non-agentive verb, past

(72) Kpyémboṭ on mgbahrùrmbon.
    yesterday 1SG fall-PAST.1SG
    ‘Yesterday I fell over.’

(73) Kpyémboṭ ni mgbahrùrmbon.
    yesterday 1SG.P fall-PAST.1SG
    ‘Yesterday I fell over.’

(74) Kpyémboṭ mi mgbahrùr-mbih.
    yesterday 3 fall-PAST.3
    ‘Yesterday s/he fell over.’

(75) *Kpyémboṭ ndo mgbahrùr-mbih.
    yesterday 3SG.P fall-PAST.3
    ‘Yesterday s/he fell over.’

When a non-agentive verb is used in a non-past tense, then the pronoun choice is once again fixed, with all persons being coded by the nominative pronouns.

Non-agentive verb, non-past

(76) On mgbahrùrnten.
    1SG fall-FUT.1SG
    ‘I’m going to fall over.’

(77) * Ni mgbahrùrnten.
    1SG.P fall-FUT.1SG
    ‘I’m going to fall over.’

This section has shown us some of the ways in which morphology can be less than straightforward. In the following section I shall elaborate on some of the syntactic ways in which alignment can be determined.

2.4 Constructions, pivots, and alignment

Often a language that has an ergative morphological system will show a different alignment in terms of syntax (the opposite, morphological accusativity and syntactic ergativity, is also, but rarely, attested).
Semantic alignment systems: what's what, and what’s not

Commonly attested dimensions of morphology/syntax splits include:

- Coordination almost always (universally?) selects an S,A pivot as the unmarked case.  
- Constructions that are close to the verb syntactically or semantically (such as suppletive verb stems, resultatives, or in some cases floating quantifiers) tend to have an S,P pivot.

Other preferential pivots can be established for various constructions, and will not be enumerated here. I shall illustrate, or at least mention, some of the more common splits that can be found.

2.4.1 Ergative morphology; nominative pivots

This is perhaps the most common of alignment splits. We have already seen an example of this in Warlpiri, which has ergative case, but in which intraclausal syntax consistently marks S,A as being distinct from P. When we examine the syntax of the language, it becomes clear that the S,A grouping is the one that is relevant for syntactic purposes. In (78) we can see that the subordinate verb uses the suffix -karra to indicate coreference of the A of the subordinate clause with the S of the main clause; the person whistling is coreferent with the person trimming the boomerang. (79) shows that the same suffix is used when the main clause contains an A that is coreferential with the A of the subordinate clause, and (80) shows that when the main clause is monovalent the same -karra suffix is used. If the S or A of the subordinate clause is coreferent with the P of the main clause, a different suffix, -kurra (identical to the allative suffix), must be used, as shown in (81). Finally, when the S or A of the subordinate clause is not related to any of the arguments of the main clause, the suffix -rlarni is used. This clearly establishes a contrast between the grouping of S and A on the one hand and P on the other.

Warlpiri

(78) *Ngarrka ka wirrapirli-mi karli jarni-mlinja-karra.*

man AUX whistle-NPST boomerang trim-INF-PROX

‘The man is whistling, while trimming the boomerang.’

(79) *Ngarrka-ngku ka purlapa yunpa-ri karli jarni-mlinja-karra-rlu.*

man-ERG AUX corroboree sing-NPST boomerang trim-INF-PROX-ERG

‘The man is singing a corroboree, while trimming the boomerang.’

---

8 Claims have been made about S,P-pivot coordination in various languages. Closer examination of textual materials reveals that there is, at best, a lack of any particular preference, rather than a preference for, or a restriction to, a pivot that targets a non-S,A category.
Mark Donohue

child-erg AUX-RLA water-DAT search-dig-NPST speak-INF-PROX-ERG
‘The child is digging for water, while speaking.’

(81) Ngarrka-ngku marlu pantu-rnu, marna nga-rninja-karra.
man-ERG kangaroo spear-PST grass eat-INF-OBV:OBJ
‘The man speared the kangaroo while it was eating grass.’

(82) Ngarrka-ngku ka karli jarnti-rni, kurdu-ku maliki
man-ERG aux boomerang trim-NPST child-DAT dog
wajili-pi-nja-rlarni.
chase-VERB-INF-OBV
‘The man is trimming a boomerang while the child is chasing the dog.’

These data were drawn from Hale (1982), where there is a more detailed discussion of the issues raised here. Further analysis can be found in Simpson (1991).

2.4.2 Accusative morphology, absolutive pivots

This split is rarely attested, but can be found. Oirata, from southeastern Indonesia (de Josselin de Jong 1937, Donohue and Brown 1999), shows nominative-accusative morphological alignment, as determined by the case choices on pronouns. Similarly, Oirata shows a semantic alignment in the switch-reference system (with some tendency towards event reference—see the discussion of Central Pomo below). By contrast, relativization is restricted to the S or P argument.

The basic morphosyntax can be seen in (83)–(86). Verbs show no agreement, but nominative is a marked category on pronouns.

Oirata

(83) In-te ee asi.
1pl.excl-nom 2SG.POLITE see
‘We saw you.’

(84) Ee-te in asi-ho.
2SG.POLITE-NOM 1PL.EXCL see-NEG
‘You didn’t see us.’

(85) An-te ete na’a ipa.
1sg-nom roof obI fall
‘I fell off the roof.’

(86) In-te Ahum na’a ma’u.
1pl.excl-nom Ambon obI come
‘We arrived from Ambon.’

The switch-reference morphology also monitors S/A in the first clause being coreferential or not to an (A+S) or (S+P) in the second clause. In (87) we see that the coreference of the A of the first clause with the S of the second means that the same-reference suffix -le is used. In (88), by contrast, it is the P of the first clause that is coreferent with the ellipsed S of the second clause. Since there is no identity of S or A with S or A, the different-reference suffix -le must be used.
Note that (88), which features the agentive verb tipare 'run' in the second clause, cannot be interpreted as showing the A of the first clause coreferent with the S of the second clause at a different point in time; this becomes relevant with the discussion of (90).

(87)  In-te  ihar asi-le  ____  lalare.  
      1PL.EXCL-NOM  dog  see-SAME   walk  
      'We, saw a dog and then Øi walked (away).'

(88)  In-te  ihar asi-to  ____  tipare.  
      1PL.EXCL-NOM  dog  see-DIFFERENT  flee  
      'We, saw a dog and then Øi ran off.'
      * 'We, saw a dog and then later Øi ran away.'

In (89) the second clause contains a non-agentive verb, ipa 'fall', the subject of which is coreferent with the subject of the first clause; just as in (87), the non-final verb takes the same reference suffix -le. In (90) the (monovalent) verbs of the two clauses are identical to those seen in (89), and the different reference suffix, seen earlier in (88), is used to link the clauses. The difference between (89) and (90) is that the second clause in (90) must be interpreted as occurring at a different point in time, as an event unconnected to the climbing seen in the first clause. As seen in (88), this use of different reference markers to indicate a different temporal reference, rather than a different argument reference, is not possible when the verb of the second clause is agentive but only when it is non-agentive. We can see, then, that the syntax monitors the difference between Sa and Sp in the switch-reference system, in that an Sp may optionally be treated in the same way as a P for the purposes of selecting switch-reference marking. (91) shows a more 'canonical' use of the different reference suffix with bivalent clauses, with an interpretation that clearly indicates two separate events.

(89)  In-te  ete  i'a  iamo-le  ____  ipa.  
      1PL.EXCL-NOM  roof  on  climb-SAME  fall  
      'We, climbed onto the roof and straightaway Ø fell off.'

(90)  In-te  ete  i'a  iamo-to  ____  ipa.  
      1PL.EXCL-NOM  roof  on  climb-DIFFERENT  fall  
      'We, climbed up onto the roof and (after a while) (when we were up there) fell off.'

(91)  Ira  eme  modo  ina-to  ____  tutu.  
      water  take  child  give-DIFFERENT  drink  
      'Give the child some water to drink.'

Relative clauses are formed with the suffix -n on the verb, and optionally mark the non-head core argument in the possessive case, if pronominal (1SG.POSS: an).
In (92) and (93) we can see examples of a relative clause headed by a P or an S, respectively.

(92) \[ \text{NP } Ihar \ [ \text{an-te } asi-n] \] tipare.  
\text{dog} \quad \text{1SG-NOM see-REL flee}  
'The dog that I saw left.'

(93) \text{In-te} \quad \text{NP ihar [mara-n] } \text{asi}.  
\text{1PL-EXCL-NOM dog go-REL see}  
'We saw the dog that had left.'

In (94) and (96) we can see that a relative clause headed by an A is not possible; instead, strategies such as that seen in (95) and (97), employing switch-reference morphology, must be used. Note that the fact that switch-reference morphology must be used with subordinated, as well as coordinated, clauses means that the relative clause translation of (97) is not as unusual as it might seem at first glance.

(94) * \text{An-te} \quad \text{NP modo [ira tutu-n]} \text{asi}.  
\text{1SG-NOM child water drink-REL see}  
'I saw the child that had drunk the water.'

(95) \text{An-te} \quad \text{modo asi-to } \text{ira tutu}.  
\text{1SG-NOM child see-DIFFERENT water drink}  
'I saw the child and he/she drank the water.'  
'I saw the child drinking the water.'  
'I saw the child who was drinking the water.'

(96) * \text{Ihar [ (ani / an) } \text{asi-n]} \text{] mara}.  
\text{dog} \quad \text{1SG-ACC 1SG-POSS see-REL go}  
'The dog that saw me left.'

(97) \text{Ihar} \quad \text{ani asi-le } \text{mara}.  
\text{dog} \quad \text{1SG:ACC see-SAME go}  
'The dog saw me, and Ø left.'  
(or, to give a discoursally equally valid translation or two: 'Seeing me, the dog left', or equally, 'The dog that saw me left')

2.4.3 Semantically based coordination pivot?

McLendon (1978) describes the case marking and switch reference of Eastern Pomo as having a system of switch reference that shows semantic alignment. The difference in case marking for agentive and non-agentive participants can be seen in (98)–(100). Note particularly that the Sa of (98) is coded differently to the Sp of (99).
Semantic alignment systems: what’s what, and what’s not

Eastern Pomo

(98) Há- mi-pal š-a’k’a.
   1sg.a 3sg.m.p killed
   ’I killed him.’

(99) Há- wis-du.kiya.
   1sg.a going
   ’I’m going.’

(100) Wí c’e-xélka.
   1sg.p slipping
   ’I’m slipping.’

The switch reference system can be shown by the suffixes on the verbs of the first verb in each of (101)–(103). Here we can see that when the agentive argument of the first clause is coreferential with the agentive argument of the second clause, the same-reference suffix -y is used on the verb. By contrast, -qan is used in other instances, including those such as (103) in which the identity of the two S arguments is the same, but one is agentive and one is not.

(101) Há- kálulu-y, si-má- mérqaki-hí.
   1sg.a went.home-SAME went.to.bed
   ’I went home and then went to bed.’

(102) Há- kálulu-qan, ści-p’ mérqaki-hí.
   1sg.a went.home-DIFFERENT 3sg.m.a went.to.bed
   ’I went home and then he went to bed.’

(103) Há- xá- qákkí-qan, wi q’a-lál-ta-la.
   1sg.a take.bath-DIFFERENT 1sg.p got.sick
   ’I took a bath and then I got sick.’

This behaviour is not found in Central Pomo, which shows event-reference tracking, not (strictly) participant-reference tracking (Mithun 1993); the same analysis might well extend to Eastern Pomo, if it was examined in more detail. Lani, and probably a number of other languages of the Western Highlands of New Guinea, show evidence of switch-reference systems that are similar to that of Central Pomo, though while Central Pomo seems to favour same marking, this being the default and the different marking being used only when a list of possible same-qualifying factors are not met, the languages in New Guinea favour different marking.9 Skou, mentioned earlier (examples (17) and (18)), also shows a mixture of participant-tracking and event-tracking in its switch-reference system; it is no

9 The fact that in New Guinea the more eastern languages with switch reference tend more closely to monitor participant reference, and that these languages also show the greatest elaboration of switch-reference systems, means that we have an interesting cline in terms of the function of the switch-reference system, as well as its morphological exposition.
accident that Skou is located approximately midway between the eastern and western extremities of switch-reference systems in New Guinea.

### 2.5 Morphology, syntax, and semantic alignment phenomena

It should be clear from the preceding section that the fact that a language is morphologically aligned one way or another does not entail any categorial implications about overall alignment. This observation is not a new one; Li and Lang (1979), Van Valin (1981), and many authors since have reached the same conclusions. There are some strong tendencies, however, which are charted in Table 2.6.

At the same time, we must consider the fact that the label ‘non-nom/acc’ covers both the relatively common case of languages with ‘ergative-absolutive’ alignment and languages with semantic alignment. What are the criteria that lie behind semantic alignment? Since the term ‘semantic alignment’ is a new one (see Wichmann, this volume, for discussion of the rationale for the term), it is worth briefly examining the kinds of terminology that have been used, by different authors and in different frameworks, to describe the kinds of argument encoding we are considering here. Note that a number of the terms commonly used to describe the oppositions match descriptors employed by, for instance, Hopper and Thompson (1980), in describing the degrees of (semantic) transitivity in a clause, without restricting the sense to the argument of monovalent verbs. Table 2.7 presents the kinds of term that have been used, by different authors and from different traditions, on the left, with the 10 criteria discussed by Hopper and Thompson listed on the right. While not all of the columns on the left match categories on the right, the number of matches is too high to be random.

The fact that Hopper and Thompson assembled their list of criteria with the aim of unifying the description of monovalent and bivalent clauses is telling, and calls for the examination of bivalent clauses as well as the monovalent ones that we have been seeing so far, when discussing semantic alignment. This examination can be found in the following two sections.

#### 2.5.1 Bunak

The unity of many of these different parameters can be seen in Bunak, a non-Austronesian language of central Timor, in southern Indonesia and western Timor
Semantic alignment systems: what’s what, and what’s not

Table 2.7. Alternative terminologies/parameters of variation

<table>
<thead>
<tr>
<th>Labelled opposition</th>
<th>(cf. Hopper and Thompson 1980)</th>
</tr>
</thead>
<tbody>
<tr>
<td>active stative</td>
<td>Participants</td>
</tr>
<tr>
<td>Proto-Agent</td>
<td>Kinesis</td>
</tr>
<tr>
<td>eventive stative</td>
<td>Aspect</td>
</tr>
<tr>
<td>unergative unaccusative</td>
<td>Punctuality</td>
</tr>
<tr>
<td>controlled non-controlled/uncontrolled</td>
<td>Volitionality</td>
</tr>
<tr>
<td>Sa</td>
<td>Affirmation</td>
</tr>
<tr>
<td>Initial-1</td>
<td>Mode</td>
</tr>
<tr>
<td>agentive non-agentive</td>
<td>Agency</td>
</tr>
<tr>
<td>unaffected affected</td>
<td>Affect edness of O</td>
</tr>
<tr>
<td>external argument internal argument</td>
<td>Individuation of O …</td>
</tr>
<tr>
<td>Actor</td>
<td>Undergoer</td>
</tr>
</tbody>
</table>

Lorosa’e. Verbs show agreement with an accusative prefix, and this prefix marks person and, for 3rd persons, a distinction that has traditionally been described as one of gender, differentiating animate and inanimate. Examples of sentences that show this difference are given in (104) and (105).

Bunak: ‘animate’ vs. ‘inanimate’ based on human vs. non-human

(104) en tapol himo g-ukat.
    person fall DET.AN 3AN-lift
    ’(they) picked up the one who had fallen.’

(105) Neto r-on h-ukat loi ni’.
    1SG REFLEX-ARM 3INAN-lift good NEG
    ’I can’t lift my arm.’

The same morphological distinction is also used to show the difference between specific (= ‘animate’) and nonspecific objects (note that the inanimate prefix h- is not realized before a consonant-initial verb stem).

‘Animate’ vs. ‘inanimate’ based on specific vs. nonspecific

(106) Neto uwor g-ial gie.
    1SG vegetables 3AN-CARRY IRR
    ’I’m going to carry the vegetables.’

(107) Gereje g-ewen no eto uwor wit loi.
    church 3-FRONT OBL 2SG vegetables (3INAN-)BUY good
    ’In front of the church, you can buy vegetables.’

The same contrast is also used to show the kinds of difference that we saw with English prepositionally marked objects. The use of the ‘animate’ agreement set on
the verb implies a more agentive A, while the 'inanimate' set of prefixes has no such implication.

'Animate' vs. 'inanimate' based on agentive vs. non-agentive


1sg music 3an-hear 1sg music 3INAN-hear
'I listen to music.' 'I heard music.'

'Animate' vs. 'inanimate' based on realis vs. irrealis

(110) Neto meja g-ukat heta.
1sg table 3an-lift able
'I am able to lift the table.'
(111) Neto meja h-ukat heta.
1sg table 3INAN-lift able
'I will be able to lift the table.'

There is also a very small number of monovalent verbs which take the accusative prefixes to show agreement with their single arguments. On their own, it would appear that these predicates represent a (recent?) grammaticalization from trans-simpersonal verbs (see Malchukov, this volume, Mithun, this volume, and sections 2.7 and 2.8 below); but in light of the widespread semantic alignment found in related languages (see e.g. Klamer, this volume, for a survey, and Kolana in section 2.6.5 of this chapter), it is perhaps more likely that the accusative monovalent verbs of Bunak represent a retention of an earlier alignment system.

2.5.2 Basque

Basque is a language that has been considered, by various authors, to show ergative alignment, or else to show semantic alignment (Joppen and Wunderlich 1995, Hualde and Ortiz de Urbina 2003, C. Donohue 2005). Rather than examine this debate in detail, I refer the reader to Aldai (this volume), who examines the historical and geographic distribution of alignment in Basque varieties. The data that have been used to argue for a split in intransitive codings are shown in (112) and (113). While (112) shows absolutive case and agreement (on the auxiliary) for the S, (113) shows ergative morphology (and the use of the 'have' auxiliary, rather than the 'be' auxiliary).

Basque

(112) Mikel joan da. (113) Kepa-k tarda-tu
Mikel.abs go.perf be+tns Kepa-erg be.late-perf
du.
have+tns
'Mikel has went.' 'Kepa was late.'

Problems with an analysis that tries to differentiate these ergative and absolutive monovalent verbs in terms of semantic or syntactic features involve the fact that
both agentivity and patientivity are associated with both the ergative and the absolutive case-frame verbs.10

In addition to showing different argument codings in monovalent clauses, there is also a variety of case frames available to bivalent predicates. The common ergative-absolutive pattern is shown in (114), while (115) and (116) show ergative-dative and dative-absolutive predicates.

(113) Soldadu-ek haur guzti-ak hil zituzten.
    Soldier-PL.ERG child whole-PL.ABS kill.PERF HAVE+TNS:3PL.ERF>3PL.ABS 'The soldiers killed all the children.'

(114) Pablo-k Miren-i itxaroten dio.
    Pablo-ERG Miren-DAT wait.for HAVE+TNS:3SG.DAT 'Pablo is waiting for Maria.'

(115) Sagarr-ak gusta-tzen zaizkit.
    apple-PL.ABS like-IMPERF BE+TNS:1SG.DAT>3PL.ABS 'I like apples.'

Clearly any complete study of semantic alignment in Basque must also take into account the fact that the bivalent clauses, even more than the monovalent ones, display more than one case-marking option. While there is evidence for a system of semantic alignment in Basque, there is even stronger evidence for the dominance of semantic over syntactic marking when we examine bivalent clauses. The presence of more than one bivalent coding option is not unusual, and only the fact that there are three different choices in Basque, as well as there being more than one monovalent option, merits the inclusion of the language here.

2.6 Three-way splits and beyond

So far I have examined alignment with a two-way split: nominative-accusative, ergative-absolutive, agentive-non-agentive, etc. It is also possible for the basic alignment of a language to allow for a three-way split. The obvious way to have a three-way split is to have three (or more) different ways of marking (lexically distinct) Ss. This type of complex coding probably precludes simple order-with-respect-to-verb, such as seen in Ambonese Malay earlier, by virtue of there only being two positions with respect to the verb, preverbal and postverbal.11 Three-way splits tend to occur in head-marking languages, and tend to show splits in terms of some other alignment, either syntactic or semantic, though this is not an absolute characterization.

10 C. Donohue (2005) argues that internal causation is the semantic factor which differentiates the two classes.

11 It is possible to imagine a language with obligatory incorporation of some Ss, allowing for a three-way split. I am not aware of any such language.
We can see a three-way split in S coding in Warekena (Brazil, Maipuran: Aikhenvald 1998: 229–30), a language without much head-marking morphology. Here, not only is preverbal and postverbal position utilized to show the difference between A and Sa (preverbal) and P and Sp (postverbal) but some preverbal Ss are additionally coded in a PP, as in (120).\footnote{It is possible that the examples in (119) and (120) are transimpersonal constructions, and that a (perhaps somewhat abstract) 3sg causing A has been pro-dropped from the clause. Constructions of this type are discussed in more detail in section 7; see also Malchukov, this volume.}

\begin{verbatim}
 Warekena: bivalent clause: AVP

 (117) wa-ha wa-ja yut-sa-ha ema
       then-PAUSAL jaguar kill-PAUSAL tapir

       ‘Then the jaguar killed the tapir.’

 Monovalent clauses: V S

 (118) fupe-ha fiani-ke
       many-PAUSAL child-PL

       ‘Children are many.’

 (119) S V [PP S] V

      pea-nu-qi-tua wiyua
      one 1sg-brother die

      ‘One of my brothers dies.’

      nu-yue mawali
      1sg-for hungry

      ‘I am hungry.’
\end{verbatim}

In the following sections I shall examine the ways in which head or dependent marking can be used to split monovalent predicates into three categories, and the ways in which this use of semantic features to an extreme degree also affects the ways in which the arguments of bivalent predicates are encoded in the clause.

2.6.1 Three morphological markers

Languages of the Muskogean family (e.g. Choctaw, Chickasaw, Mikasuki, Koasati, Alabama; USA, southeast), and of the Yapen branch of the West Papuan family (Indonesia, Jones 1986, Donohue 2001, 2004b; distantly related to the North Halmahera languages discussed earlier, and in Holton, this volume) show a three-way distinction in the coding of Ss, but also have more than one way to mark an Ao raP.\footnote{Jones (1986) discusses Yawa, and while Yawa clearly has a split in marking for different Ss, it is not clear if this is a two-way or three-way split.} The possibilities are shown in outline in Table 2.8.

Muskogean languages (here exemplified with data from Koasati (Kimball 1991)) have optional nominative case marking for S or A; Yapen languages have optional (to different degrees) ergative marking for A. Muskogean languages show a switch-reference system that monitors (according to majority opinion) identity of S, A. In both Muskogean and Yapen languages, however, the verb shows a complex range of coding options. Basic clauses for Saweru, a Yapen language, are shown in (121) and (122). We can see that Ps and Sps are indexed by accusative prefixes, while As and Sas are indexed by nominative proclitics.\footnote{Agreement is complicated in Saweru by the fact that nominative agreement is by a proclitic to the VP. This means that in this SOV language the nominative agreement is separated from the V in
Table 2.8. Marking possibilities in Muskogean or Yapen languages

<table>
<thead>
<tr>
<th>S can be:</th>
<th>If A is …</th>
<th>… then P can be:</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOMINATIVE</td>
<td>NOM</td>
<td>ACC of DAT</td>
</tr>
<tr>
<td>ACCUSATIVE</td>
<td>ACC</td>
<td>ACC of DAT</td>
</tr>
<tr>
<td>DATIVE</td>
<td>—</td>
<td>—</td>
</tr>
</tbody>
</table>

Saweru: Basic bivalent and monovalent clauses

(121) a. Mo=na-ba-i.
3SG.F.NOM=2SG.ACC-hit-TNS
'She hit you.'

b. No=ra-ba-i.
2SG.NOM=3SG.F.ACC-hit-TNS
'You hit her.'

(122) a. Mo=rayan-i.
3SG.F.NOM=swim-TNS
'She swam.'

b. Ra-teson-i.
3SG.F.ACC-diarrhoea-TNS
'She has diarrhoea.'

So far the language appears to show a simple semantic alignment system, similar to many of the other languages examined already. In Saweru, the determining factor for the use of accusative agreement for an S is control: accusative Ss are uncontrolled (though some nominative Ss are also uncontrolled). Less common predicate types show further complications. Dative suffixes are used, in some cases, on the verb to index either an S or a P, as in (123). The use of dative agreement implies a low degree of affectedness on the part of the argument in question.

Dative agreement

(123) a. Manano-inai.
lonely-1SG.DAT
3SG.F.NOM=look.for-1SG.DAT
'I'm lonely.'

b. Mo=komi-inai.
3SG.F.ACC=look.for-1SG.DAT
'She looked for me.'

Further agreement patterns are found with the use of accusative prefixes to index a non-nominal A. In (124a) we can see the accusative prefixes indexing the A, which is not in control of the event, and the dative prefixes indexing the P, which is unaffected. In (124b) both the A and the P are indexed by means of accusative prefixes, on this bipartite verb stem.

Bivalent verbs with non-nominal subjects

(124) a. Ra-meme-inai.
3SG.F.ACC=think.of-1SG.DAT
'She's thinking of me.'

b. Ra-ne-na-mari.
3SG.F.ACC=forget-2SG.ACC-forget
'She forgot about you.'

the event of there being a nominal object, as in the following sentence based on (121b), employing the noun phrase ariau-o ruama 'girl': No=[ariau-o ruama] ra-ba-i 'You hit the girl.'
Koasati shows a very similar pattern. (125)–(127) show the three different coding choices with monovalent verbs; for each of the nominative, accusative, and dative possibilities a sample of verbs showing this pattern has been given.

Koasati: monovalent

(125) a.  
\[ \text{a} \text{ì} \text{í}-\text{l} \]  
\[ \text{go}-1\text{sg}.\text{nom} \]  
\[ 'I go.' \]  
\[ (\text{há:lon} 'hear', \text{ó:tin} 'gather', \text{t}:\text{sin} 'take one thing', etc.) \]

b.  
\[ \text{a} \text{ì} \text{í}-\text{cí}-\text{y} \]  
\[ \text{go<2sg}.\text{nom} \]  
\[ 'You go.' \]

(126) a.  
\[ \text{ca-\text{ficcák}} \]  
\[ 1\text{sg}.\text{acc}-\text{be.jealous} \]  
\[ 'I am jealous.' \]  
\[ (\text{íllin} 'die', \text{ó:tin} 'gather', \text{t}:\text{sin} 'take one thing', etc.) \]

b.  
\[ \text{ci-\text{ficcák}} \]  
\[ 2\text{sg}.\text{acc}-\text{be.jealous} \]  
\[ 'You are jealous.' \]

(127) a.  
\[ \text{am-akán} \]  
\[ 1\text{sg}.\text{dat}-\text{be.hungry} \]  
\[ 'I am hungry.' \]

b.  
\[ \text{cim-akán} \]  
\[ 2\text{sg}.\text{dat}-\text{be.hungry} \]  
\[ 'You are hungry.' \]

As with Saweru, while the 'basic' coding pattern for bivalent verbs is nominative-accusative, as in (128), other combinations are also found. Only one example of these different possibilities, showing a nominative-dative predicate, is given, though all of the combinations promised in Table 2.8 are attested.

(128)  
\[ \text{mán} \text{haci-\text{hi:ca-li-}laho-\text{v}} \]  
\[ \text{again} 2\text{pl}.\text{acc}-\text{see}-1\text{sg}.\text{nom}-\text{irr}-\text{phrase.final} \]  
\[ 'I will see you all again.' \]

(129)  
\[ \text{cim-há:lo-li-}laho-\text{v} \]  
\[ 2\text{sg}.\text{dat}-\text{hear/obey}-1\text{sg}.\text{nom}-\text{irr}-\text{phrase.final} \]  
\[ 'I will obey you.' \]

2.6.2 Economical use of two markers

Three-way splits in monovalent subject coding can be achieved in ways other than having three morphologically distinct paradigms. The use of two paradigms to create three classes of verbs is common in eastern Indonesia (Donohue 2004b), and will be illustrated here with data from Nuaulu (Austronesian, Indonesia; Bolton 1990: 36–42). In this language, prefixal agreement is obligatory on all verbs, while suffixal agreement is only used to mark the P of a bivalent clause, or (in combination with the prefixes) the Sp of a stative monovalent clause. As can be seen in (131b), even though the Sp is marked by suffix, the nominative prefix is still obligatory; the suffixes cannot be used on eventive verbs such as \textit{anamama} in (131a), so we can clearly talk about there being two morphologically distinct classes of verb based on the agreement strategies found on the verb.
Semantic alignment systems: what's what, and what's not

Nuaulu: bivalent

(130)  a.  \textit{U-sosa-i.}  \\
    1sg-rub-3sg  \\
    'I’m shining it.'  \\
  b.  \textit{i-na-ku \ i-hita-ku.}  \\
    mother-1sg.poss 3sg-hit-1sg  \\
    'My mother hit me.'  \\

Monovalent: prefix for both Sa and Sp, suffix obligatory for Sp

(131)  a.  \textit{U-anamana.}  \\
    1sg-speak  \\
    'I’ll speak.'  \\
  b.  \textit{U-ampeta-ku.}  \\
    1sg-wet-1sg  \\
    'I am wet.'

In addition to these two monovalent verb classes, one showing agreement by prefix and one by prefix and suffix, there is an additional class of intradirective verbs (roughly, motion verbs, the class of verbs whose sole argument is simultaneously an agent and also a theme), for which agreement by prefix is obligatory and suffixes are optional. There is no reported semantic distinction between the use of a single agreement affix, as in (132a), and double exponence, as in (132b), so treating this semantically unified class of verbs as showing labile behaviour is not easily justified. While this does not represent a clear and undeniable case of a three-way split such as was seen in the Muskogean or Yapen languages, it does nonetheless distinguish three semantically defined classes of verbs.

(132)  a.  \textit{U-eu ria.}  \\
    1sg-go inland  \\
    'I’ll go inland.'  \\
  b.  \textit{Ia \ i-hoka-i \ tewa.}  \\
    3sg 3sg-go-3sg neg  \\
    'He didn’t come.'

A similar economical use of only two marked options to code a three-way set of distinctions can be found in Haida (Enrico 2003). Pronouns show two different sets, which Enrico calls the ‘agentive’ and ‘objective’ sets. Additionally, a small set of verbs can occur with either the agentive or the objective set pronouns, and this set of verbs shows a relatively consistent semantics.\(^{15}\) Just as with the Muskogean and Yapen languages, this versatility of S-marking also extends to A-marking. Table 2.9 shows examples of the coding options for Ss and As with different verbs.\(^{16}\)

The systems seen in this section do not have simply one way to mark As, with some Ss marked in the same manner, and one way to mark P, with some other Ss marked in that manner. Rather, the marking system allows for a split in As and (in Muskogean and Yapen) a split in Ps. What we have is a coding system which

\(^{15}\) The fact that in Haida the verbs showing the split behaviour include bodily or mental acts that are susceptible to different degrees of control suggests that a ‘labile’ analysis, in which verbs such as ‘sneeze’ can be treated as more or less controlled, is warranted.

\(^{16}\) Some verbs vary in their coding options between the Masset and Skidegate dialects. There does not appear to be any consistent relationship between the two dialects in terms of alignment coding. Thus \textit{sk’al.aaw} ‘have diarrhoea’ is objective in Masset, and shows variation between agentive and objective in Skidegate. \textit{Q’usahlda} ‘cough’ is agentive (M), or variable (S), and \textit{q’anda} ‘belch’ is variable (M) or agentive (S). See Mithun, this volume, for further discussion of the complexities of semantic alignment in Haida.
Table 2.9. Haida three-way coding split

<table>
<thead>
<tr>
<th></th>
<th>Monovalent</th>
<th>Bivalent</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Agentive subject</strong></td>
<td>rad 'run', srayhla 'cry', kusad 'fart', qaa 'come/go',</td>
<td>qing 'see, look at', da.a 'have, keep', qii.a 'find, receive',</td>
</tr>
<tr>
<td><strong>Objective subject</strong></td>
<td>skaa'hda 'hiccup', dladabla 'fall down', sk'al.aaw 'have diarrhoea', q'i.id 'remember PP', kaa.âydà 'feel and act playful'</td>
<td>q'al.a 'be unaquainted with', graa7ala 'resemble', go'lâa 'like',</td>
</tr>
<tr>
<td><strong>Agentive/objective subject</strong></td>
<td>skin.äng 'keep waking up', hats'asaa 'sneeze', q'anda 'belch', 7anggung 'be curious about PP'</td>
<td>tlâgâng 'vomit up'</td>
</tr>
</tbody>
</table>

is much more sensitive to semantic distinctions everywhere than is one which really cares about A, S, and P. Should this be described as semantic alignment, or simply semantically explicit marking? Crucially, is there a principled difference between the two labels? If we accept coding in semantically explicit ways as a form of semantic alignment, then a number of semantically motivated morphosyntactic splits (or at least splits that have semantic origins) should perhaps be considered under this label.

2.6.3 Icelandic

In northeastern Europe we find a split in the coding of Ss in Icelandic, a language that has received a great deal of attention from syntacticians (see e.g. Andrews 1990a, 1990b, 2001, Eythórsson and Barðdal 2005, Faarlund 2000, Rögnvaldsson 1994, Van Valin 1991, Zaenen and Maling 1990, Zaenen, Maling, and Thráinsson 1985). Because the Icelandic data have been so well studied, and because they show so many interesting features, I shall present a number of features associated with grammar, while noting that I am not even attempting to describe the well-documented diachronic situation.

Empirically, we can observe a four-way split in the coding of Ss in Icelandic, with any one of four cases, nominative, accusative, dative, or genitive, being used to mark these arguments, depending on the verb used. This complex set of case-marking alternatives corresponds to a split in case possibilities for As, and a similar split for Ps (though the possibilities available for the marking of a P are restricted by the choice made for the A). Examples of the possibilities for marking monovalent subjects are shown in (133)–(137). Note that the ‘default’ nominative case can be used not only for activities, such as in (133), but also for non-agentive events such as sinking, as seen in (134). Dative subjects are generally experiencers, such as in (135), while accusative subjects typically involve a lack of control. There
Semantic alignment systems: what’s what, and what’s not

are too few genitive subjects to try to characterize this group semantically, and in
the modern language they are very unstable (though see the literature cited above).

Icelandic

(133) Stelpurnar hláu.
   thegirls:NOM laughed:PL
   ‘The girls laughed.’

(134) Skipið sökk.
   the:ship:NOM sank
   ‘The ship sank.’

(135) Mér er kalt.
   1SG.DAT is cold
   ‘I am cold.’

(136) Drengina rak á land.
   the:boys:ACC drifted to land
   ‘The boys drifted to land.’

(137) Vindsins getir ekki.
   the:wind:GEN matters not
   ‘The wind does not matter.’

The very loose semantic characterizations given above obviously lead to the possi-
bility of semantically similar arguments being assigned different cases by different
verbs, and thus of there being evidence of degrees of lexicalization. This is indeed
the case; in (138) and (139) we see two verbs which are semantically extremely
similar, but which assign different cases to their subjects.17

(138) Mig velgir við setningafraði.
   1SG.ACC am.nauseated by syntax
   ‘I am nauseated by syntax.’

(139) Mér byður við setningafraði.
   1SG.DAT am.nauseated by syntax
   ‘I am nauseated by syntax.’

Furthermore, we can note the existence of what might be termed ‘covert quirky
case’ subjects in Icelandic. The reflexive verbs in (140) and (141) both take nomi-
native subjects, but the obligatory reflexive element can be accusative, as in (140),
or dative, as in (141). Once more there is a split in the case marking of monovalent
predicates, though it is disguised.

(140) Hann rækti sig.
   3SG.M.NOM clear.throat self:ACC
   ‘He cleared his throat.’

(141) Hann snýtti sér.
   3SG.M.NOM blow.nose self:DAT
   ‘He blew his nose.’

In (142)–(144) we can see a few of the possible combinations of case found
with bivalent verbs. The most common case frame is nominative-accusative,
but other cases may appear on the object, as seen in (143) and (144). While

17 There is also the well-discussed phenomenon of ‘dative sickness’, whereby objects that are
expected to appear in accusative case are starting to appear with dative case endings (see e.g. Smith
1994). Similarly, non-nominative subjects are being replaced by nominative ones, and genitive objects
are being replaced by either accusative objects or arguments inside PPs.
we cannot state that certain semantic roles are assigned to particular cases, it
is generally true that dative-marked objects are experiencers (this characteriza-
tion is more successful than an attempt simply to characterize dative-marked
subjects).

(142) \textit{Strákurinn} kitlaði stelpuna.
\text{the:boys:NOM} tickled \text{the:girls:ACC}
'The boys tickled the girls.'

(143) \textit{Han} bjargaði mér.
\text{3SG.M:NOM} saved \text{1SG.DAT}
'He saved me.'

(144) \textit{Ég} mun sakna hans.
\text{1SG.NOM} will miss \text{3SG.M:GEN}
'I will miss him.'

Furthermore, the subject of a bivalent may also appear in a non-nominative case,
and if that happens then the nominative may be used for the object (there are no
double-nominative case frames), as seen in (145)–(147).

(145) \textit{Stúlkuna} vantar efni í ritgerðina.
\text{the:girl:ACC} lacks \text{material in the:paper:NOM}
'The girl lacks material for the paper.'

(146) \textit{Stráknum} líkar slíkir bílar.
\text{the:boy:DAT} likes such \text{cars:NOM}
'The boy likes such cars.'

(147) \textit{Barninu} batnaði veikin.
\text{the:child:DAT} bettered \text{the:disease:NOM}
'The child recovered from the disease.'

As with the monovalent examples in (138) and (139), it is possible to find bivalent
predicates that are semantically very similar, and yet which take different case
frames. This means that the event structure and semantic roles of the arguments
of two different verbs can be extremely similar, without this dictating that the case
frames for the two different predicates are identical.

(148) \textit{Hann} hjálpaði mér. \textit{(149) Hann} aðstoðaði mig.
\text{3SG.M:NOM} helped \text{1SG:DAT} \text{3SG.M:NOM} assisted \text{1SG:ACC}
'He helped me.' 'He assisted me.'

For some verbs involving incremental themes we can identify different senses
correlating with different case patterns, (150) and (151); for these verbs, the dative
is commonly used to mark a human \text{P}, rather than a non-human one (see Barðdal
1993, Svenonius 2002a). \text{(The accusative is also grammatical for the human Ps, but
the dative cannot be used with a non-human P.)}
Semantic alignment systems: what's what, and what's not

(150) Hann óvöði gólfð.  
3sg.m.nom washed the:floor:acc  
‘He washed the floor.’

(151) Hann óvöði barninu.  
3sg.m.nom washed the:child:dat  
‘He washed the child.’

An accusative-dative alternation is also found in cases that can be considered analogous, in an abstract way, to the dative alternations that are found in English. In (152) we see the verb for ‘scratch’ used with an accusative object, while in (153) the same verb appears with an object that can be construed as a beneficiary, and this object is marked with the dative.

(152) Kötturinn klóraði mig.  
thecat:nom scratched 1sg.acc  
‘The cat scratched me.’

(153) Ég klóraði kettinum.  
1sg.nom scratched the:cat:dat  
‘I scratched the cat.’

A similar alternation exists with verbs denoting directed motion away from the subject, depending on whether the object of the verb is the affected target or the instrument propelled. The is analogous to the use of dative case with verbs of propelled motion that do not involve the accompaniment of the subject, compared to the use of accusative on objects that are propelled without the subject accompanying (assisted motion); examples of these two different verb classes are shown in (154) (Svenonius 2002a).

(154) a. skjóta fuglinn  
shoot the:bird:acc  
‘shoot the bird’

b. skjóta kúlunni  
shoot the:bullet:dat  
‘shoot the bullet’

(155) a. draga acc ‘pull, drag’  
farra acc ‘move, bring’  
hakka acc ‘raise’  
henda dat ‘throw away, discard’  
velta dat ‘roll (a) barrel’

b. kasta dat ‘throw, fling, hurl’

Other verbs show alternations similar to these based on whether the object is a field or a theme; field takes accusative case, while themes that are removed from the field to which the predicate is applied appear in dative case (Svenonius 2001, 2002b).

(156) Hann sópar gólfð.  
3sg.m.nom sweeps the:floor:acc  
‘He sweeps the floor.’

(157) Hann sópar ruslinu í poka.  
3sg.m.nom sweeps the:garbage:dat in bag  
‘He sweeps the garbage into a bag.’
This sketch by no means exhausts the possible discussion of case in Icelandic, but it adequately serves to demonstrate the fact that there are multiple splits in the coding choices available to As, Ss, and Ps in the language. The split in the marking of Ss corresponds somewhat to semantic categories, but is also to a large extent lexicalized (a common fate for such morphological processes in these languages). The overwhelming tendency is for arguments of any semantic role to be accommodated in a nominative(-accusative) case frame.

These examples show that different kinds of motion away from the subject are as important in characterizing the classes of verbs that govern different cases as are more argument-based semantic notions such as bearing a semantic role of 'experiencer'. The fine aspectual distinctions seen in (154)–(157), and the fact that verbs with very similar semantics can display different case possibilities, suggests that the split in marking is no longer strongly dependent on semantic roles, but is moving towards being more of an aspectually governed system. A similar state of having more than one semantic factor underlying the split in marking of S in the clause is shown for Otomi (Palancar, this volume), while the notion that the kind, or direction, of motion can have a bearing on argument coding is also explored for two South American languages Pilagá (Vidal, this volume) and Guaraní (Velázquez-Castillo, this volume).

We should ask whether this sort of multiple splitting of argument-encoding options is the same kind of alignment as is found in, say, Warekena or even that or the simpler Waris. It is true that there is a nominative set of verbal agreement suffixes in the language that link S to A, but the case-marking system is more complex than that, with some Ss being coded in the same way as canonical As, while some are coded as canonical Ps. We have seen there is no reason to disallow a semantic alignment analysis simply because there is a split in the coding of alignment across different primary coding devices, and that one part of the grammar might be aligned semantically (the case system in Icelandic), while another might be arranged more syntactically (the Icelandic agreement system). In the following section I present an account of Tagalog, a well-known Austronesian language. Here the nominal case-marking system cannot be said to show semantic alignment, but the verbal affixation shows strong semantic leanings.

2.6.4 Tagalog

In Tagalog, an Austronesian language of the northern Philippines (see Schachter and Otanes 1972, Schachter 1976, Kroeger 1993, and many others), verbs appear with one affix, which indexes which participant in the clause has the status of subject (see Tsukida, this volume, for a discussion of similar behaviour in Amis, another Austronesian language with a similar morphosyntactic profile; similar behaviour is found in most northern Austronesian languages). Table 2.10 shows the different affixes that are used to encode As, Ss, and Ps as subjects on the verb;
there is a great deal of lexical stipulation of affix identity for different roots, but the less common classes formed do show general semantic clusterings. The primary (in terms of lexical frequency and productivity, and extensibility into smaller lexical classes) coding options for clauses in which an A or an S is the subject are the affixes mag- and -um-, and for Ps -in, as shown in (158). Here we have examples of the most semantically bleached affixes, the infixes -um- and -in-, being used to mark the A or the P (respectively) of the clause as the subject; morphologically, the subject is indicated by the use of the nominative case (ang with common nouns) on the subject NP; syntactically a host of constructions favour this argument (Kroeger 1993 presents the data most clearly). The non-subject appears in genitive case.

18 Many of these same affixes, particularly those from the bottom of the different columns, are also used to index a dative argument (D), or a range of other adjuncts.
Tagalog

(158) a. $P<um>\text{-}\text{utol siya ng punungkahoy.}$
   $\text{cut}<\text{um}>\text{ 3SG.NOM GEN tree}$
   ‘He cut trees.’

b. $P<in>\text{-}\text{utol niya ang puno ng mangga.}$
   $\text{cut}<\text{in}>\text{ 3SG.GEN tree GEN mango}$
   ‘He cut down the mango tree.’

In (159) we can see the same affixes used with monovalent verbs, indexing a more agentive, and less agentive S. The use of -\text{in-} with monovalent verbs is rare, with \text{mag-} and -\text{um-} generalizing as nominative markers.

(159) a. $D<um>\text{-}\text{ating siya sa halamanan.}$
   $\text{arrive}<\text{um}>\text{ 3SG.NOM DAT garden}$
   ‘He arrived in the garden.’

b. $<\text{In}>\text{-a-antok siya.}$
   $\text{RED}<\text{IN}>\text{-sleepy 3SG.NOM}$
   ‘He’s sleepy.’

Another verb such as \text{bukas} ‘open’ marks the A or P as subject with a different paradigm of affixes. Here the marker for A-as-subject is the generic \text{mag-}, while the marker for P-as-subject is -\text{an}, historically an applicative that now marks low-affect Ps. Note that a monovalent use of the verb is also possible, as shown in (160c), using the infix -\text{um-}.

(160) a. $\text{Mag-bukas siya ng pinto.}$
   $\text{MAG-open 3SG.NOM GEN door}$
   ‘He opened a door.’

b. $\text{Buks-an niya ang pinto.}$
   $\text{open-AN 3SG.GEN door}$
   ‘He opened the door.’

c. $B<um>\text{-}\text{ukas ang pinto sa lakas ng hangin.}$
   $\text{open}<\text{um}>\text{ NOM door DAT strong GEN wind}$
   ‘The door opened in the strong wind.’

We have seen that the same morphology that indicates an A-as-subject in (158a) can also be used to show an S-as-subject, in (159a) and (160c). As can be seen in Table 2.10, approximately half of the affixes that can be used to index an S are also used with As, and half of them are also used with Ps, with one prefix, \text{ma-}, being employed for all three syntactic roles with different verbs (it is no accident that \text{ma-} is also a productive derivational affix to a degree that is not true for the other affixes shown here). Table 2.10 also offers a rough characterization of the qualities of an S that are marked with the different affixes (following Schachter and Otanes 1972); to various extents these characterizations also apply to the A
and P arguments marked by these affixes. In (161) we can see the same prefix \textit{ma-} used to index either an A, an S, or a P, depending on the verb used; note that with an S the use of \textit{ma-} implies an inchoative event, while the same predicate without the \textit{ma-} is interpreted statively. (162) shows sentences corresponding to (161a) and (161c) in terms of verb selection that mark the other argument as the subject of the clause. Common correspondences between A-marking affixes and P-marking affixes are shown in Table 2.11.\footnote{Sells (1998) suggests that the ‘voice markers’ in Philippine-type languages such as Tagalog can best be thought of as pronominal affixes. This analysis would result in the semantic split in verb marking ‘fitting in’ more closely with the analysis of other languages discussed here.}

Tagalog

\begin{tabular}{lccc}
\toprule

\textbf{A} & \textbf{P} & \textbf{Example} \\
\midrule

\textit{mag-} & \textit{i-} & \textit{balita} ‘tell something’ \\
& \textit{-in} & \textit{alis} ‘remove’ \\
& \textit{ipa(n)g-} & \textit{tiis} ‘endure’ \\
& \textit{-an} & \textit{bantay} ‘watch’ \\
\textit{-um-} & \textit{-in} & \textit{bati} ‘greet’ \\
& \textit{-an} & \textit{hawak} ‘hold’ \\
\textit{mang-} & \textit{-in} & \textit{kailangan} ‘need’ \\
& \textit{ipa(n)g-} & \textit{anak} ‘bear child’ \\
& \textit{-an} & \textit{kuwala} ‘rob’ \\
& \textit{pa((n)g)- -an} & \textit{nood} ‘watch’ \\
\textit{maka-} & \textit{ma -an} & \textit{limut} ‘forget’ \\
& \textit{ma-} & \textit{kita} ‘see’ \\
\textit{ma-} & \textit{pa((n)g)- -an} & \textit{kingig} ‘listen to’ \\
& \textit{ma -an} & \textit{tuto} ‘teach’ \\
\bottomrule
\end{tabular}

\section*{Semantic alignment systems: what’s what, and what’s not}
Mark Donohue

(162) a. **Ma-tutu-an ng mga bata ang matematika.**
    ma-learn-an gen pl child nom maths
    'The children learn the maths.'

b. **Maka-kita ng mga bata ang maestra.**
    maka-see gen pl child nom teacher
    'The teacher saw children.'

As with Icelandic, the marking found with different predicates is not unique to monovalent clauses, and the same attention to semantic distinctions is found in bivalent clauses as well. There is quite a degree of lexicalization in the choice of affixes with different Tagalog verbs, just as different predicates, even semantically very similar ones, can show different case arrays in Icelandic.20 The Tagalog affixes have variously been argued to be derivational, inflectional, or pronominal; the ultimate analysis will most probably combine these different characterizations.

2.6.5 Kolana

Kolana is a non-Austronesian language of eastern Alor, in southern Indonesia. Prefixes show agreement on the verb in Kolana, but only for S or P arguments (Donohue 2004b). We can, therefore, say that there are absolutive prefixes on verbs; this can be seen in (163).

<table>
<thead>
<tr>
<th>Kolana: bivalent</th>
<th>Monovalent</th>
</tr>
</thead>
<tbody>
<tr>
<td>(163) a. <strong>Geta n(a)-poin.</strong></td>
<td>b. <strong>N(a)-tati.</strong></td>
</tr>
<tr>
<td>3sg.erg 1sg.abs.i-hit</td>
<td>1sg.abs.1-stand</td>
</tr>
<tr>
<td>'She hit me.'</td>
<td>'I stood up.'</td>
</tr>
</tbody>
</table>

In addition to the most general prefixes (exemplified for 1sg in (163)), there are two more prefix sets. In both cases they are available to mark agreement only for Ss and Ps, and so are also absolutive prefixes. The kinds of predicate that can be marked with the set II and set III prefixes are more semantically restricted than are found with the set I prefixes, the examples in (164) and (165) being representative of their classes. Here we have a language with syntactic alignment, in that the primary split is between the A and the S,P grouping, which introduces a set of semantic divisions within this absolutive category.

| (164) a. **Geta nai-suai.** | b. **Nai-lalan** |
| 3sg.erg 1sg.abs.ii-stab | 1sg.abs.ii-fever |
| 'She stabbed me.' | 'I have a fever.' |

| (165) a. **Geta nadi-modo.** | b. **Nadi-wiri.** |
| 3sg.erg 1sg.abs.iii-discard | 1sg.abs.iii-cold |
| 'She discarded me.' | 'I'm cold.' |

20 For instance, in addition to kita 'see', which inflects for av and pv with maka- and ma-, respectively, we also find tingin 'watch', which inflects with -um- and -an.
The alignment found in Kolana suggests strongly that an earlier, 'cleaner' system of semantic alignment is in the process of being replaced by a more syntactically governed one. The fact that other languages related to Kolana, such as Bunak (section 2.5.1) and Oirata (de Josselin de Jong 1937; Donohue and Brown 1999), have all but lost the semantic aspects to their agreement, and others such as Kui have accusative prefixes (with a semantic split), supports this position (see Klamer, this volume, for further data on languages related to Kolana and Bunak). The Tanglapui data in the following section show another way in which semantic factors maintain a tenuous presence in the alignment system of a language of this family.

2.7 A split in 'intransitives'? Approaching 'S' via P and the limits of 'semantic alignment'

Another major construction type that impinges on the typology of semantic alignment involves the so-called 'transimpersonal' constructions, constructions in which the verb (at least) is formally transitive, but the clause shows the behaviour of an impersonal construction, in that the subject does not appear in the case normally associated with monovalent subjects (see Malchukov, this volume, and Nichols, this volume, for further discussion of the implications of viewing this construction as an instance of semantic alignment, and Mithun, this volume, for thoughts on the role played by this construction in the evolution of semantic alignment systems).

In Tanglapui (Trans New Guinea, Alor, southern Indonesia; Donohue 1996b, and Klamer, this volume) we can observer a split in the coding of monovalent clauses by the use of the inverse prefix on verbs with either an affected P or an affected Sp. (166) shows that monovalent clauses with non-affected Ps do not require use of the inverse, while in (167) we can see that when the P is affected by the predicate, and the P outranks the A in terms of the animacy hierarchy, the inverse prefix na- must be used.

(166) a. Ya-di-a.
   2sg-see-asv
   'You saw (her/him/it).'
   2sg-asp
   b. Di-a.
   hit-asp
   '(He/It/She) hit (you/her/him/it).'
   Bivalent: affected P

(167) a. Ya-baba.
   2sg-hit
   'You hit (her/him/it).'
   2sg-inverse-hit
   b. Ya-na-baba.
   '(He/It/She) hit you.'

When we examine predicates with a single argument we see a similar split in behaviour: when the argument (the S) is not affected, by the predicate, the verb appears with the agreement prefix alone. When the single argument of the predicate is affected, the agreement prefix appears in conjunction with the inverse...
prefix. In other words, the affected S is formally coded identically to an affected P—but not identically to all Ps.

Monovalent: non-affected and affected S

(168) a. Ya-ve
   2SG-go
   'You went.'

b. Ya-na-tansi.
   2SG-INVERSE-fall
   'You fell.' ('it fell you'?)

If we consider this to be an instance of semantic alignment, then a range of other transimpersonal constructions need to be considered. In Skou (Macro-Skou family, north-central New Guinea), a number of predicates code the experiencer as the object of the clause. Thus in (169) the burper is coded as the object of the clause, clearly signalled as such by the use of láng, a suppletive form of the verb ‘hit’ used with feminine objects. In (170) löøŋri ‘snout’ is the subject of the clause, insofar as it determines the use of 3SG.f forms of the verb, though notice that ni ‘1SG’ occupies the clause-initial position. This suggests that there are some subtle differences in the way the predicates ‘burp’ and ‘be snotty’ are conceptualized, as well as encoded.

Skou

(169) Oe pe ke=láng.
   burp 3SG.F 3SG.NF=burp
   'She burped.'
   (literally, '(A) burp hit her. ; the verb is phonologically identical to ‘hit (FEM.OBJ)')

(170) Ni löøŋri tue e tue.
   1SG snot 3SG.F.do 3SG.F.be 3SG.F.do
   'I’m full of snot.'

The kind of impersonal constructions seen in Skou are common in many parts of the world, including New Guinea. Tauya (Trans New Guinea, Papua New Guinea; MacDonald 1990) is one such language. (171) shows the experiencer marked as the P of the verb sepame ‘sick’, though there is no nominal representing an A in the clause, only the S,A suffixed. In (172) we have a similar construction, with the experiencer marked as the P of the clause even though this is a bivalent clause with an obliquely marked second argument. It is clear that the -?a suffix is purely pleonastic. 21

Tauya

(171) Ya -sepame-ti-a-?a.
   1SG.P-sick-INTENS-3SG.S/A-IND
   'I am really sick.'

21 Donohue (2005b) discusses these constructions in New Guinea languages, arguing that the P-coded arguments are in fact objects, and not subjects with ‘quirky’ agreement patterns.
Semantic alignment systems: what’s what, and what’s not

(172) Na-ra awa na-pi-pe na-sa-fe-a-?a.
   2SG-TOP father 2SG-GEN-BEN 2SG-P-ANGRY 3SG.S/A-IND
   ‘You’re angry at your father.’

Warembori is (probably) an Austronesian language from northwest New Guinea. In Warembori we find similar examples of experiencer-as-object constructions. In some cases there is an overt cause mentioned in the clause, such as monggena in (173). In others the cause must be incorporated into the verb, as seen in (174), and in (175) we can see a construction similar to that found in Nuaulu (section 2.6.2).

Warembori (Austronesian, Donohue 1999b: 41–2)

(173) Mongge-na ban-e-o.
   snot-PL make-1SG.P-IND
   ‘(My nose) is running with snot.’

(174) Doro-pai-tan-e-o.
   rain-affect-APPL-1SG.P-IND
   ‘I got soaked by the rain.’

(175) A-vaitumban-e-o.
   1SG-tired-1SG.P-IND
   ‘I’m tired.’

Another construction similar to the inverted clauses here forming a minor part of its repertoire is Tukang Besi (Austronesian, central Indonesia: Donohue 1999a: 96, 134). Tukang Besi is not part of the New Guinea ‘area’, and makes less use of these constructions than languages in New Guinea, but nonetheless shows some experiencer-as-P codings. In (179) we can see another example of the double marking that we saw in Nuaulu and Warembori (for further examples of this type of morphosyntax in the eastern Indonesian area, see Donohue 2004b).

Tukang Besi

(176) No-motiti=a ku te ‘olo o s<um>:io.
   3REAL-dry=1SG.P core sun afternoon.stl
   ‘I dried in the afternoon sun.’
   (∼ ‘The afternoon sun dried me.’)

(177) No-raho=ka mi te wande.
   3REAL-affect=1PA.P core wind
   ‘We were tossed about in the wind.’
   (∼ ‘The wind affected us.’)

(178) O-raho=’e te watu.
   3REAL-affect=3P core stone
   ‘He banged himself on a stone.’
   (∼ ‘A stone affected him.’)

(179) To-langke-ono’o-ngkita.
   1PL-REAL-sail-be.six-1PL.NUM.P
   ‘Six of us went sailing.’
Ambai is another Austronesian language of northwest New Guinea, probably reasonably closely related to Warembori (Silzer 1983). Basic clauses see an S or an A preverbal, with agreement on the verb, and the P postverbal, appearing either as a separate (pro)nominal, or (for 3rd persons) as a suffix on the verb. Examples are shown in (180)–(183). In (180) we can see a preverbal A, and in (181) and (182) are two examples of preverbal Ss. A postverbal P is shown in (183).

(180) Yau y-eo y-ang diang.  
1sg 1sg want 1sg -eat fish 
'I want to eat fish.'
(181) Yau i-minohi na Warironi.  
1sg 1sg-live prep Warironi 
'I live in Warironi.'
(182) Yau i-matai i-fafa we reirei.  
1sg 1sg -afraid 1sg -swim prep shore 
'I was afraid and swam for shore.'
(183) D-eti yau kaha.  
3sg 3sg -see 1sg NEG 
'He didn’t see me.'

There are two exceptions to this characterization. First, a number of verbs require a reflexive subject, in that a free pronoun must appear in the object position as well as there being nominative agreement on the verb. The verbs that behave in this way are all verbs of motion, and their behaviour is identical to that of predicates with reflexive objects, such as seen in (186).

(184) a. I-karobu yau na airauri.  
1sg-dive 1sg prep water 
'I jumped/dove into the water.'

b. *I-karobu na airauri.  
'I jumped/dove into the water.'

(185) a. Aha mani y-arabera yau ma.  
tomorrow top 1sg-return 1sg hither 
'Tomorrow I’ll return here.'

b. *Aha mani y-arabera ma.  
'Tomorrow I’ll return here.'

(186) a. D-eparandin-i.  
3sg-praise 3sg 
'He glorified himself.'

b. Y-eparanding yau.  
1sg-praise 1sg 
'I praised myself.'

The other exception involves predicates of a different syntactic category. When a nominal predicate appears with the copular di, the pronoun must appear following the copular. Rather than considering this to be an ‘accusative’ subject, it is better to think of it as an inverted clause, in which the predicate is initial, rather than the subject being initial (see the Tukang Besi examples in (35) and (36) for another example of nonverbal clauses showing a different syntax to verbal clauses).
Semantic alignment systems: what's what, and what's not

(187) (Yau mani) mam-biriu di yau.
1SG TOP NOM-strong COPULA 1SG
'I’m the champion.'

Northern Australia is another area that shows similar trans impersonal constructions (Walsh 1987). In Ndëbbana (MacKay 2000: 270, 272) the verb in (188) is simply ‘I sit’. (188), however, also contains the nominal barakángka. This nominal is part of the construction with ‘sit’ that specifies '(be) worn out', and the predicate can be considered as a complex N+V predicate with ‘regular’ coding of the S. In (189), however, the agreement on the verb is 3rd person, clearly agreeing with mangúya, a nominal that is logically the possessor of the experiencing S.

Ndëbbana

(188) Barakángka nga-nó-ra.
worl.out 1MINIMAL.S-sit-contemporary
'I’m worn out.'

(189) Mangúya ka-ddjúwa ka-nó-ra.
throat 3MINIMAL.S-suffer 3MIN.S-sit-contemporary
'I’m really sad.'

Similar constructions are found in a great range of languages, including those of western Europe. While this is archaic in English (in idioms such as methinks), we find that it is productive in Dutch, as in (190)—(192). In (190) the experiencer of the ‘cold’ adjectival predicate is marked in accusative case, with a pleonastic ‘subject’ and 3rd person singular verb agreement. In (191) the subject is marked in nominative case, and the verb shows the lack of overt agreement that is characteristic of 1st person singular in the present tense, but a reflexive ‘object’ also appears in the clause. (192) shows a pattern similar to that in (190), with a verbal predicate, and pragmatic fronting of the experiencer.

Dutch

(190) Het is mij te koud.
it is 1SG.ACC too cold
'I’m cold.'

(191) Ik voel me niet goed.
1SG.NOM feel 1SG.ACC not good
'I don’t feel well.'

(192) Mij lijkt het niet gezellig.
1SG.ACC strike-2/3SG it not ‘cozy’
'I find it impersonal.'

The material in this and the previous section suggests that we should perhaps consider a analysis of some of the data as not involving split-intransitive phenomena.
While they clearly represent semantic alignment, a more detailed examination of the syntax of these languages might reveal that many of the non-agentive S arguments are in fact Ps, with pleonastic (or elided) ‘dummy’ subject. For instance, in (193) we have an example of a non-agentive verb from Lakota (Boas and Deloria 1941: 81). Given that 3rd persons do not show agreement on the verb, it is hard to see why (193) could not be analysed as also showing a null 3rd person, such as the Dutch example in (190), with more literal translations resembling ‘it chills me’.

(193) *ma-c’úwita*  
1sg.NONACTIVE-cold  
‘I am cold.’

(194) *(Ø-)*ni-c’úkša  
2sg.NONACTIVE-wrestle  
‘He wrestles with you.’

We have seen that in many languages in which an S shows more than one coding option, either the A or P will also show variation in coding. This is not a feature unique to languages with semantic alignment realized in the behaviour of single-argument verbs; in English it is well known that not all Ps are encoded with the same morphosyntax, as can be seen in examples (195)–(197), showing the conative construction and the alternations in P encoding that this construction requires.

As indicated, the use of the conative in (195a) implies a more agentive event than that encoded without the conative, while not using the conative implies a greater likelihood of the P being affected, something that is not specified with the conative. Semantically close predicates, such as *look at* and *see* or *listen to* and *hear,* sometimes also show this alternation: the event in which the A must be construed as being more agentive appears with a prepositionally marked object.

(195) a. *The police shot at the robbers.*  
(+ agentive), [± affected]  
b. *The police shot the robbers.*  
(± agentive), [+ affected]

(196) a. *The police looked at the robbers.*  
(+ agentive)  
b. *The police saw the robbers.*  
(− agentive)

(197) a. *The police waited for the robbers.*  
(− affected)  
b. *The police looked for the robbers.*  
(− affected)

Of course, most verbs in English do not participate in this construction, but we still see evidence that the semantic conditions evident in the conative alternation carry through to other predicates in which objects are marked with prepositions. In these cases the prepositional object shows a lower degree of affect as a result of the predicate than it would have as the object of a primary transitive verb such as *kill,* *break,* etc.

As implied earlier in the discussion of the semantic factors that govern the splits in semantic alignment systems (see section 2.5), the semantic factors that underlie the choice of bivalent verb classes are the same as those we find in ‘semantic alignment’ systems (Blume 1998, Testelec 1998, Tsunoda 1981, 1985, 1999, and in
Semantic alignment systems: what’s what, and what’s not

Table 2.12. (Pseudo-)transitive verb types

<table>
<thead>
<tr>
<th>Control</th>
<th>Affectedness</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>A, P</td>
<td>speak to</td>
</tr>
<tr>
<td>II</td>
<td>A, P</td>
<td>fight with</td>
</tr>
<tr>
<td>III</td>
<td>A &gt; P</td>
<td>ask, threaten</td>
</tr>
<tr>
<td></td>
<td></td>
<td>follow, meet</td>
</tr>
<tr>
<td></td>
<td></td>
<td>agree, resist</td>
</tr>
<tr>
<td>IV</td>
<td>A</td>
<td>make, write</td>
</tr>
<tr>
<td>V</td>
<td>A</td>
<td>pull, take</td>
</tr>
<tr>
<td>VI</td>
<td>(A)</td>
<td>praise, search</td>
</tr>
<tr>
<td>VII</td>
<td>A, P</td>
<td>stick to</td>
</tr>
<tr>
<td>VIII</td>
<td>A (?)</td>
<td>see, depend on</td>
</tr>
</tbody>
</table>

Table 2.13. Singular agreement prefixes in Toba

<table>
<thead>
<tr>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>na, ne, ni</td>
<td>s, ja, ji</td>
</tr>
<tr>
<td>2sg</td>
<td>ʔan, ʔana</td>
<td>ʔaw, ʔawa, ʔa(ri)</td>
</tr>
<tr>
<td>3sg</td>
<td>n, na</td>
<td>r, ri, ʔa, ya, yi</td>
</tr>
</tbody>
</table>

passing Donohue 1998). Table 2.12 (adapted from Testelec 1998) illustrates a range of different verb classes that can be distinguished when discussing bivalent verbs.

These same categories are often reflected in monovalent constructions. Toba (Argentina and Brazil, Guaykuruan: Manelis Klein 2001) has three prefixes that are used to mark either S or A (see Vidal, this volume, for a discussion of the related language Pilagá). Just as Kolana can be characterized as displaying a three-way-split absolutive system, Toba can be thought of as exhibiting a three-way-split nominative system, but with complications. Of the three agreement prefix sets, Set II is the most common set; set III is only found with three verbs, ‘be afraid’, ‘want to go’, and ‘reach for’. Set I is complex, and is used for predicates that involve some level of (loosely termed) ‘centrifugal activity’: ‘direction inward towards the body’, ‘reflexivity, patient orientation’, or ‘reciprocity’ (as well as marking possession on nominals).

Toba

(199) な- pilottak

1sg.1-wash

'I’m washing (myself).'

(200) s-iyoGon

1sg.1i-wash.feet

'I wash my feet.'
74

Mark Donohue

(200) \( \text{ni-waGagatet} \quad \text{so-\text{-}iriGawa.} \)
1sg.1-am.fighting cl-several.people
'I was fighting with several people (and they were fighting with me).'

(201) \( \text{so-waGagatet} \)
1sg.11-am.fighting
'I’m hitting someone (with something like a whip, an outward motion).'

(202) \( \text{hayem ni-paGana} \quad \text{naroq\text{-}ilaqtak} \)
I 1sg.1-educate white.man’s.tongue
'I study Spanish.'

(203) \( \text{hayem sa-paGanek} \quad \text{naroq\text{-}ilaqtak} \)
I 1sg.11-educate white.man’s.tongue
'I speak Spanish.'

2.8 Semantic alignment: a summary

What can we conclude about semantic alignment? We have said that semantic alignment is:

- a split in the morphological encoding of arguments according to some feature of the lexical semantics of the verb.

The factors that determine these splits tend to involve something to do with agency, affectedness, or lexical aspect. It is also clear that monovalent verbs do not have an exclusive licence on variation according to these parameters; many, if not all, languages have more than one way of encoding two arguments of bivalent predicates. Furthermore, the ways in which variation according to these parameters is marked are often identical for both monovalent and polyvalent predicates; that is, the same case or agreement choices are found with the arguments of both monovalent and bivalent predicates. We therefore need to note that

- Variation in degree of affectedness can be located on the P argument of a polyvalent verb just as well as it can be on the S argument of a monovalent verb.
- Variation in degree of agency can be located on the A argument of a polyvalent verb just as well as it can be on the S argument of a monovalent verb.

Alternations such as the conative in English, or ‘quirky case’ objects (or subjects), are simply a special case of the same explicit semantic marking found with semantically aligned languages. Or, put another way, the phenomenon known as split intransitivity is a special case of the kind of semantically explicit subsystems that are found in most languages, although generally confined to marking one or other argument of a bivalent predicate. The fact that there is in some cases more
than one possible coding choice for the sole argument of a monovalent verb is a natural consequence of the fact that most languages have more than one, or even two, ways to code the arguments of bivalent predicates. This multiplicity of coding options is clearly a challenge to configurational explanations of semantic alignment (‘unaccusativity’) which assume that the split in coding reflects two structural positions, one being that of a ‘normal’ subject and one that of an object.
3

Split intransitives, experiencer objects, and ‘transimpersonal’ constructions: (re-)establishing the connection

ANDREJ MALCHUKOV

3.1 Introduction*

In his review of Uhlenbeck (1916), Edward Sapir (1917) suggested that ‘inactive’ (object-inflecting) intransitive verbs in Native American languages are better analysed as transitives: “Thus, forms like “I sleep” or “I think” could be understood as meaning properly “it sleeps me”, “It seems to me” ” (Sapir 1917: 85). In essence, Sapir suggests that such constructions (i.e. the type of construction labelled ‘transimpersonal’ by Haas 1941) are transitive impersonal constructions with object experiencers. In the typological literature dealing with the phenomenon of split intransitivity this insight has not been pursued,* since Sapir’s proposal faces some problems both on the functional and on the formal side. Thus, Merlan (1985: 327) points out that Sapir’s analysis is functionally obscure as it is not clear ‘how any form can mean “it sleeps me”’. Further, there are also formal differences between the constructions under discussion. Indeed, within the patient subject construction in a split intransitive system the verb is intransitive by definition and takes an experiencer-patient as subject, while within the

* I am very grateful to the editors for the helpful comments on the earlier version of this chapter. I am also indebted to audiences in Leipzig, Nijmegen, and St Petersburg for their comments; special thanks are due to Balthasar Bíckel, Martin Haspelmath, Helen de Hoop, A.E. Kibrik, Akio Ogawa, Matt Shibatani, Anna Siewierska, Leon Stassen, and Peter de Swart, as well as to Nick Evans for the discussion of the Iwaidjan data. I acknowledge the Netherlands Organization for Scientific Research (NWO) (grant 220–70–003 for the PIONIER project ‘Case cross-linguistically’), as well as the Japan Society for the Promotion of Science (JSPS) for financial support.

1 It might be interesting to note that some analyses of patitive subjects within the generative tradition (related to the work on unaccusativity) are reminiscent of transimpersonal analysis in that they treat patitive subjects as (underlying) objects (see also Wichmann, this volume, for a similar observation).
transimpersonal experiencer constructions the verb is transitive and takes an experiencer as its object. Finally, it is not clear how general Sapir’s analysis is conceived to be, that is, whether it is meant to pertain to all split-intransitive languages. Due to the work of Merlan (1985), Mithun (1991), and Nichols (1992), typologists have become increasingly aware of the heterogeneity of split-S languages, which can be classified on different criteria. On the one hand, split intransitivity may be related to differences either in role properties of the subject or in lexical aspect (cf. the distinction between agent/patient and active/stative languages in Mithun 1991). Second, split intransitivity can have either an ‘accusative base’, with the patient subject construction as a minor pattern, or an ‘ergative base’, with the agent subject construction as a minor pattern (Nichols 1992; cf. Merlan 1985).

Yet Sapir’s analysis is appealing, since transimpersonal experiencer constructions and patient subject constructions indeed involve significant similarities. First, both constructions usually involve experiential predicates pertaining to emotions and sensations (as a transimpersonal construction such as German *Mich friert* [me freezes] ‘I freeze’ illustrates). Secondly, the formal similarities between these constructions, where experiencers are usually cross-referenced through the same inflection used to cross-reference objects, cannot be overlooked. Thirdly, as will be clear from the subsequent discussion, the distinction between experiencer object constructions and experiencer subject constructions is often not clear-cut, as the former have been often alternatively analysed as non-canonical subjects (see e.g. the papers in Aikhenvald et al. 2001). In what follows I shall provide evidence to the effect that Sapir’s analysis can be upheld with certain qualifications. I shall suggest that the connection between the two types of construction should be conceived in a diachronic rather than synchronic perspective (which would counter Merlan’s objection concerning the functional inconsistency of this analysis in synchronic terms). That is, I shall argue that in a number of cases, experiencer subject constructions have evolved from the reanalysis of transimpersonal experiencer constructions. The second qualification pertains to the generality of the analysis: it will be claimed that this diachronic scenario is best applied to languages where patientive intransitives constitute a minor class (for split intransitives with an ‘accusative base’, in Nichols’s terms), and where the split is role-based rather than aspect-based. This qualification, however, still leaves a large share of split-intransitive languages within the purview of my analysis, since languages with the agentive/patientive split are more frequent than languages with the active/stative split (Mithun 1991, 1999: 220), and also because split intransitive systems with an accusative base are more frequent than those with an ergative base (Nichols 1992: 105).

In this chapter I shall provide evidence for a diachronic connection between transimpersonal experiencer construction and the patient subject construction by focusing on the cases when the distinction between the two constructions is not clear-cut. In section 3.2 I shall present evidence from a number of Native
American languages, where the Split-S pattern arguably arose from the reanalysis of transimpersonal experiencer constructions. This section shows similarities both in data and in interpretation with the contribution by Mithun (this volume) on the origins of agentive systems in the languages of North America. In section 3.3 I give additional evidence for the connection from Papuan languages where object experiencers gradually develop into non-canonical subjects. Section 3.4 broadens the scope of the discussion, extending it to other cases where transimpersonal and object experiencer constructions are subject to reanalysis, but where reanalysis does not yield a split-intransitive pattern. Generally, it will be argued that split intransitivity arises through a conspiracy of universal functional motivations and language-specific structural properties. Section 3.5 summarizes the main results of the chapter.

3.2 Reanalysis of the transimpersonal experiencer construction into the patient subject construction in Native American languages

3.2.1 Basic and extended uses of transimpersonal constructions: Slave and Eskimo

In this section I shall provide some data from languages of the Americas where patientive intransitives have evolved from transimpersonal experiencer verbs. To avoid misunderstanding, it should be mentioned from the outset that transimpersonal constructions are not intrinsically connected to split intransitivity. Indeed, in many Native American languages we find transimpersonal constructions which are clearly structurally distinct from the patientive pattern in a split-intransitive system. Consider examples (1) and (2) from Slave, where an unspecified subject—either intransitive, as in (1), or transitive as in (2)—is encoded by the special ‘indefinite subject’ agreement marker -tse-:

\[
\begin{align*}
\text{Slave (Rice 1989: 1020)} \\
(1) \quad \text{ts'ejji} & \quad (2) \quad \text{k'ina-ts'e-reyo} \\
\text{‘Someone is singing.’} & \quad \text{‘S/he is chased’} \\
& \quad \text{‘(lit. someone chased him/her).’}
\end{align*}
\]

The construction in (2) is clearly distinct from the patientive pattern; in particular, the verb morphology remains transitive. Yet it is symptomatic that the transitive construction with indefinite A has been rendered by Rice through the passive gloss ‘s/he is chased’; this suggests that a transimpersonal construction deviates functionally from the canonical transitive construction.

In some other cases, however, the distinction between transimpersonal and patientive intransitive constructions is less clear-cut. Consider the situation in Eskimo languages. In West Greenlandic (Fortescue 1984: 59–61), verbs referring to time and weather expressions (e.g. ukiuli- ‘become winter’, api- ‘snow’) are impersonal. The understood subject referring to some natural force is not expressed lexically but is cross-referenced by the default 3sg agreement suffix.
Importantly in the present context, some impersonal verbs are used transitively, as in (3):

West Greenlandic (Fortescue 1984: 81)

(3) **Anurliliup-paatigut**

storm-INDF.3SG.A>1PL.P

'We were caught by storm.'

In this example the verb takes the portmanteau transitive agreement suffix -paatigut indicating a 3rd person singular subject acting upon a 1st person plural object (so a more literal translation of this sentence would be 'It stormed us').

In some other Eskimo varieties, however, transimpersonals have extended their use far beyond weather expressions. Thus, in (Siberian) Yupik (Emeljanova 1967, Vaxtin 1995), trans impersonal inflection can be used with any intransitive verb to indicate the lack of control on the part of the subject. Consider the following two examples, where (4) uses the regular intransitive inflection and is noncommittal with regard to volitionality/control, while (5) uses the transitive (trans impersonal) inflection and indicates that action is spontaneous/unintended:

Yupik (Emeljanova 1967: 89)

(4) **Taqnyxaq awxsaxtaq-a**

child.ABS crawl.start-INDF.3SG.S

'The child starts to crawl.'

(5) **Taqnyxaq awxsaxtaq-a**

child.ABS crawl.start-INDF.3SG.A>3SG.P

'The child starts to crawl (unintentionally).'

According to Vaxtin (1995: 46), the trans impersonal construction is used to refer to spontaneous, inactive, often unexpected or adversative events. Consider the following examples, where the use of the trans impersonal construction in (7) implies that an event is unexpected:

Yupik (Vaxtin 1995: 46)

(6) **Ityxta naqu-a**

mother.my come-INDF.3SG.S

'My mother came.'

(7) **(S’una) ityxta naq-a**

suddenly mother.my come-INDF.3SG.A>3SG.P

'Suddenly my mother entered.'

Functionally, the alternation of intransitive and trans impersonal constructions as illustrated above is similar to split of the agent/patient type, in that a large number of verbs allow for more than one coding option, one of which resembles
the coding used for objects (more specifically, this alternation is reminiscent of 'fluid-S' languages in terms of Dixon 1994). Structurally, however, the transimpersonal pattern is distinct from canonical cases of the patientive-S pattern, since the verbal agreement remains unmistakably transitive. Yet, as will be clear from the following discussion, the distinction between transimpersonal and patientive-S constructions is often not clear-cut (see also Donohue, this volume).

### 3.2.2 Transimpersonals reanalysed as patientive intransitives: Tunica, Koasati, and Yurakaré

A next step in the reanalysis of transimpersonal constructions is attested in Tunica (Haas 1941; cf. Merlan 1985). Tunica is described by Haas as a split-intransitive language, with object-inflecting experiencer verbs constituting a minor pattern. Tunica also has a distinct construction termed 'trans impersonal' by Haas, referring to events originating from some natural force, and where the understood subject is cross-referenced with 3rd person singular feminine agreement. Interestingly, object-inflecting verbs in the inchoative form are constructed as transimpersonals, and some of these verbs occur only in this form. On the basis of these data, Haas concluded that in Tunica the object-inflecting experiential verbs developed from transimpersonals (Haas 1941: 59). Haas also provides evidence that other involuntary action verbs were earlier constructed as transimpersonals, but were later assimilated to plain intransitives (see Mithun, this volume, for further discussion of Tunica and exemplification).

A similar case of reanalysis has arguably occurred in Koasati, as reported by Kimball (1991). Koasati, like many other Muskogean languages, is usually characterized as split intransitive on the basis of its agreement system, which distinguishes agent ('subject') agreement, patient ('object') agreement, and dative agreement series (cf. Mithun 1999: 237–8). Agreement with agentive and patientive subjects is illustrated in (8) and (10), respectively, and should be compared to the transitive verb in (9):

Koasati (Kimball 1991: 260, 251).

(8)  nokće:ba-lí-t
     stop-1sg.A-PAST
     'I stopped.'

(9)  ca-nokće:bá-ci-t
     1sg.P-stop-caus-PAST
     '(It) stopped me.'

(10) ca-libátli-t
     1sg.P-burn-PAST
     'I got burned.'

The overall alignment system of Koasati is complicated by the fact that NP case operates on an accusative basis, and the overall pattern can be properly
characterized as a combination of agentive and accusative systems (Mithun 1999: 236–8). The patientive subject in (10), if expressed by a pronoun, takes the nominative case, as would the agentive subject in (8). There is another reason why Kimball is reluctant to call Koasati split intransitive. He notes that, morphologically, patient subject verbs look like plain transitives; compare the patientive S construction in (10) above with the transitive construction (11):

Koasati (Kimball 1991: 251)

(11) Nihahcí ikba-k ca-libáli-t
    grease  hot-NOM 1sg.p-burn-PAST
    'The hot grease burned me.'

Note that in both (10) and (11) the patient/experiencer is cross-referenced through the same inflection that is used for most objects (3sg ‘subject’ agreement is not overt). Indeed, if we disregard for a moment the nominative case on the subject in (10), one could analyse the purported intransitive construction as transitive with an omitted unspecified 3rd person subject (i.e. ‘It burned me’). It is not clear from Kimball’s description how many patient subject verbs also have transitive uses (note, incidentally, that the verb -nokcoba- ‘stop’ as exemplified in (8) and (9) is also labile). Generally, however, we can follow Kimball in his suggestion that diachronically the patientive subject pattern originated from reanalysis of transitive forms with a 3rd person (singular) agent (see Kimball 1991 for additional arguments for the transitive origin of stative verbs based on parallelism in derivation of the negative forms and forms of verbal nouns).

A somewhat similar development can be attested for the experiencer object verbs in the Amazonian language Yurakaré, although reanalysis has not proceeded as far here (data on Yurakaré stems from van Gijn 2005: 163–4 and p.c.). In Yurakaré, experiencer verbs referring to emotions/sensations take object experiencers. In a few cases they cross-reference the experiencer through the regular (direct) object agreement prefixes (as in (12)); more often through the use of object agreement in combination with different applicative markers (sometimes fused with the agreement markers, as in the case of the ’cooperative’ agreement markers in (13)):

Yurakaré (van Gijn 2005: 163–4)

(12) ti-justu-ø
    1sg.p-want-3sg.a
    'I want it (lit. it wants/attracts me).'

(13) té-dyumme-ø
    1sg.coop-cold/freeze-3sg.a
    'I am cold (lit. it is cold with me).'

The latter cases are better analysed as constructions with an (oblique) object experiencer, as there is no compelling evidence that they have been reanalysed into
the subject experiencer construction. In the case illustrated by (12) the situation is not as clear, though. Indeed, currently, the verb kusu ∼ jusu cannot be regarded as a regular transitive, as it cannot take the whole person paradigm of agreement suffixes to refer to the subject (the stimulus) (van Gijn 2005: 165). Interestingly, there are indications that around a generation or two ago this verb was a regular transitive verb capable of taking a full series of subject agreement suffixes, including the 1st and 2nd person (van Gijn 2005: 165, citing M. Day):

(14) nish ta-jusu-m
    NEG 1PL-P-WANT-2SG.A

'We do not want you' (or you don’t attract us).

Since in the Yurakaré dialect studied by van Gijn such a construction is ungrammatical, van Gijn concludes that it has been reanalysed into a transimpersonal or possibly an intransitive construction with a subject experiencer (van Gijn 2005: 165). The final decision is difficult to make, since, unlike Koasati, Yurakaré lacks case marking of core arguments, and some of the other syntactic tests for subjecthood are not applicable. The examples from Yurakaré further demonstrate that the functional reanalysis of transitive verbs can already occur at the transimpersonal stage: notably, van Gijn invariably translates the verb kusu ∼ jusu as ‘want’, although syntactically the translation ‘attract’ would be more appropriate.

The data from Yurakaré can also illustrate the point that a patientive intransitive of a transimpersonal origin need not display lability, as the transitive use can be lost. Furthermore, establishing a diachronic connection between transimpersonals and patientive intransitives in a split-intransitive language does not amount to stating that every single patientive intransitive had a transitive use at an earlier stage, since some verbs could be assimilated to a transimpersonal/patientive class by analogy (cf. e.g. the expansion of the transimpersonal construction in Eskimo languages, discussed in 3.2.1 above, where analogy must have played an important role). On the other hand, lability on the part of a sizeable group of patientive intransitives is indeed symptomatic of their transimpersonal origin.

The diachronic connection between patientive intransitives and transimpersonals can also explain the prominence of the class of labile verbs in many languages displaying split intransitivity, which Klimov (1977) even considered as a characteristic property of active-stative languages in general.

3.2.3 Convergence of transimpersonal and patientive-S constructions in other Split-S languages

There are two features of the verb structure in Koasati and Yurakaré that contribute to the conflation between transitive and intransitive patterns and facilitate the reanalysis of the former into the latter: the availability of object agreement on the verb and the absence of an overt 3SG subject agreement marker. Now, it should be noted that in this regard these two languages are not at all exceptional among split-intransitive languages, which commonly manifest both features. In
Intransitives, objects, and ‘transimpersonal’ constructions

fact, object/patient agreement is generally acknowledged to be a typical feature of languages of the active type, which are typically radically head-marking (Nichols 1992). The absence of overt marking for 3rd person A arguments has also been noted as a characteristic property of active languages (Klimov 1977: 226). These two features will be illustrated in the remainder of this section for a selection of Amerindian languages showing split intransitivity (see Mithun, this volume, for further examples from North American languages).

Dakota (Lakota), which has figured prominently in the discussion of split-intransitive languages, is representative in both respects. Note that the absence of a formal marker for the 3rd person agent in combination with overt patient agreement results in a formal identity of intransitive verbs with a patientive argument (as in (15a)) and transitive verbs with a 3rd person agent in (15b):

\[
\begin{align*}
\text{Dakota (Boas and Deloria 1941: 76)} \\
(15) & \quad \text{a. } ni-\text{s'\text{ca}} \quad \text{b. } ni-\text{k\text{te'}} \\
& \quad \text{2pl.p-bad} \quad \text{2pl.p-kill} \\
& \quad \text{‘You are bad.’} \quad \text{‘He kills you.’}
\end{align*}
\]

The structural convergence of intransitive and transitive structures observed above for Dakota is in fact typical of Siouan languages in general, where 3rd person agents are commonly unmarked and patientive arguments of intransitives are cross-referenced through the patient (‘object’) inflection (Mithun 1999: 508).

Similarly, in Ika, patientive intransitives, as in (16a), are structurally indistinguishable from transitives with a 3rd person agent, as in (16b):

\[
\begin{align*}
\text{Ika (Frank 1985: 11)} \\
(16) & \quad \text{a. } Na-\text{'\text{tikuma-na} } \quad \text{b. } N\text{â-tsua-na} \\
& \quad 1sg.p-\text{forget-dist} \quad 1sg.p-\text{see-dist} \\
& \quad \text{‘I forgot.’} \quad \text{‘He saw me.’}
\end{align*}
\]

In (colloquial) Guaraní these patterns are again identical, although here the absence of agent agreement on a transitive verb is restricted to cases when the patient/experiencer is 1st or 2nd person:

\[
\begin{align*}
\text{Guarani (Gregores and Suarez 1967: 131)} \\
(17) & \quad \text{a. } \text{\text{še-\text{manu\text{'\text{\d{a}}} } b. } \text{\text{še-pet\text{'}}} \\
& \quad 1sg.p-\text{remember} \quad 1sg.p-\text{hit} \\
& \quad \text{‘I remember.’} \quad \text{‘He (it/they) hit me.’}
\end{align*}
\]

For Kiowa, as for Guaraní, it is generally not true that the 3rd person agent marker is zero; however, in certain transitive and ditransitive patterns this marker is left unexpressed, since there are restrictions on the number of arguments which

\footnote{Even though it has no bearing on the proposed analysis, it should be mentioned for the sake of completeness that most Split-S languages which have zero agreement for 3rd person As have zero agreement for 3rd person Ps as well (A. Siewierska, p.c.).}
can be simultaneously encoded on the verb. Consider the following pattern of experiencer verbs which show a restricted loss of the A-marking:

Kiowa (Watkins 1984: 137)

(18)  yq-táy
     (2,3SG.A+)1SG.P+PL.THEM-awake

‘I awoke/something woke me up.’

Watkins (1984: 136) qualifies such experiencer verbs as intransitive (which would make Kiowa a split-intransitive language), although the pattern is clearly modelled on the transitive, or rather a ditransitive pattern. Note that the verb in (18) cross-references both the patient-addressee (P) and a dummy (plural) object-theme (THEM) (the terminology here reflects that fact that Kiowa distinguishes between primary and secondary objects rather than direct and indirect objects). Compare (18) with the ditransitive construction (19):

Kiowa (Watkins 1984: 138)

(19)  Kút yq-kón
     book (3,2SG.A+)1SG.P+PL.THEM-bring

‘He brought me the books.’

Note that similarly to the experiencer verb in (18), in (19) both objects—patient-addressee and theme—are explicitly encoded, while the agent is left unexpressed, but ‘is implied’ (Watkins 1984: 114).

Broadening the perspective from split-S languages with a split manifested in agreement to those with a split in the pronominal system, one can also find languages where the 3rd person agent pronouns are lacking. A relevant case is Haida (Enrico 2003: 93), which displays split intransitivity in free pronoun marking (see also Mithun, this volume, and Donohue, this volume, for further discussion of Haida and exemplification). Significantly, there are no overt inanimate (‘low-potency’) agent pronouns in Haida, which makes an intransitive clause with a patient argument indistinguishable from the transitive clause with an unexpressed (inanimate) agent. The use of zero pronouns for inanimate/non-prominent agents/subjects is of course not confined to split intransitive languages; it has been reported for Niuean and Mixtec, to name but a few cases. In the present context, it is important that absence of an overt 3rd person marking for (non-prominent) agents, coupled with the presence of patient agreement on the verb, leads to convergence of intransitive and transitive structures.

3.2.4 Transimpersonals and transitivity decrease: Navajo and Tlingit

Above, we have noted that absence of an overt marker for the agent/subject argument blurs the morphological distinction between transitives and intransitives and thus opens the way for syntactic reanalysis. However, a language may possess other means for distinguishing between transitive and intransitive clauses. Navajo
Intransitives, objects, and ‘transimpersonal’ constructions

may serve as an example, even though it is accusative rather than split intransitive in its agreement pattern. Navajo makes a strict distinction between transitive and intransitive verbs through the use of different valency markers (called ‘classifiers’ in the literature). In the canonical transitive construction with a definite subject the verb takes a transitive classifier:

Navajo (Kibrik 1996: 291, citing Young and Morgan 1987)

(20)  Né-i-a-l-zho?
    MOD-3SG.P-3SG.A-TR-hunt.ITER
    ‘He repeatedly hunts it.’

When the transitive subject is indefinite, as in (21), the verb may be alternatively marked by an intransitive classifier:

Navajo (Kibrik 1996: 291)

(21)  Ná-a-ðá-l-zho?
    MOD-3SG.P-INDEA-DETR-hunt.ITER
    ‘Someone repeatedly hunts it.’

Kibrik (1996) concludes on the basis of such examples that A-(in)definiteness is an important feature contributing to canonical transitivity in Navajo. More generally, he proposes to add the parameter of subject individuation to the list of Hopper and Thompson’s (1980) transitivity parameters, along with the parameter of object individuation. In the present context the important point is that, since transitive constructions with an indefinite subject deviate from the transitive prototype, they can pattern intransitively even in languages which otherwise make a strict distinction between transitive and intransitive verb patterns. In the case of Navajo, it is not the marking of morphological transitivity but rather the presence of an overt marker of (indefinite) subject that prevents us from analysing the verb in (21) as intransitive. Interestingly, according to Mithun (this volume), in Tlingit, which is a distant relative of Navajo, the transimpersonal construction has been reanalysed as intransitive with a subject experiencer. This development, along with the absence of the 3rd person subject marker on transitive verbs, has led to convergence between transitive and intransitive constructions, and has ultimately resulted in an alignment shift from a nominative-accusative system to an agentive-patientive system.

3.2.5 Conclusions

In this section we considered examples of split-intransitive patterns which originated from the realanalysis of transimpersonal experiencer constructions. In some cases the transitive origin of patientive intransitives is obvious, as the verb retains the agent-subject agreement marker (cf. examples from Yupik Eskimo, but also Tunica). The same development can also be suggested for other languages, where in the absence of an overt 3rd person subject marker the distinction between
transitives and intransitives is blurred. Thus, in Koasati, some patientive intransitives could alternatively be analysed as transitives, provided that the transitive use is attested elsewhere. Here it is only the presence of case marking on nouns that indicates that reanalysis into a construction with subject experiencers has indeed taken place. In most split-intransitive languages under discussion case morphology is, however, lacking, which leads to a structural ambiguity in cases where a verb takes a patient/object inflection, and an agent/subject marker is absent. Admittedly, some of the examples of split-intransitive languages discussed in 3.2.3 above provide stronger cases for this path of reanalysis than others. Thus, the reanalysis scenario is more plausible for Ika and Kiowa, since in both languages patientive intransitives represent a minor pattern, and split intransitivity is role-related rather than aspect-related. In other cases this connection is less transparent. For example, for Dakota this analysis is less straightforward, as patientive intransitives do not constitute a closed class. For Guarani, such an interpretation is complicated by the fact that split intransitivity is aspect-related rather than role-related (Mithun 1991). This does not exclude the possibility that split intransitivity may be historically related to a transitive pattern here as well. After all, patientive marking may spread by analogy to substantial parts of the verbal lexicon (cf. Dixon 1994: 187ff. for some discussion), and the semantic basis for split intransitivity may shift over time (in fact, Mithun 1991 has suggested that a shift from agentive-patientive to active-stative systems has occurred in Iroquoian, and possibly also in Tupi-Guarani). For such less clear cases, comparative data is necessary to determine the origins of split intransitivity (see Mithun, this volume, for comparative data from a number of families in the North America supporting the conclusion that patientive intransitives developed from a transitive pattern).

In other languages, however, the connection between these two patterns is still obvious; what is at issue is rather whether we are dealing with a split intransitive or still a transitive (transimpersonal) pattern, as in the case of Yurakaré. Similarly to Mithun (this volume), I have emphasized the role of the structural factors which facilitate reanalysis of transimpersonal verbs into patientive intransitives. As noted above, two important structural factors are the marking of object agreement on the verb, on the one hand, and the absence of an overt 3sg subject marker, on the other. The latter factor contributes to the conflation of transitive and intransitive patterns and thus opens the way for reanalysis. On the other hand, if object agreement is missing, it is more difficult to identify the split in subject encoding (note, however, the small group of languages, like Haida, that have a split

3 Note, however, that at least some patient subject verbs in Dakota are unmistakably of transimpersonal origin. Thus, some of the verbs with the applicative (‘instrumental’) prefixes can be used both intransitively and transitively ‘with an indefinite actor understood’ (Boas and Deloria 1941: 47). To take one example (which could have inspired Sapir’s analysis): ma-ka-ist’me [1sg.P-app.-sleep] ‘I have fallen asleep; he/it/circumstances put me to sleep’ (Boas and Deloria 1941: 47). Some of these verbs are exclusively object-inflecting, but some can take either objective or subjective prefixes to refer to an experiencer. This is indicative of the process of reanalysis of the type observed for Tunica.
in their free pronouns). Mithun (this volume) also notes that language-external factors, such as language contact, can promote reanalysis of accusative pattern into an agentive one. Of course, in this as in other cases (see Heine and Kuteva 2005 for general discussion), external and internal factors in language change need not exclude each other.

3.3 From experiencer object constructions to Split-S in Papuan languages

In the previous sections we have seen that in a number of languages in the Americas the split-intransitive pattern developed from reanalysis of the transimpersonal experiencer construction. The Papuan languages considered in this section also provide evidence for the tendency of experiencer objects to be reanalysed as non-canonical subjects. This path of reanalysis is clearly similar to what has been observed for Native American languages, although the source construction does not always qualify as transimpersonal proper. The occurrence of this type of reanalysis will be demonstrated by way of comparison with experiencer object constructions in a number of Papuan languages where reanalysis has proceeded to varying degrees.

3.3.1 From transitive experiencer object construction to transimpersonal: Usan and Amele

In Usan, as in many other Papuan languages (Foley 1986), most experiential verbs take object experiencers (Reesink 1987: 139). The experiential construction in Usan is based on a semi-auxiliary verb having the literal meaning ‘strike, shoot’, which takes an experiencer as its object and a stimulus as its subject. Note in (22) that the experiencer is cross-referenced by the object agreement prefixes and that the stimulus is cross-referenced by the 3rd person subject marker:

\[
\text{Usan (Reesink 1987: 139)}
\]

\[
(22) \quad \text{Munon isig toar wA-r-a in-Ab igo}
\]

\[
\text{man old sickness him-shoot-3sg.DS lie-ss be-3sg.PRES}
\]

‘The old man is sick and lying down.’

Reesink notes that structurally these verbs are similar to ordinary transitives, except for the fact that the experiencer invariably appears in the clause-initial topic position (unlike other objects). Note also that since Usan—like the other Papuan languages discussed in this section—lacks case morphology, the form of a nominal argument reveals nothing of its syntactic function.

A similar construction has been reported by Roberts (1987) for Amele, which is closely related to Usan.
In (23), too, the experiencer is cross-referenced on the (auxiliary) verb as its syntactic object, and the subject agreement is always 3rd person singular. Again as in Usan, the experiencer must appear in the clause-initial subject/topic position. However, in Amele this construction reveals further features of grammaticalization or reanalysis: in particular, the verb stem is truncated, and some nominals referring to the stimulus do not occur outside this construction at all (Roberts 1987: 316). Further, the experiencer reveals certain properties of the syntactic subject: apart from the positional property of appearing sentence-initially, for example, it can control switch-reference marking on a preceding medial verb. Consider the following example, where the medial verb takes the same-subject form to indicate coreferentiality of the previous subject with the experiencer of the second clause:

Amele (Roberts 1987: 166)

(24) Ija bi-bi-g *wen* te-i-a
    1SG simul-come.up-1SG-SS hunger 1SG.P-3SG.A-PAST
    ‘As I came up I became hungry.’

In fact, as noted by Roberts (2001), the experiencer displays most syntactic subject properties, both ‘interclausal’ (e.g. control of switch reference) and ‘intraclausal’ (e.g. reflexivization). These facts support Robert’s conclusion that the experiencer in these constructions is in fact a non-canonical subject. Note, however, that it is still cross-referenced by the object agreement markers, which make this pattern similar to a transimpersonal construction, and equally similar to patient-object construction in a split-intransitive language. Thus the syntactic status of the experiencer is ambiguous at the intermediate stages of reanalysis.

3.3.2 From transimpersonal to experiencer subject construction: Tanya and Galela

While the languages considered above still qualify as transimpersonal, some other Papuan languages have gone further in the reanalysis. In Hua, ‘grammatical pressure to identify the experiencer of impersonal verbs as a subject’ manifests itself in the fact that some of experiencer object verbs like auiahu- ‘be ashamed’ pattern
as transimpersonals when in the indicative mood, but as intransitive experiencer subject verbs in the imperative mood (Haiman 1980: 359). Another scenario of reanalysis has been reported for experiential constructions in Tauya (MacDonald 1990). In this language, experiential verbs usually take experiencer as object in a transimpersonal pattern.

Tauya (MacDonald 1990: 187)

(25) ya-sepame-a-ʔa
   1SG.P-sick-3SG.A-IND
   ‘I am sick’ (lit. it sickens me).

However, these verbs can appear, albeit marginally, in another construction, where subject agreement, rather than being 3sg by default, doubles the object agreement, redundantly cross-referencing the experiencer argument. Note that here the subject marker is chosen to match the features of P instead of taking the regular 3sg subject prefix -a-:

Tauya (MacDonald 1990: 191)

(26) Ni-pa sen-foʔu-fe-ene-ʔa
   eat-SS 2PL.P-full-TR-1/2PL.A-IND
   ‘We ate until we were full.’

A further and final step of reanalysis is attested in the West Papuan languages Galela, Tobelo, and Sahu, as described by Holton (this volume). In these languages the dummy subject agreement prefix in the transimpersonal experiencer object pattern undergoes phonetic erosion, resulting in its optionality and ultimate loss on certain verbs:

Galela (Holton 2005)

(27) (*i-)ni-kiola
   (3SG.A)-2SG.P-asleep
   ‘You are asleep’

Note that this pattern is similar to canonical cases of split intransitivity, as the verb is object-inflecting, but intransitive. As noted by Holton, many of these verbs cannot be used transitively any more (recall the discussion of Yurakaré in 3.2.2). Holton concludes that split intransitivity in the West Papuan languages of North Halmahera is a recent phenomenon which developed from reanalysis of a transitive pattern aided by processes of phonetic erosion.

3.3.3 Conclusions

Papuan languages also reveal an ongoing process of reanalysis of the construction with object experiencers into a construction with subject experiencers. At intermediate stages of reanalysis the construction becomes transimpersonal, even if it started as a cognate subject construction with an inanimate stimulus (as in Usan).
At more advanced stages of reanalysis, the experiencer acquires syntactic subject properties, but transitive agreement morphology is retained. Thus, experiencer verbs in Amele are similar to a patientive intransitive construction inasmuch as the experiencer is cross-referenced by the object agreement markers, but distinct inasmuch as the agreement pattern remains transitive (transimpersonal). At later stages of reanalysis the default agreement, which has become functionally dispensable, is either deleted (as in the languages of North Halmahera), or else is put to other uses, as in Tauya.

Papuan languages also quite clearly show that the rise of split intransitivity is partly due to the fact that morphological reanalysis lags behind syntactic reanalysis in the processes of language change. This tendency, repeatedly observed in the functional typological literature (see Keenan 1976, Cole et al. 1980, Givón 1997b on the gradual ‘acquisition of subjecthood’), is borne out by the Papuan data. The process of reanalysis starts with the acquisition of topicality-related properties (e.g. linear position), followed by syntactic (‘behavioural’ in terms of Keenan) properties, and finally by coding properties pertaining, in particular, to agreement (recall that case morphology is lacking in the languages under discussion). At a more advanced stage of reanalysis this results in a pattern where the experiencer has acquired syntactic properties of subject (and therefore may qualify as ‘non-canonical subject’), but is still cross-referenced by object agreement, which is more conservative in the processes of language change (see also Donohue 2005b for more discussion of transimpersonal constructions in Papuan languages).

3.4 The evolution of transimpersonal constructions and experiencer object constructions: beyond split intransitivity

Above we have seen how transimpersonal experiencer constructions give rise to a split-intransitive pattern in various languages of the Americas and Papua New Guinea. The motivation behind this reanalysis is transparent, and is ultimately due to a well-known functional tendency to align the most prominent argument within a clause with the subject (cf. e.g. Givón 1990, Blake 1994). Thus, on the one hand, all experiencer object constructions are subject to a functional pressure to syntactically upgrade the experiencer, which is the most discourse-pragmatically prominent argument. On the other hand, in transimpersonal constructions, and more generally in constructions involving a non-prominent (i.e. indefinite-nonspecific, inanimate, or cognate) subject, the subject tends to be syntactically demoted or suppressed. Thus, the motivation behind the reanalysis of a transimpersonal experiencer construction is actually twofold.5

5 The same two factors may be responsible for the diachronic instability of experiencer object constructions that are not impersonal, and transimpersonal constructions that do not involve an experiential predicate.
However, as is clear from the discussion above, even if the functional pressures underlying the reanalysis of transimpersonal constructions are universal, these pressures yield a split-intransitive pattern only under specific structural conditions, such as the presence of object agreement and the non-overt expression of subject. If such conditions are not met, the functional pressures may lead to reanalysis, but not necessarily to a Split-S pattern. In what follows I shall consider different scenarios for reanalyses, leading to ‘subject absorption’ in experiencer object constructions in Iwaidjan (Ilgar), covert reanalysis of experiencer verbs in Germanic, partial realignment of transimpersonal experiencer object constructions in Himalayan, and evolution of indefinite subject constructions into passives, attested in many languages.

3.4.1 ‘Subject absorption’ in experiencer object construction in Iwaidjan

Consider the evolution of the experiencer object construction in Iwaidjan (a Non-Pama-Nyungan Australian family), discussed by Evans (2004; see also Walsh 1987). The initial stages of reanalysis are represented by the transitive construction with object experiencers reminiscent of the Papuan pattern, where the stimulus clearly qualifies as a transitive subject (Evans 2004: 170); cf. the following example from Ilgar:

Ilgar (Evans 2004: 170)

(28) Nga-ni-ma-ny wunyarru
1sg.p-3m.sg.a-get-past sickness
‘I got sick (lit. sickness got me).’

This type of experiencer construction, termed constructions with ‘subcategorized nominal subject’ by Evans, contrasts with constructions with a ‘frozen nominal subject’, where the status of the formal subject (the noun ngok in (29)), cross-referenced through the subject agreement, is problematic as it does not occur outside this construction:

Ilgar (Evans 2004: 170)

(29) Nga-ni-mi-ny ngok
1sg.p-3m.sg.a-get-past ngok
‘I am full.’

A further stage of reanalysis is attested in the impersonal construction, which cannot take a (lexical) subject at all; cf. (30):

Ilgar (Evans 2004: 170)

(30) Iny-ni-ngardbanbu-n
3f.sg.p-3m.sg.a-cause.headache-npast
‘She has a headache.’
Synchronously, these construction types can be seen as a ‘cline’ (Evans 2004), leading from the regular transitive to the impersonal construction. Diachronically they represent different stages of reanalysis. As noted by Evans (2004: 172), there is evidence that the ‘frozen subjects’ in constructions like (29) were once fully-fledged (subcategorized) subjects. Thus the noun ngok in (29), which is restricted to this construction in Ilgar, has cognates in other Iwaidjan languages with the meaning ‘guts; belly’, so the original meaning of (29) would be something like ‘(my) guts did/said (to) me’ (Evans 2004: 172). Although the origin of the construction with a dummy subject is more obscure, it arguably must have evolved from a subsequent reanalysis of the ‘frozen subject’ construction, mediated through a stage of noun incorporation. For example, the verb -ngardbanbu- in (30) originally meant ‘it-me-head-bites’ (Evans 2004: 173), which in its turn could originate from a construction meaning ‘my head bites me’ (cf. the construction ‘my tooth bites me’ used for ‘I have toothache’ in another Iwaidjan language Amurdak; Evans 2004: 181).

We have seen that Iwaidjan languages also present evidence for the diachronic instability of the experiencer object construction: the non-prominent inanimate subject is gradually ‘absorbed’, leaving the construction impersonal, as in (28)–(30). However, the result is clearly distinct from canonical cases of the patient subject constructions inasmuch as the verbal morphology remains transitive (cf. the extended use of transimpersonals in Papuan languages). Interestingly, in some cases reanalysis has proceeded further, since some verbs in Ilgar alternate between a transitive and an intransitive pattern:

Ilgar (Evans 2004: 182)

(31)  a. I-nga-mardalkanyi-ny arrkbi
   3M.SG.P-3F.SG.A-sneeze-PAST man
   ‘He is sneezing’ (lit. ‘it has made him sneeze’).

   b. I-mardalkanyi-ny arrkbi
   3M.SG.S s-sneeze-PAST man
   ‘He is sneezing.’

As these examples show, the only difference between the transimpersonal construction in (31a) and the intransitive construction in (31b) is that the former displays an overt 3rd person A marker while the latter does not. Note also that agreement prefixes for the 3rd person object and intransitive subject are identical; thus some parts of the agreement paradigm are organized on an ergative-absolutive basis (Evans 2004: 171). This may have facilitated the reanalysis of transimpersonal verbs as plain intransitives, and may explain the fact that reanalysis has not yielded a split-intransitive pattern even at final stages.

6 A similar cline has been described in detail by Donohue (to appear) for Skou, a (non-Austronesian) language of New Guinea.
Intransitives, objects, and 'trans impersonal' constructions

3.4.2 Covert reanalysis of experiencer object verbs in Germanic languages

More evidence for the diachronic instability of experiencer object constructions comes from the covert reanalysis of experiencer object verbs in Germanic languages. A celebrated case of reanalysis comes from English, where some subject experiencer verbs such as like started as object experiencer verbs; compare the following examples from Old English in (32a), Middle English in (32b), and Modern English in (32c); (Jespersen 1927, Lightfoot 1979, Faarlund 1990, Allen 1995):

(32) a. ðam cynge lícodon peran
the king DAT please, PL pears
b. the king likeden peares
c. the king liked pears

Here we can observe how the pressure for syntactic upgrading of a prominent (animate) experiencer object has resulted in a reversal of the mapping from semantic roles to syntactic functions. Similar cases of reanalysis of trans impersonal structures into personal constructions with subject experiencers abound in other Germanic languages (see Seefranz-Montag 1983; Bauer 2000 for more discussion of the 'personalization' of impersonal constructions in various branches of Indo-European). Compare the following examples from Swedish and German (from Seefranz-Montag 1983), where the impersonal pattern in (33a, 34a) is more archaic and the personal one in (33b, 34b) is an innovation:

(33) a. Det lyckades honom
it succeeded him
b. Han lyckades
he succeeded
‘He succeeded.’

(34) a. Mich hungert
me hungers
b. Ich hungere
I hunger
‘I am hungry.’

Here again a prominent object experiencer is promoted to subject status, although it does not lead to an inverse marking of arguments, as in the case of English. Still, at the later stages of reanalysis this does not yield a split-intransitive pattern, since, given the absence of object agreement, there is nothing to reveal the erstwhile object status of the experiencer argument.7

3.4.3 Split intransitivity with a person split: Yamphu and Totonac

In Himalayan languages, in particular Tibeto-Burman, experiential verbs frequently occur with object experiencers (or 'goal experiencers' in the terminology of Bickel 2004). Some of these constructions can appropriately be qualified as

7 As noted by Donohue, this volume, the impersonal patterns with 'non-canonical subjects' can be compared with patient subject constructions in split-intransitive languages; however, the subject status of experiential arguments in these constructions remains problematic. In fact, in only a few languages (like Icelandic) does the experiencer pass syntactic tests for subjecthood.
transimpersonal, as the verb takes the default 3sg agreement to agree with a non-referential A; cf. the following example from Limbu:

Limbu (van Driem 1987: 75):

(35)  
\[
\begin{align*}
Khe\,ya\,\text{m}\,\text{yusi} \\
\text{they & inebriate, 3SG.A > 3PL.P}
\end{align*}
\]
‘They are drunk (lit. it inebriates them).’

Notably, even though the experiencer is cross-referenced through the object agreement in constructions like (35), it may already display certain syntactic subject properties at this stage (see Bickel 2004 for arguments that experiencers in transimpersonal constructions in Tibeto-Burman languages are syntactic subjects or ‘pivots’). An interesting development of this pattern is found in Y amphu, another Kiranti language (Rutgers 1998: 109), where some experiencer verbs (called ‘semi-transitive’ by Rutgers) display a person-based split in their agreement paradigm. Constructions with 3rd person experiencers are transitive with the experiencer cross-referenced through object agreement, and the stimulus cross-referenced through the 3sg subject agreement:

Yamphu (Rutgers 1998: 109)

(36)  
\[
\begin{align*}
Wai\,?m\,-\text{æ}\,?\,\text{si}-s-w-e? \\
\text{thirst-erg attach-3SG.A > 3SG.P.FACT}
\end{align*}
\]
‘Is he thirsty?’

However, if the experiencer is a 1st or 2nd person, the verb takes an intransitive agreement to agree with the experiencer, as in (37):

Yamphu (Rutgers 1998: 109)

(37)  
\[
\begin{align*}
Sag\,-\text{æ}?\,\text{sis}-\text{ïp-ma} \\
\text{hunger-erg attach-exper-1PL.S}
\end{align*}
\]
‘We were hungry.’

This pattern is similar to split intransitivity, inasmuch as some forms in a paradigm are object-inflecting, while other pattern as plain intransitives (recall the situation in Tunica, where inchoative forms of stative verbs pattern as transimpersonals). Here, however, the split cuts across the person paradigm in a way reminiscent of split-ergative languages (of the Australian type) with an NP-based split. Apparently the emergence of this ‘split’ split intransitivity can be partially attributed to the fact that Yamphu displays features of split ergativity in its verbal agreement: Rutgers (1998: 9) notes that the intransitive agreement markers for the 1st and 2nd person shows similarities to those of a transitive verb with a 1st and 2nd person patient. This may have facilitated the reanalysis of the structure with a 1st and 2nd person patient-experiencer, as the verbal structure needed minimal readjustment in agreement morphology. Third person patient markers, however, are distinct from agent-subject markers, and therefore these forms are more resistant to reanalysis. However, functional factors have contributed to this pattern
Intransitives, objects, and ‘transimpersonal’ constructions

as well. Importantly, the split follows the predictions of the animacy hierarchy, since experiencers which are highest on the hierarchy (referring to speech act participants) are upgraded earliest to subject status, as far as verbal agreement is concerned, while the 3rd person experiencers, which are lower on the hierarchy, are still cross-referenced as objects. Another peculiarity of the experiencer constructions in Yamphu is that syntactically both patterns count as transitive, since in both (36) and (37) the stimulus retains its ergative marking. Thus, also in this case we witness the same functional pressures for upgrading the most prominent experiencers to subject position, but due to the structural features of languages involved they give rise to an unconventional split.

A similar case of partial reanalysis, observed for Yamphu, has also been reported for transimpersonal verbs in Totonac (Beck 2004: 31). Totonac is different, however, in that it treats the 2nd person as highest on the Person Hierarchy (that is, 2 > 1,3), while Yamphu ranks the 1st and 2nd person over the 3rd (1, 2 > 3). In Totonac, verbs taking an indefinite subject and a 3rd or a 1st person patient-experiencer cross-reference the latter by means of object agreement markers (through the 3pl object prefix ka- in (38)):

Totonac (Beck 2004: 32)

(38) ka-muSu:-má:-ka
3PL.P-kiss-PROG-INDF.A
‘They/some persons are kissing them.’

However, in the indefinite subject construction with the 2nd person experiencer, the argument is encoded through the intransitive subject agreement:

Totonac (Beck 2004: 31)

(39) ma:-ma?tāŋ l-ni?pá:-ka
caus-guard-caus-prog.2SG.S-INDF.A
‘They are making you guard it.’

In this example the subject status of the experiencer is confirmed by the fact that the progressive aspect marker takes a special form -ni?pá:- as it does elsewhere in constructions with a second person subject (Beck 2004: 32). Thus, as in case of Yamphu, the transimpersonal forms in Totonac have been reanalysed as intransitive in one part of the paradigm. Since structural factors cannot account for this partial reanalysis, we conclude that it is due to functional factors, i.e. the pressure to syntactically upgrade experiencers which are highest on the person hierarchy.

3.4.4 From indefinite subject constructions to (impersonal) passives

As a final illustration of the reanalysis of transimpersonal constructions, let us consider a well-known grammaticalization path leading from indefinite

\^ Independent evidence for special status of the 2nd person as the unmarked form in the person paradigm is that it shows most irregularity in combination with TAM markers (Beck 2004: 29).
subject constructions to impersonal passives and ultimately to personal passives (Greenberg 1959, Givón 1979, Shibatani 1985; see also Pustet and Rood, this volume). Greenberg was the first to propose this origin for the passive construction in the African language Masasi (cf. also Heine and Claudi 1986: 79–84); later on the same scenario was suggested for a number of other African languages (see Givón 1979 on Bantu languages), but also elsewhere (see Shibatani 1985 on Ainu, Trukic, and Indonesian). In Ainu, for example, the 1st person plural marker is also used as an indefinite subject marker with both intransitive and transitive verbs:

Ainu (Tamura 2000: 71; cf. Shibatani 1985)

(40) *Itak-an*

 speak-1PL

‘We speak’ or ‘One speaks.’

Transitive verbs with the impersonal marker, as in (41), are used as a kind of impersonal passive. The construction is clearly impersonal, as the patient-experiencer is still cross-referenced through the object agreement. On the other hand, reanalysis to passive has indeed taken place, as the agent can be introduced as an oblique object.

(41) *A-e-kóyki na*

 INDF-2SG.P-scold mod

 ‘You will be scolded/one will scold you.’

Constructions with an oblique agent can no longer be analysed as active with an indefinite subject (compare the anomalous literal translation of (42)):

Ainu (Tamura 2000: 72):

(42) *Unuhu oro wa an-kóyki*

 mother place from INDF-scold

 ‘He was scolded by (his) mother’ (lit.?? We/one scolds him by his mother).

The latter construction is indeed a passive, but an impersonal passive rather than a personal one. Indeed, as noted by Shibatani (1985: 824), the patient reveals few subject properties apart from positional ones.

A similar development is observed in Itelmen, not discussed in the typological literature in this connection. In this Paleosiberian language, constructions with a 3rd person plural subject marker can also be used as an (impersonal) passive (Georg and Volodin 1999). Compare a regular transitive clause with a 3PL subject in (43), and an impersonal passive construction in (44):

A related grammaticalization path not discussed here leads from the indefinite subject construction to the inverse construction. This path has been tentatively suggested by Rude (1997: 331, 349) for Sahaptin, where the verbal inverse marker *pá*- may be related to 3PL subject prefix. See Donohue, this volume, for a discussion of similar data from Tanglapui.
Intransitives, objects, and ‘transimpersonal’ constructions

Itelmen (Georg and Volodin 1999: 164)

(43) *Sillatumx-e*n komma n-an’çp-mitj*
   brother-PL 1SG 3PL.A-teach-1SG.P
   ‘The brothers taught me.’

(44) *Komma n-an’çp-mitj sillatumx-enk/ sillatumx-e*-n*
   1SG 3PL.A-teach-1SG.P brother-LOC/ brother-PL-LOC
   ‘I was taught by the brother/ brothers.’

Even though in both (43) and (44) the verb takes the same agreement morphology, the latter construction is impersonal, as the A argument appears in the oblique case and need not be plural to match the form of the agreement marker (Georg and Volodin 1999: 164). Georg and Volodin (p. 163) also note certain minor differences between the transitive ('active') and 'passive' paradigms (in particular in constructions with a 3rd person object), which means that morphological reanalysis is also under way. The authors do not specify which of the syntactic subject properties, apart from positional, are acquired by the patient of a 'passive' construction.

In some languages reanalysis has proceeded even further, and an impersonal passive construction was further reanalysed as a personal passive. According to Givón’s (1979) analysis, Kimbundu (a Bantu language) is such a language:

Kimbundu (Givón 1979: 211)

(45) *Nzua a-mu-muno* (kwa mame)
   John they-him-saw (by me)
   ‘John was seen by me.’

According to Givón, the original 3rd person plural subject marker has been reanalysed as a passive marker in (44), and the former object marker has become a new subject marker. Note that this construction bears certain similarities to patient subject constructions in split intransitive languages, inasmuch as some subjects are cross-referenced through object agreement. The construction, however, is different in that the agent can be introduced by an oblique phrase, and the erstwhile indefinite subject agreement is reanalysed as a passive marker. In that respect the evolution of the indefinite subject construction is clearly different from the development of transimpersonals. This can be attributed to functional differences between the two constructions. Even though in both constructions the original subject may be non-referential, only transimpersonals conceptualize it as an unknown natural force. In the indefinite subject construction the agent is human, and even if pragmatically defocused, it may be reintroduced in a later stage of reanalysis as an oblique constituent. There are also structural differences between the patient subject construction and (impersonal) passives. Thus, the latter need not contain an object agreement marker, and the subject marker remains overt even at later stages of reanalysis. This grammaticalization path again shows that overt encoding of the subject need not prevent reanalysis but would yield a
structure distinct from split intransitivity (see also Helmbrecht 2005 on Hocak and Pustet and Rood, this volume, on Lakota for other instances of languages where the indefinite subject construction is on the way to being reanalysed into the passive).

3.4.5 *The rise of split intransitivity: functional pressures and structural conditions*

In this section we looked at how universal functional pressures which promote reanalysis of transimpersonal constructions and experiencer object constructions interact with structural properties of the languages in question. It has been argued that although the functional pressures are universal they would produce a split-intransitive pattern only under certain conditions. The structural conditions include availability of object agreement on a transitive verb, and covert marking of a 3rd person (non-prominent) subject. If these conditions are not met, the reanalysis may be suspended or may lead to an outcome structure distinct from canonical cases of split intransitivity. Thus, if there is no object agreement on the verb, covert reanalysis (of the *please > like* type observed for English) is more likely, because reanalysis would leave no traces in the verb structure which reveal the erstwhile object status of the experiencer-patient. On the other hand, if the subject is overtly marked in a transitive construction, reanalysis may be suspended at a transimpersonal stage. At this stage the construction will still look morphologically transitive, and therefore should be considered as an extended use of the transimpersonal rather than as split intransitive (cf. Iwaidjan, but also some Papuan languages). The availability of overt subject agreement will not necessary prevent reanalysis; but a language would have to take some extra measures, either suppressing the dummy A marker (as in Galela) or reinterpreting it as a passive marker (as in Ainu, Itelmen, and Kimbundu).

The rise of split intransitivity on the basis of transimpersonal experiencer constructions is also constrained by the alignment pattern of the language in question, as already anticipated by Klimov (1977). Consider a speculative case of an ergative language where both case and agreement operate on an ergative-absolutive basis (of the type found, for example, in Dagestani languages, which show (gender) agreement with the absolutive argument; cf. e.g. Comrie 2004 on Tsez). In such a language, if the transimpersonal construction yields to functional pressures and is reanalysed, the changes in the structure will be minimal: the experiencer argument will still be marked by the absolutive case and cross-referenced through absolutive agreement. So in the absence of a separate transitive subject agreement marker, one would need hardly any readjustment in either morphology or syntax when a construction is reanalysed as intransitive. This supposition would predict that a consistently ergative language cannot easily develop split intransitivity on the basis of a transimpersonal construction (which, however, does not exclude the possibility of reanalysis on the basis of different constructions; see below). Thus, ergative alignment for agreement (and case) promotes a covert reanalysis to an
Intransitives, objects, and 'transimpersonal' constructions

intransitive pattern, rather than the rise of split intransitivity. We have observed this for Yampu, where ergative patterning in a part of the agreement paradigm has led to the emergence of split intransitivity complicated by a person 'split'.

3.5 Concluding remarks

We have seen that there is a universal functional pressure towards the reanalysis of transimpersonal experiencer constructions, the reasons for which are twofold: there is pressure on one hand to demote a non-prominent indefinite subject of a transimpersonal construction and, on the other, to syntactically upgrade a prominent experiencer-object. These universal functional pressures, however, will yield a split-intransitive structure only under particular structural conditions. The conditions which lead to a partial convergence of transimpersonal and intransitive structures and open the way for reanalysis include, in particular, the availability of object agreement on transitive verbs, and a covert marking of 3rd person (non-prominent) subjects. If these conditions are not met, the reanalysis may be suspended or may lead to an outcome structure distinct from canonical cases of split intransitivity. Note also that this account straightforwardly explains why both structural features persist in many split-intransitive languages, in particular, why they are overwhelmingly head-marking.

To conclude, unlike the approaches which explain the phenomenon of split intransitivity through role domination (direct mapping from semantic functions to case marking), I have argued that this alignment pattern may be a secondary phenomenon which arises through a conspiracy of universal functional tendencies and language-particular structural properties.\(^{10}\) Recall, however, that this scenario is proposed in the first place for split-intransitive languages with an accusative base, where the patient subject construction is a minor pattern. For split-intransitive languages with an ergative base (with agentive subject construction as a minor pattern) other explanations may be needed; yet this latter pattern could also involve a reanalysis of a transitive construction. Thus it has been claimed that the minor agentive subject pattern in Georgian and Basque derives historically from a transitive construction with an indefinite or cognate object (see Harris 1985, Hewitt 1995, and Wier 2005 on Georgian, and Trask 2002 and Aldai, this volume, on Basque). Still other scenarios for the rise of a split-intransitive pattern have been proposed, in particular scenarios involving a shift between an ergative and an accusative alignment type (Plank 1985, Dixon 1994:

\(^{10}\) Note that opting for a diachronic explanation for the phenomenon of split intransitivity does not entail that the existence of a semantic motivation behind the split has to be questioned. Indeed, even in those cases where split intransitivity is due to the reanalysis of a transitive construction, as long as there is a choice between the two alternative patterns, the contrast is likely to be semanticized and spread by analogy to other verbs. This means that a semantic and a diachronic explanation of this phenomenon are not necessarily incompatible.
I hope, however, to have demonstrated that transimpersonal experiencer constructions constitute an important source for the rise of split intransitivity, as already anticipated in Sapir’s (1917) account of object-inflecting verbs in Native American languages.

Thus, split intransitivity may develop from reanalysis of an ergative structure into accusative, leading to alternative encoding of subjects at intermediate stages. For example, in Wakhi (an Iranian language), the gradual extension of ergative marking to intransitive subjects has yielded a pattern reminiscent both structurally and functionally of split-intransitive system of the ‘fluid’ type (Bashir 1986; cf. Dixon 1994: 34).
Thematic roles, event structure, and argument encoding in semantically aligned languages

PETER M. ARKADIEV

4

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 Guarani 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 Shuinsky, 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' than all other proto-properties regardless of their number. That is, if a verb entails for its argument a specific property \( x \) from, say, the set of agentic 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. 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) \( \text{gà té} \)  
\( \text{1pl.excl.sbjv 3pl.objv see dur} \)  
'\text{We see them.}'

(5) \( \text{gà ká zu} \)  
\( \text{1sg.sbjv see+3sg.objv dur} \)  
'\text{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) \( \text{gà li zu} \)  
\( \text{1pl.excl.sbjv go dur} \)  
'\text{We are going.}'

(7) \( \text{tóa lo zo} \)  
\( \text{3sg.sbjv fall dur} \)  
'\text{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

<table>
<thead>
<tr>
<th>Person</th>
<th>‘Subjective’</th>
<th></th>
<th>‘Objective’</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Singular</td>
<td>Plural</td>
<td>Singular</td>
<td>Plural</td>
</tr>
<tr>
<td>1st</td>
<td>gà</td>
<td>gá (excl.)</td>
<td>low tone*</td>
<td>gé (excl.)</td>
</tr>
<tr>
<td></td>
<td>dà (incl.)</td>
<td></td>
<td>dè (incl.)</td>
<td></td>
</tr>
<tr>
<td>2nd</td>
<td>já</td>
<td>wà</td>
<td>è</td>
<td>ü</td>
</tr>
<tr>
<td>3rd</td>
<td>tóa</td>
<td>tá</td>
<td>high tone*</td>
<td>té</td>
</tr>
</tbody>
</table>

* 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.

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 than ‘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)).
Thematic roles, event structure, and argument encoding

Table 4.2. Verb classes in Georgian

<table>
<thead>
<tr>
<th>Class</th>
<th>Case marking in Series II</th>
<th>Agreement (3sg Present, 3pl Present, 3pl Aorist)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>A: Ergative—P: Nominative</td>
<td>-s/-en/-es</td>
</tr>
<tr>
<td>II</td>
<td>S: Nominative</td>
<td>-a/-ian/-nen</td>
</tr>
<tr>
<td>III</td>
<td>S: Ergative</td>
<td>-s/-en/-es</td>
</tr>
</tbody>
</table>

(11) té gé ya ne (12) gé li ni
3pl.objv 1pl.excl.objv see pfv 1pl.excl.objv go pfv
'They saw us.' '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

<table>
<thead>
<tr>
<th>Class I ‘paint’</th>
<th>Class II ‘die’</th>
<th>Class III ‘work’</th>
</tr>
</thead>
<tbody>
<tr>
<td>3sg Present</td>
<td>xat’av-s</td>
<td>k’vdeb-a</td>
</tr>
<tr>
<td>3pl Present</td>
<td>xat’av-en</td>
<td>k’vdeb-ian</td>
</tr>
<tr>
<td>3pl Aorist</td>
<td>daxat’-es</td>
<td>mok’val-nen</td>
</tr>
</tbody>
</table>

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*  *mō-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-myeyr-*d-a  
    man-NOM INCH-sing-INCH-AOR.3SG.SBJV  
    ‘The man began to sing.’

(19) *c’q’al-ma*  *i-duy-*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.
Thematic roles, event structure, and argument encoding

('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)):

\[
(20) \quad \theta : \{P: \text{change of state}\} \rightarrow \langle \text{Pred: Class II}, \langle S: \text{Nominative}\rangle \rangle \\
\theta : \text{elsewhere} \rightarrow \langle \text{Pred: Class III}, \langle S: \text{Ergative}\rangle \rangle 
\]

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.
Bats 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 'Isova-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 b. so wože
1SG.ERG fell 1SG.NOM FELL
'I fell (on purpose).' '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’-yordar 'freeze', daq'dalar 'dry up', daxdalar 'go mindlessly, unconsciously', mildar 'be cold', qeč'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', luvor 'talk', darc'dalar 'take off one's clothes', daxar
'live', lap'i:ar 'play';
3. those which allow both Nominative and Ergative Ss, e.g. dah"daxar 'get
drunk', derc'ar 'turn into', xarc'dalar 'change', 'opdalar 'hide, come out of
sight', k'urˇcdalar '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 uncon-
sciously'. 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 'opdalar
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
gavr'€dalar '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):

\[ (23) \quad \frac{\varnothing \{ A: \text{volitinality} \}}{\varnothing \{ \text{elsewhere} \}} \rightarrow \frac{S: \text{Ergative}}{S: \text{Nominative}} \]

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).

\[ (24) \quad uu\; uu\; uu\; uu\; \text{zurc} = \text{amt}-\text{za-vu} \]
\[ 1SG\; 2SG\; beat-1SG.AGT-2SG.PAT \]

'I have beaten you.'
Thematic roles, event structure, and argument encoding

Table 4.4. Agentive and Patientive agreement in Tabassaran

<table>
<thead>
<tr>
<th>Person</th>
<th>Agentive</th>
<th>Patientive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Singular</td>
<td>Plural</td>
</tr>
<tr>
<td>1st</td>
<td>-za</td>
<td>-a</td>
</tr>
<tr>
<td>2nd</td>
<td>-va</td>
<td>-ću,ā</td>
</tr>
</tbody>
</table>

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’, ružun-za ‘I began to cry’, rușun-za ‘I came’; (ii) verbs which take only Patientive agreement, e.g. kabqun-za ‘I drowned’, warun-za ‘I froze’, ergra-za ‘I got tired’; and (iii) verbs which may take both Agentive and Patientive agreement, e.g. ružun-za ‘I remained (voluntarily)’ vs. rușun-za ‘I remained (against my will)’; aqun-za ‘I fell (intentionally)’ vs. aqun-za ‘I fell (by accident)’; hilirqun-za ‘I shook (on purpose)’ vs. hilirqun-za ‘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). 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):

   1SG.NOM 3SG-ACC chase.away
   ‘I chased him away.’

   b. muɾ-ʔo: ?ɛː¿ədiw
   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. ?a: wâq'île
   1sg.nom would.go
   ‘I would go.’

b. ?a: sbî'c'
   1sg.nom get.up
   ‘I got up.’

c. ?a: ñadên
   1sg.nom swim
   ‘I swam.’

d. ?a: cêc'
   1sg.nom escape
   ‘I escaped.’

(27) a. ?a: yâ:qa'c'in
   1sg.nom careful
   ‘I’m careful.’

b. ?a: ?nâc'
   1sg.nom hiding
   ‘I’m hiding.’

c. ?a: ?e q'âm q'dî.
   1sg.nom cop kindhearted
   ‘I’m kindhearted.’

d. ?a: ?e kúc'
   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 Georgian5).

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. ?o: lô:ya
   1sg.acc fall
   ‘I fell.’

b. ?o: madâts'îw
   1sg.acc slip
   ‘I slipped.’

c. ?o: ?k'êm.nâda.
   1sg.acc get.well
   ‘I’m getting well.’

d. ?o: qâmâ:leč
   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. *to: *kasila$_{1sg}$$_{acc}$ cold
   'I'm cold.'

b. *to: *kìts’ˇ ciw$_{1sg}$$_{acc}$ scared
   'I'm scared.'

c. *to: *?$t^d$á$^\ell$$_{1sg}$$_{acc}$ painful
   'I'm in pain.'

d. *to: *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. *?a: *sma mtí:ˇ c’$_{1nom}$ go.to.bed
   'I went to bed.'

b. *?a: *sma mtí:ˇ c-ka$_{1acc}$ go.to.bed-INF
   'I must have fallen asleep.'

(31) a. *?a: *k’lú:k’luw$_{1nom}$ cough
   'I coughed (intentionally).'

b. *?a: *k’lú:k’luw$_{1acc}$ 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. <$\emptyset$: [A: volitional]$>$ <$S$: Nominative$>$

b. <$\emptyset$: [P: affected; A: sentient]$>$ <$S$: Accusative$>$

c. <$\emptyset$: 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 *mat’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 patiientive 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 patiientive set of entailments, and not even a group of patientive entailments, but rather a combination of a patiientive 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 than 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 ‘sentience’ 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 here7).

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 20008).

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:

\[ \langle \theta | A: +\text{volitional} \rangle, \langle P: –\text{change of state} \rangle \rightarrow \text{Agentive marking}; \]
\[ \langle \theta | A: –\text{volitional} \rangle, \langle P: +\text{change of state} \rangle \rightarrow \text{Patientive marking}. \]

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:
\[
\langle \theta \mid [A: +volitional], [P: +change of state] \rangle \rightarrow \text{Agentive marking if} \\
[\pm \text{volitionality}] >> [\pm \text{change of state}]; \\
\langle \theta \mid [A: +volitional], [P: +change of state] \rangle \rightarrow \text{Patientive marking if} \\
[\pm \text{change of state}] >> [\pm \text{volitionality}].
\]

In Georgian, where telicity ([\pm \text{change of state}]) is more prominent than agentivity ([\pm \text{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 than 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...

(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).
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 reclassified 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.
Part II
Eurasia
Why are stative-active languages rare in Eurasia?
A typological perspective on split-subject marking

JOHANNA NICHOLS

5.1 Introduction

What have classically been known as stative-active languages were first described in the Americas and are still best attested there (for surveys see Comrie 2005, Siewierska 2005, Dixon 1994: 70–110, Mithun 1991, Merlan 1985), although the number of attestations in Oceania is rapidly gaining as fieldwork and its consolidation proceed (Donohue 2004b, Foley 2005). The type is rare in Eurasia, however: the only good examples are Georgian and its sister Kartvelian languages (Harris 1991), where the split-S type may be millennia old; Batsbi (or Tsova-Tush; Holisky 1987), where the fluid-S type is recent, not found in sister languages, and possibly due to the throes of language death; Basque, split-S (Aldai, this volume; Comrie 2005: 399, citing Hualde and Ortiz de Urbino 2003: 364); and some Tibeto-Burman languages in which fluid-S marking is bound up with evidentiality and a conjunct/disjunct system (DeLancey 1985, Hargreaves 2005). Ket (Yeniseian), identified as stative-active (Comrie 1982, Siewierska 2005), has recently been shown by Vajda (2003, 2004) to have no particular alignment, every verb dictating its own valence pattern. Figure 5.1 shows the languages coded as stative-active in the large Autotyp sample (Bickel and Nichols 2002–).

Part of the reason for its rarity in Eurasia is probably the rarity there of the head-marking type, which strongly favours stative-active alignment in the Americas and the Pacific (Nichols 1992: 100–105). However, it will be argued here that a more principled typological definition of alignment patterns and the criteria for setting them up makes it possible to detect a number of exemplars in Eurasia. Let us begin with terminology. Here I will deal only with split types and not fluid types. Terms used for the split type include 'statative-active', 'split-S', 'split-intransitive', 'active',...
'active-inactive', and others, depending partly on individual preference and partly on the linguist's understanding of the putative semantic basis for the split, with 'split-S' being noncommittal as to semantics. Now, in descriptions of split-S languages the relevant verbs often include some that plausibly or demonstrably have objects, such as 'like' or 'fear, be afraid', so we need a term that is noncommittal as to syntax as well. I will use 'split subject (marking, language)' as a provisional generic term that covers the coding of S and A (with no implication that linguists whose descriptions I have used apply the term 'subject' in this way). I also assume a maximally syntactic and minimally prototype-based definition of the notions A, O, and S, so that A comprises experiencer, possessor, etc. subjects as well as agent-like ones and so that A and O apply not only to the arguments of transitive verbs but also to oblique subjects and/or objects.

The approach taken here is that of lexical typology (Nichols et al. 2004) as opposed to the more usual morphosyntactic typology which underlies most work on alignment. In morphosyntactic typology, the basic building blocks of clauses are predicates and arguments of relatively discrete and determinate types; there are roles and alignments, again of fairly discrete and determinate types, and a language has an alignment type that determines the coding of its roles. The verbs fit into that alignment type, except for a few that are deponent, assign semantic case, etc. That is, the alignment type is a property of the language that determines various aspects of its morphosyntactic behaviour.

In lexical typology, the essential building blocks of clauses are individual verbs, each of which has its own argument structure and valence. Alignment types, role types, and argument types are not building blocks but linguists' generalizations
over the behaviour of individual verbs. For instance, an ergative language is one in which at least some criterial percentage of the verbs code S and O alike, and the threshold percentage is determined by surveying a set list of verb glosses across a sample of languages. In this approach, types are not discrete but form a continuum. Every language can be expected to display multiple splits in its argument coding, and a language can be said to belong to an alignment type if it displays an above-average (or, better, significantly above average) frequency of verbs with that kind of argument coding. For most languages, a typical or usual type of argument coding is easy to discern. The task for lexical typology, however, is to determine the kinds and frequencies of all coding patterns and not just the typical ones. Thus, depending on the verb, S can be coded like A or like O, and also like a possessor or an indirect object; A can be coded not only typically but also like S or O or possessor or indirect object; and so on.

The notion of object needs to go beyond Dixon's O, which I define here as object of monotransitive verb. For ditransitive verbs we need two additional object types, and I will label these T (most theme-like object of ditransitive) and G (most recipient-like or goal-like or beneficiary-like object of ditransitive). (For this typology of syntactic roles, and more examples, see Bickel and Nichols, to appear.) In direct/indirect object languages (Dryer 1986), the alignment of objects is T=O; in primary/secondary object languages, object alignment is G=O. Across the world's languages, some verbs or clauses code T like O, others code G like O, others may code one or both objects like possessors, like goals, like locations, like adverbials, etc. In (1)–(4), money is T and children is G regardless of coding; the morphological coding of money is like O in (1)–(3) and oblique in (4), and that of children is like O in (2) and (4) and like a goal in (1) and (3). The goal-like coding of G, i.e. indirect object, can be symbolized as G = goal.

(1) I gave money to the children. [G = goal, T = O]
(2) I gave the children money. [G = O, T = O]
(3) She left her money to her children. [G = goal, T = O]
(4) They provided their children with enough money to start a business. [G = O, T = oblique]

This chapter is a first lexical-typological investigation of alignment and argument coding patterns cross-linguistically. It is a pilot study, designed to set up a manageably sized list of verb glosses, drawn chiefly from those figuring in descriptions of split-subject languages, and to see whether the cross-linguistic treatment of these glosses reveals relatively discrete types such as ergative, accusative, and split-S. It also aims to show how Eurasian languages do and do not resemble American, Oceanian, and African ones.

There are two kinds of survey that can reveal lexical types: text frequency surveys and lexical surveys. A text frequency survey probably gives the most sensitive
indication of a language's overall preferences, but is labour-intensive and requires close control of genres, stylistic levels, etc. in text corpora. I estimate (from pilot studies for Nichols 1993 and unpublished work on medieval Slavic languages) that a corpus of about 1,000 clauses, exclusive of those containing the verb 'be' that swamps all frequencies in some Indo-European languages, is enough to yield reliable information on natural frequencies of verbs: which verbs are the most frequent, which is the basic one corresponding to a survey gloss, what is its usual argument coding, etc. From 1,500 to 2,000 clauses need to be surveyed in order to extract 1,000 exclusive of 'be'. Finding such corpora is difficult and surveying them very time-consuming. I know of almost no corpora (other than my own pilot studies and the data for Bickel 2004) that are annotated in such a way as to make argument status and argument coding types clear. Thus cross-linguistic corpus work for the number of languages needed in typological surveys like this one is simply impracticable.

Lexical surveys are much less time-consuming, though still labour-intensive, and yield reliable if more schematic information. The procedure for lexical surveys is described in Nichols et al. (2004): the linguist sets up a list of (in this case) verb glosses designed to test for (in this case) S and A coding in a range of semantic circumstances, tests it out on a few languages to verify that the gloss can be found in most dictionaries, draws up a list of proxy verbs that can be used in a pinch and an explicit set of coding guidelines, and looks up that set of glosses in bilingual dictionaries, or elicits them from a native speaker, to determine their argument coding. This requires dictionaries large enough to contain all or most of the survey verbs and complete enough to include valence information. Few dictionaries include valence information explicitly, but a number of dictionaries for well-described languages endeavour to give an example or two showing the 'usage' of each verb (i.e. valence and argument treatment).

The results of preliminary text surveys and lexical surveys of dative subject coding in six languages of Europe and the Caucasus are shown in (5). The text surveys are pilot studies in progress and too short to be reliable in detail. Both surveys reveal a large difference between the Indo-European languages, which have low frequencies of dative subjects, and the other two (indigenous Caucasian families), which have higher frequencies. The text survey shows that dative subjects are absolutely infrequent in all six languages, justifying the classification of all six as basically ergative or accusative or split ergative/accusative but not dative (if that is considered to be an alignment type). The lexical survey is that described in section 5.5.2 below. For one Ossetic verb the valence information could not be recovered from examples, so only 19 were surveyed, and one verb more or less makes an appreciable difference in the calculated percentage. A

1 A frustrating practice in European lexicography is citing all such illustrative examples as infinitives, thereby obscuring information about subject case.
larger survey, covering the same semantic fields with more duplication, would reduce the impact of gaps and flukes but increase the time. The text surveys required, for each clause, approximately 15 minutes (Russian) to an hour for inputting, analysis, and annotation. (Corpora for the first four languages are on BITC (Sprouse 1997) and valence-related annotation for all six is in Excel®.) Thus a minimal adequately sized corpus can take hundreds of hours per language to assemble. Lexical surveys took from an hour or two to several hours per language.

(5) Text frequencies of dative subjects as a percentage of all clauses in six languages. In all of them, dative = the case of indirect objects, i.e. coding of G as goal in languages with predominant coding of G as goal (direct/indirect object alignment). Datives = % of clauses with dative S or A. Verbs = number and % of verbs out of the 20-verb survey list shown in Appendix 5.1 that take dative subjects. Chechen was not surveyed lexically but would be approximately identical to Ingush. EC = East Caucasian, IE = Indo-European.

<table>
<thead>
<tr>
<th>Family</th>
<th>Text survey</th>
<th>Lexical survey</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Datives</td>
<td>Clauses</td>
</tr>
<tr>
<td>Ingush EC (Nakh)</td>
<td>4%</td>
<td>406</td>
</tr>
<tr>
<td>Chechen EC (Nakh)</td>
<td>8%</td>
<td>270</td>
</tr>
<tr>
<td>Georgian Kartvelian</td>
<td>6%</td>
<td>104</td>
</tr>
<tr>
<td>Ossetic IE (Iranian)</td>
<td>1%</td>
<td>89</td>
</tr>
<tr>
<td>Russian IE (Slavic)</td>
<td>0%</td>
<td>177</td>
</tr>
<tr>
<td>Bulgarian IE (Slavic)</td>
<td>1%</td>
<td>80</td>
</tr>
</tbody>
</table>

A last detail. It is customary to abbreviate (morphosyntactic, whole-language) alignment types as S=O, S=A, etc., and I have extended this to T=O, G=O. It also seems expedient to have a notation for indicating that a particular verb codes its S like a typical O in that language, regardless of the overall alignment type or preference of the language. I will provisionally use the dot and a lower-case letter for this: S.o, A.g, etc. Thus in Latin, which has overall S=A alignment, the coding of S in (6) is S.o:

(6) me piget
1SG.ACC shame 3SG.PRES
'I'm ashamed,' lit. '(It) shames me.'

5.2 Survey

This study began with the observation that several of the verb glosses most prone to oblique experiencer marking in the survey of Bossong (1998) were also
among those likely to surface as ‘statives’, object-coded, or specially coded in Merlan (1985), Mithun (1991), and other literature on split-S marking—despite the fact that work on split-S marking generally saw the semantics of the split as involving agency and non-agency. (Merlan was more or less alone in emphasizing that the split involved markedness relations or open vs. closed sets, not a positive semantic descriptor such as agency.) As I subsequently discovered, Walsh (1987) had much earlier observed the essential structural commonality between object-coded experiencers and classic split intransitivity. I assembled a list of verbs drawn from Bossong (1998) and Merlan (1985) with some additions of my own and some adjustments after the first few surveys. The list of verbs, together with proxies, comments, and elicitation instructions, is given in Appendix 5.1. (This is the worksheet which I distributed to linguists who served as expert consultants or fieldworkers on some of the languages. See note 7.)

What is surveyed is not verbs but more precisely verb glosses. Appendix 5.1 spells out the exact sense intended for potentially ambiguous glosses. For each language these 20 glosses were sought in English–X (or Russian–X or Spanish–X) dictionaries and then checked in X–English (X–Russian, X–Spanish) dictionaries. (Before doing the lexical survey, a grammar was read or skimmed to make morphosyntactic analysis of lexical entries possible.) Ten languages were surveyed by native-speaking linguists or expert linguists, and some of these surveys were done by elicitation rather than dictionary lookup. One language (Caddo) was surveyed not with dictionaries but by going through all examples in grammars in search of the 20 glosses, and for several languages grammar searches supplemented the lexical searches.

Semantically, the glosses have a mix of agent, patient, and experiencer subjects. The survey excluded glosses likely to surface as adjectives in languages having such a part of speech: durable or permanent properties and qualities such as ‘young’, ‘old’, ‘tall’, etc. In split-subject languages having few or no adjectives, the subject coding of these verbs is of interest, but in languages with adjectives the subject coding is assigned by the copula or otherwise differs from what determines subject treatment with verbs. Though these predicates have some of the most patient-like subjects, subject coding with them is due to factors orthogonal to the split-subject coding at issue here.

Forty-one languages were surveyed (Appendix 5.2). They constitute a sparse 21-language, genealogically fairly well-distributed sample of Eurasia; all known American split-subject languages for which I could obtain lexical materials, and a few others added to expand and diversify the sample with languages I considered likely to have split-subject coding; and a similar mix of diverse known and possibly split-subject languages from Oceania (western insular Southeast Asia, Melanesia (including New Guinea), and northern Australia). Thus, there are 21 languages from Eurasia and 20 from parts of the world known to be rich in classic stative-active languages. Since the sample design targeted known split-subject
languages, they are (presumably) more common in the sample than in the world at large.

In addition to the 20 verb glosses, I surveyed typical subject alignment and typical object alignment for each language, using the Autotyp database (Bickel and Nichols 2002–).

5.3 Results

5.3.1 Types of subject coding

Eight types of subject coding occurred frequently in this survey. They are shown in (7) and illustrated below.

(7) Types of subject coding. For S, A, O, G see section 1. x = internal (defined below). Poss = (adnominal) possessor.

<table>
<thead>
<tr>
<th>Role</th>
<th>Coded as</th>
<th>Notation</th>
<th>Terms used in literature; common case names (in case-using languages)</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>A</td>
<td>S.a</td>
<td>Active, A-coded; ergative/nominative</td>
</tr>
<tr>
<td>S</td>
<td>O</td>
<td>S.o</td>
<td>Stative, O-coded; absolutive/accusative</td>
</tr>
<tr>
<td>S</td>
<td>A + Ox</td>
<td>S.a + Ox</td>
<td>Internal (dummy) object; ergative/nom.</td>
</tr>
<tr>
<td>S</td>
<td>O + Ax</td>
<td>Ax + S.o</td>
<td>Transimpersonal</td>
</tr>
<tr>
<td>S, A</td>
<td>G</td>
<td>S.g, A.g</td>
<td>Dative subject; dative</td>
</tr>
<tr>
<td>S</td>
<td>Poss</td>
<td>S.poss + Sx</td>
<td>Psycho-collocations; genitive</td>
</tr>
<tr>
<td>S</td>
<td>A or O</td>
<td>S + Sx</td>
<td>Psycho-collocations; absolutive/nominative</td>
</tr>
</tbody>
</table>

Types S.a and S.o can be illustrated by the following from Batsbi (or Tsova-Tush; East Caucasian: Nakh; Georgia)

(8) S.a:  as *vuizhn-as*
          1SG.ERG M.fall-AOR-1SG.ERG
          ’I fell (on purpose)’ (v = a gender class)

(9) S.o:  so *vozhen-so*
          1SG.ABS M.fall-AOR-1SG.ABS
          ’I fell (accidentally)’

Regardless of the coding, there is only one argument in each clause, the S. Batsbi is a fluid-S language, so S.a and S.o coding are both equally direct as subject coding and neither is oblique (to use traditional but still valid terms). In an accusative language, S.a (as in the English translations of (8) and (9)) has direct subject coding and S.o (as in (5) above) oblique. In an ergative language it is the other way around. Whether S.a or S.o is oblique or direct is not discussed here further (this is always indicated in the survey data, however).
S.a + Ox coding is common in light verb constructions, which are common in southwest Asia. These have what I call an internal object, illustrated in (10) and (11):

(10) Persian  *man nafas mi-kesh-am*

\[ \text{1sg breath cont-pull:pres-1sg} \]

‘I breathe’

(11) Ingush (East Caucasian: Nakh; southern Russia)

\[ \text{aaz sa daux} \]

\[ \text{1sg.erg breath d.take.pluralional} \]

‘I breathe’

Superficially, each of (10) and (11) contains two arguments, but one of them, ‘breath’, is lexically internal (annotated Ox).


(12) Tunica (isolate; southeastern US; Haas 1941: 59)

\[ \text{ih-e ha-kati} \]

\[ \text{1sg.o-breathe-3sg.f.a} \]

‘I am breathing’ (lit. ‘(it) breathes me’)

The clause is impersonal, with default or dummy 3rd singular feminine A morphology and the experiencer treated as object. A non-impersonal variant of this type is in (13), where the verb is not impersonal but agrees with an overt internal agent:

(13) Wishram (Chinookan; northwestern US; Dyk 1933: 148)

\[ \text{aGa kwapt ikmakan ga-ch-(a)-u-Xa} \]

now then \[ \text{m-anger rempst-3sgm.a-(3sgf.o)-deix-make} \]

‘Thereupon she became angry’ (lit. ‘anger did her’)

A.g coding is illustrated by dative subjects in Ingush and Russian:

(14) Ingush  *suona yz gu*

\[ \text{1sg.dat 3sg.abs see} \]

‘I see him/her/it’

(15) Russian  *mne nravitsja eta kartina*

\[ \text{1sg.dat like-3sg-refl this picture:nom} \]

‘I like this picture’

In both examples the subject (A) is 1st person singular. The object is a normal absolutive direct object in Ingush but an oblique nominative object in Russian.

S.Poss coding is illustrated by the light verb constructions that Matisoff (1986) calls ‘psycho-collocations’ for Southeast Asian languages. Typically these have an internal or incorporated body-part noun and the experiencer subject is coded as its possessor:
Typological perspective on split-subject marking

(16) Limbu (Tibeto-Burman, Nepal; Michailovsky 2002)

*a-ningwa*  mutch-ε

1SG.GEN-mind forgot-3SG

'I forgot', lit. 'my mind forgot'

A variant of this, found e.g. in Skou of the Macro-Skou family of New Guinea, treats the subject like an ordinary S or A but otherwise has the same structure involving internal body-part noun and light verb.

5.3.2 Frequencies of coding types

Appendix 5.3 shows the 20 verb glosses and the total number of attestations of each subject coding type for each gloss. (Half points occur where one verb has two different subject treatment possibilities, and where two different verbs each had equal claim to be regarded as the best candidate for some gloss in some language.) Several patterns emerge:

5.3.2.1 A coding A coding (S.a and A.a) is the most frequent overall, making up about half (49 per cent) of the total entries.

5.3.2.2 O coding O coding is more frequent for the verbs of perception and emotion (which have experiencer subjects) than for the verbs of action and sound (including those with patient or theme subjects). G and Possessor coding are at their most frequent for the verbs of perception and emotion.

5.3.2.3 G and Possessor coding G and Possessor coding are found only with experiencer subjects.

Note that, in general, it is semantic experiencers that are most prone to have non-A coding, be it O, G, or Possessor. This is true in classic stative-active languages and in dative-subject languages of Eurasia.

5.3.3 Is there a stative-active type?

The sample allows the various kinds of subject coding to be plotted against each other for a graphic view of whether and how subject-coding preferences cluster. Figure 5.2 plots percentages of verbs with S.a vs. S.o coding in this fashion, for 40 languages. (One language, Mawng, is not included since it has three-way alignment and does not fit in easily.) Ergative languages cluster in the upper left, accusative languages in the lower right, and stative-active languages in the middle—along with a number of languages not generally labelled stative-active. The plot shows that, if S.o vs. S.a is the criterion, then at least Tauya, Motuna, Chol/Tzotzil, and Basque should also be classified as stative-active, and perhaps

2 Gutiérrez and Zavala (2005) do in fact classify Chol as stative/active; to my knowledge this has not been pointed out for Tzotzil, however. Basque has been called both ergative and stative/active, as noted above.
Figure 5.2. Plot of S.a coding (horizontal axis) vs. S.o coding (vertical axis). *Italics, bold* = languages generally regarded as stative-active. Touching plot symbols = actually identical scores. The languages at the bottom, with no S.o coding, are (from left to right) Tukang Besi, Skou; Latvian, Nanai; Ossetic; Russian, German, Erzja Mordvin; Kayardild; Hungarian, Tuva, Nunggubuyu, French; Korean.

Caddo should be reclassified as ergative. Then the plot will give a relatively satisfactory view of split vs. non-split S, and the three clusters are somewhat discrete. This picture could change with a larger sample of languages, however.

Languages displaced to the lower left in Figure 5.2 either have a good number of G-coded or Possessor-coded subjects (e.g. Hindi, Lezgi) or lack datapoints (e.g. Haida, Caddo) or both.

Split-S languages are more dispersed than either ergative or accusative languages, and range from the near-ergative Caddo to the near-accusative Lakhota and Seneca. Within the split-S group there is no clustering at all into more ergative vs. more accusative types, but simply a continuum.

Figure 5.3 gives a similar plot of both S and A marking, and plots A coding against all non-A coding (O, G, Possessor). Here clustering is less evident and classic stative-active languages more dispersed among others. For example, Ossetic and Lakhota plot in the same place (4.5 A-coded subjects, 4.5 non-A-coded), though the non-A subjects of Ossetic are dative (i.e. G coding) while those of Lakhota are S.o agreement. This plot shows a notable spread within the Indo-European and Uralic language families (due almost entirely to differences of A vs. G coding), while in Figure 5.2 they clustered tightly together.
Again the split-subject languages are more dispersed than either ergative or accusative, with Caddo and Lakhota again at the extremes. Again there is no visible clustering within the split-subject languages but a continuum from more ergative to more accusative.

On either of these plots (and on others not shown here, e.g. S.a vs. S.g), and even if all the languages plotting between Caddo and Lakhota are classified as split-subject, the split-subject languages do not make up a very clear type. They are in the middle of the plot but occupy a large range, and they extend from near-ergative to near-accusative without a clear division between these two. The dispersal would presumably be lessened if property and durable state verbs such as ‘be young’, ‘be tall’, were included in the survey. These are lexicalized as adjectives in most of the Eurasian languages, and when used predicatively they occur with a copula, which in all ergative languages known to me takes an S.o subject. In languages with few or no adjectives, these predicates are lexicalized as verbs, and they generally take S.o marking. Such predicates are well represented among the S.o-coded verbs of Lakhota, for instance. Inclusion of a few such verbs would push Lakhota in Figure 5.3, and Lakhota and Seneca in Figure 5.2, closer to the centre of the plot and would cause at least the classic stative-active languages, and perhaps the entire split-subject set, to form a more coherent cluster. As explained in section 5.5.2, these predicates were left out because they would give disproportionate representation to the verb ‘be’ in the Eurasian languages, and I see no way to do...
Johanna Nichols

a worldwide lexically balanced verb survey including enough of them to make a
difference in the position of Lakhota and languages like it.

Thus, from a lexical-typological perspective, ’split subject’ covers a diffuse and
fairly diverse range of languages, without a very clear distinction from ergative at
one end of the range to accusative at the other.

Note, incidentally, that in both plots a fairly large number of languages cluster
at the extreme end of the accusative range but few at the extreme end of the
ergative range: Figure 5.2 has only one pure ergative language with exclusively S.o
marking, but 14 with exclusively S.a marking, and Figure 5.3 has no languages with
exclusively non-A coding but three with exclusively A coding. Partly this is because
fewer ergative than accusative languages were surveyed, but partly it seems to be
because ergativity is generally less thoroughgoing than accusativity (Silverstein
1976 observes that ergative languages usually have at least some splitting, while
accusative languages often have none; Nichols 1993 finds good support for this
claim in a survey of 194 languages).

5.3.4 Split subjecthood and object alignment

The main reason for grouping O and G coding together in defining split subject-
hood is that these prove to be in a sort of complementary distribution based on
object alignment: direct-object languages are likely to use G coding on their non-
A-coded subjects, while primary-object languages are likely to use O coding. (17)
and (18) show the numbers of languages with low and high numbers of O-coded
and G-coded subjects, broken down by direct vs. primary object. Both break the
range between 1.5 and 3, where there is a gap. The same asymmetry is observed in
the much larger Autotyp database (Bickel and Nichols 2002–), shown in (19).

(17) Number of languages with low (0–1.5) or higher (3 or more) numbers of So
and Ao verbs.

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>DO</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>PO</td>
<td>3</td>
<td>8</td>
</tr>
</tbody>
</table>

(18) Number of languages with low or higher numbers of Sb and Ab verbs
(cutoff points as in (17))

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>DO</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>PO</td>
<td>14</td>
<td>0</td>
</tr>
</tbody>
</table>

^3 One expected value is under 5 (high × PO is exp. 3.23), which can weaken the validity of the test.
However, this is because of the categorical nature of the distribution, which strengthens the validity.
(19) Stative-active vs. other (accusative, ergative) subject alignment and object alignment in the Autotyp database (65 languages)

<table>
<thead>
<tr>
<th></th>
<th>St-Act</th>
<th>Ergative, Accusative</th>
</tr>
</thead>
<tbody>
<tr>
<td>DO</td>
<td>5</td>
<td>29</td>
</tr>
<tr>
<td>PO</td>
<td>13</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>(p &lt; 0.025, (X^2) [1] = 6.05)</td>
<td></td>
</tr>
</tbody>
</table>

Both subject and object alignment have uneven geographical distributions, shown in (20)–(22). DO alignment predominates in Eurasia but PO elsewhere (see (20)). S.o and A.o coding (shown in (21)) are geographically asymmetrical in the same way, with O-coding predominating outside Eurasia, while S.g and A.g coding (shown in (22)) are also uneven but in inverse geographical distribution, G-coding predominating in Eurasia.

(20) Object alignment in three large areas

<table>
<thead>
<tr>
<th></th>
<th>DO</th>
<th>PO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eurasia</td>
<td>18</td>
<td>2</td>
</tr>
<tr>
<td>Pacific</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Americas</td>
<td>4</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>(p &lt; 0.005, (X^2) [2] = 11.01)</td>
<td></td>
</tr>
</tbody>
</table>

(21) Low and high numbers of verbs with O-coded subjects. Low, high as in (17) and (18).

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eurasia</td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td>Pacific</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Americas</td>
<td>0</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>(p &lt; 0.025, (X^2) [2] = 9.13)</td>
<td></td>
</tr>
</tbody>
</table>

(22) Low and high numbers of verbs with G-coded subjects

<table>
<thead>
<tr>
<th></th>
<th>Low</th>
<th>High</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eurasia</td>
<td>12</td>
<td>9</td>
</tr>
<tr>
<td>Pacific</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Americas</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>(p &lt; 0.025, (X^2) [2] = 8.23)</td>
<td></td>
</tr>
</tbody>
</table>

Thus, the uneven geographical distribution of classic stative-active languages worldwide is connected to, and probably a consequence of, large areal differences in object alignment. The reason for regarding object alignment as the causal factor is that non-A-coded subject alignment can then be expressed as a grammatical generalization: the preferred non-A subject coding is the one identical to the treatment of recipient or goal arguments of ditransitives: primary object (where O=G) in classic stative-active (S.o, A.o) languages, indirect object (where O\(\neq\)G) in dative-subject (S.g, A.g) languages. That is, the broader tendency is what Bickel (2003) calls *experiencer as goal*. It is a tendency but not a mechanical grammatical requirement, as both kinds of non-A subject coding can be found in the same

\(^4\) Low expected values for PO in the Pacific and Americas. On a two-way breakdown (Eurasia vs. other), which removes the problem, p < 0.001.

\(^5\) Low expected values for High in Pacific and Americas. On a two-way breakdown (Eurasia vs. other), which raises the expected values, p < 0.025.
language and both are well represented in several: Ingush, Lezgi, Georgain, Karata, Hindi, Chickasaw. (All of these are direct object except that Hindi is split primary/direct object.)

Thus, S.o and S.g—and to a considerable extent S.poss, though that has been less investigated here—have similar distributions across verbs and verb classes and are in roughly complementary distribution across languages (largely predicted by object alignment) and geographically. Recall also (section 5.3.2 and Appendix 5.3) that it is chiefly semantic experiencers that are most prone to have non-A coding, whether S.o or S.g. In all of these respects G-coded subjects are little different from the 'statives' of stative-active languages.

5.4 Conclusions

Why are stative-active languages rare in Eurasia? On the basis of what has been argued here, three different answers might be given to this question. The first is that they are not in fact rare in Eurasia; S.g and S.poss, which are variants or counterparts or allo-codings of S.o, are common in Eurasia, where they take the form of dative experiencer subjects.

A different, narrower answer can be given using the classical definition of stative-active and excluding S.g marking: they are rare in Eurasia because primary object alignment is rare there.

A third answer would be that they are only rare in northern Eurasia. S.g coding of experiencer subjects is common across southern Eurasia from the Pyrenees through the Caucasus to the Himalayas and South Asia. There is a northward extension in the form of Germanic and Balto-Slavic, but the north central and northeast of Eurasia (including Siberia, Manchuria, Mongolia, and Central Asia) is almost entirely lacking in oblique subject marking of any kind.

Africa is not rich in split-subject coding (neither Comrie 2005 nor Siewierska 2005 shows any tokens in Africa), but the pattern has been reported from there. In Orig (Kordofanian; Schadeberg and Elias 1979: 53), the two verbs ‘afraid’ and ‘hungry’ have what appears to be object agreement. In Loma (Mande; Rude 1983; Arkadiev, this volume), pronoun inflection is stative/active in some tenses (Rude 1983: 279–81), a distinction applying only to intransitives. The only verbs illustrated by Rude are ‘go’ (active) and ‘be big’ (stative). The only experiencer-subject verb illustrated in the paper is ‘see’, which is an ordinary transitive. In Yoruba (Benue–Kwa, Nigeria; Awoyale 1997: 11–13), several verbs have object-coded subjects (e.g. ‘feel sorry (for)’ [‘mercy did me’ = ‘I feel sorry’], ‘suffer’, ‘be ashamed’, ‘laugh’) or possessor-coded subjects (e.g. ‘be frightened’ [‘I am frightened’ = ‘my chest is broken’], ‘be deaf’, ‘be happy’), both with the stimulus coded as A. The

6 In addition, Loma verbs undergo lenition after S/O nouns, so Loma is a rare case of ergativity in Africa.
same treatment is found for several verbs in languages of the Eastern Lower Cross group (Benue-Kwa; Urua 1997: 204–5), (‘hurt’, ‘be hungry’, ‘be thirsty’).

The lexical-typological approach taken here has shown the complementarity and fundamental non-distinctness of S.o and S.g coding, and it has also shown that alignment is a continuum. Once a set of the same verb glosses is surveyed across a sample of languages, discrete types such as accusative, ergative, and stative-active begin to fade and run together. Furthermore, we have seen that, even if discreteness is not required for identifying types, stative-active or split-subject is not a third major alignment type; the difference between it and either ergative or accusative is one of degree.7

Appendix 5.1 Verb list and survey instructions

This is a pilot study for a project that will rank languages for the extent to which they are stative-active, ergative, accusative, etc. and also rank verbs for their propensity to take one or another valence pattern. The list of verbs is based on Bossong (1998), Masica (1976, 1991), Merlan (1985), Nichols et al. (2004), and the results of previous surveys.

(1) Give the language name, the family it belongs to, and the continent where it is spoken (language name alone suffices for well-known languages). Give the general subject alignment type of the language, if known, and the usual (default) treatment of S and O in the language. Give the general object alignment of the language (direct/indirect, primary/secondary), if known. Mention any general valence-relevant alternations such as tense-based split ergativity. List bibliographical sources you use.

(2) For each of the English verb glosses listed below, try to find the single best and most basic equivalent (ideally, a high-frequency verb from a normal stylistic register), and for that verb give:

- The verb in its citation form, or an abstract root or stem
- If possible, a clause using the verb in the present tense (or, if appropriate, some other tense; ideally the same tense for all the verbs) with first person subject. For purposes of this survey, subject is defined as whatever is subject of the English verb.

I am grateful to several linguists who worked through the 20-verb list in their languages of expertise: Mark Donohue (Saweru, Skou, Tukang Besi, others not in this sample); Lisa Encisco (Hindi); Marc Ettlinger (Kuki Thaadow); Jisup Hong (Korean); Ruth Singer (Mawng); Maziar Doustdar Toosarvandani (Persian); Bernhard Wälchi (Latvian, Izrija Mordvin). I thank Mark Donohue, Matthew Dryer, David Harrison, Lynette Melnar, Boyd Michailovsky, Catherine O’Connor, Masayuki Onishi, and Fernando Zúñiga for other consultation and advice. Collection and analysis of Ingush data and the basics of research design for lexical typology were partly supported by National Science Foundation grants 96-16448 and 92-22294 respectively. Ruth Singer’s work on Mawng has been funded by the University of Melbourne and the Australian Research Council.
The valence: formal treatment of each core argument (e.g. case assigned by the verb, agreement). Say whether the usual coding of the subject (A, S, experiencer) for that verb is:
- same as A
- same as O (object of monotransitive; a.k.a. T)
- same as indirect object or second object of ditransitive
- same as possessor
- other (specify)

It is likely that one or more of the verbs will be not simplex verbs but phrasal constructions such as adjective with copula (as with English hungry etc.), light verb construction (e.g. draw breath for ’breathe’), etc. Choose the usual copula, auxiliary, or light verb for this lexeme, attempt to describe the valence of the entire phrase, and identify the elements of the phrase.

If the language has primary/secondary rather than direct/indirect object alignment, indicate that and make any notes needed to be sure your descriptions of valence are unambiguous.

List of verbs
This list is for use in surveys that focus on lexical semantics and subject role, not aspect or Aktionsart. For any and all of them, assume that either an imperfective verb (stative, durative, etc.), or a perfective one (ingressive, punctual, etc.) will do equally well. (If you find both and they behave alike, just list whichever is morphologically more basic.) Some of the common imperfective–perfective pairings are listed below, separated by a semicolon. If subject coding depends on aspect or Aktionsart in your language, indicate that; if this choice is available to only some of the survey verbs, it is very important to indicate which ones.

Synonymous and alternative glosses are shown with commas and parentheses. Feel free to use other synonyms. Proxies (other verbs you can use if you can’t find the survey gloss) are listed for a few that may be hard to find in some languages.

1. see
2. forget
3. remember
4. (be) hungry
5. (be) thirsty
6. (be) cold
7. (be) glad, (be) happy (about something; not just ’be cheerful’)
8. (be) sorry, regret (reaction to something, not just ’be gloomy’)
9. like
10. (be) afraid, fear
11. (be) angry
Typological perspective on split-subject marking

12. sneeze [proxies: belch, hiccup, fart]
13. breathe
14. stand; stand up
15. jump
16. fly; fly off [proxies: crawl, run]
17. fall; fall down
18. shout, yell
19. weep, cry
20. laugh

If there is more than one common, basic verb for a given sense, choose the one most likely to have a human subject. If you are still left with more than one, the choice of one or another is important only if they have different valences; if they do, list both or all verbs.

Note: Languages vary as to whether certain of these, e.g. afraid, forget, angry, are one-argument or two-argument verbs (‘be afraid’ vs. ‘be afraid of, fear’). Also, descriptions vary as to whether they describe one-argument or two-argument valence patterns. List whatever you find and indicate what (if anything) the description says about the valence.

Appendix 5.2 Languages in sample

<table>
<thead>
<tr>
<th>Stock</th>
<th>Branch</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eurasia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ingush</td>
<td>East Caucasian</td>
<td>Nakh</td>
</tr>
<tr>
<td>Lezgi</td>
<td>East Caucasian</td>
<td>Lezgian</td>
</tr>
<tr>
<td>Karata</td>
<td>East Caucasian</td>
<td>Andic</td>
</tr>
<tr>
<td>Georgian</td>
<td>Kartvelian</td>
<td>Georgian-Zan</td>
</tr>
<tr>
<td>Ossetic</td>
<td>Indo-European</td>
<td>Iranian: NE</td>
</tr>
<tr>
<td>Persian</td>
<td>Indo-European</td>
<td>Iranian: SW</td>
</tr>
<tr>
<td>Russian</td>
<td>Indo-European</td>
<td>Slavic</td>
</tr>
<tr>
<td>Latvian</td>
<td>Indo-European</td>
<td>Baltic</td>
</tr>
<tr>
<td>German</td>
<td>Indo-European</td>
<td>Germanic</td>
</tr>
<tr>
<td>French</td>
<td>Indo-European</td>
<td>Italic</td>
</tr>
<tr>
<td>Basque</td>
<td>(isolate)</td>
<td></td>
</tr>
<tr>
<td>Finnish</td>
<td>Uralic</td>
<td>Finnish: W</td>
</tr>
<tr>
<td>Erzha Mordvin</td>
<td>Uralic</td>
<td>Finnish: E</td>
</tr>
<tr>
<td>Hungarian</td>
<td>Uralic</td>
<td>Ugric</td>
</tr>
<tr>
<td>Tuva</td>
<td>Turkic</td>
<td></td>
</tr>
<tr>
<td>Nanai</td>
<td>Tungusic</td>
<td></td>
</tr>
<tr>
<td>Korean</td>
<td>(isolate)</td>
<td></td>
</tr>
<tr>
<td>Hindi</td>
<td>Indo-European</td>
<td>Indic</td>
</tr>
</tbody>
</table>

(cont.)
### Johanna Nichols

<table>
<thead>
<tr>
<th>Stock</th>
<th>Branch</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Santali</td>
<td>Austroasiatic</td>
<td>Munda</td>
</tr>
<tr>
<td>Limbu</td>
<td>Sino-Tibetan</td>
<td>Kiranti</td>
</tr>
<tr>
<td>Kuki Thaadow</td>
<td>Sino-Tibetan</td>
<td>Kuki-Chin</td>
</tr>
<tr>
<td>Pacific</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saweru</td>
<td>West Papuan</td>
<td>Yapen</td>
</tr>
<tr>
<td>Motuna</td>
<td>South Bougainville</td>
<td>Oceania</td>
</tr>
<tr>
<td>Tauya</td>
<td>Adelbert Range</td>
<td>Brahman</td>
</tr>
<tr>
<td>Yagaria</td>
<td>Gorokan</td>
<td></td>
</tr>
<tr>
<td>Tukang Besi</td>
<td>Austronesian</td>
<td>Western</td>
</tr>
<tr>
<td>Samoan</td>
<td>Austronesian</td>
<td>Polynesia</td>
</tr>
<tr>
<td>Skou</td>
<td>Macro-Skou</td>
<td>Western Skou</td>
</tr>
<tr>
<td>Kayardild</td>
<td>Tangtic</td>
<td></td>
</tr>
<tr>
<td>Nunggubuyu</td>
<td>Gunwinyguan</td>
<td>Ngandi-Nunggubuyu</td>
</tr>
<tr>
<td>Mawng</td>
<td>Iwaidjan</td>
<td></td>
</tr>
<tr>
<td>Americas</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haida</td>
<td>(isolate)</td>
<td></td>
</tr>
<tr>
<td>Karok</td>
<td>(isolate)</td>
<td></td>
</tr>
<tr>
<td>Lakhota</td>
<td>Siouan-Catawba</td>
<td>Siouan</td>
</tr>
<tr>
<td>Caddo</td>
<td>Caddoan</td>
<td></td>
</tr>
<tr>
<td>Seneca</td>
<td>Iroquoian</td>
<td></td>
</tr>
<tr>
<td>Tunica</td>
<td>(isolate)</td>
<td></td>
</tr>
<tr>
<td>Chickasaw</td>
<td>Muskogean</td>
<td></td>
</tr>
<tr>
<td>Chontal</td>
<td>Tequislatec-Jiqaque</td>
<td>Central</td>
</tr>
<tr>
<td>Chol</td>
<td>Mayan</td>
<td></td>
</tr>
<tr>
<td>Guarani</td>
<td>Tupian</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 5.3

Subject-coding frequencies across the entire sample. Sem = semantic role of subject: Ex(periencer), Pat(ient), Ag(ent). Other = chiefly Possessor. Verbs are grouped by lexical semantics (cognition; perception; emotion; events and actions; sounds). Bold = notably high frequencies of O- and G-coded subjects.

<table>
<thead>
<tr>
<th>Gloss</th>
<th>Sem.</th>
<th>S.a</th>
<th>S.o</th>
<th>S.g</th>
<th>Other</th>
<th>A.a</th>
<th>A.s</th>
<th>A.o</th>
<th>Ag</th>
<th>Other</th>
<th>n.d.</th>
</tr>
</thead>
<tbody>
<tr>
<td>see</td>
<td>Ex</td>
<td>1.5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>35.5</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>forget</td>
<td>Ex</td>
<td>3</td>
<td>6</td>
<td>3</td>
<td>1.5</td>
<td>17</td>
<td>1.5</td>
<td>3</td>
<td>2.5</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>remember</td>
<td>Ex</td>
<td>2</td>
<td>6</td>
<td>1.5</td>
<td>3</td>
<td>19.5</td>
<td>0</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>hungry</td>
<td>Ex</td>
<td>13.5</td>
<td>13.5</td>
<td>4.5</td>
<td>8.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>thirsty</td>
<td>Ex</td>
<td>12.5</td>
<td>9</td>
<td>5.5</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>cold</td>
<td>Ex</td>
<td>11.5</td>
<td>13</td>
<td>10.5</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>glad</td>
<td>Ex</td>
<td>14.5</td>
<td>11</td>
<td>6</td>
<td>3</td>
<td>4</td>
<td>0</td>
<td>0.5</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>sorry, regret</td>
<td>Ex</td>
<td>14</td>
<td>9.5</td>
<td>4</td>
<td>1</td>
<td>2.5</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>like</td>
<td>Ex</td>
<td>3.5</td>
<td>3.5</td>
<td>2</td>
<td>5</td>
<td>15.5</td>
<td>1</td>
<td>3</td>
<td>7.5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>afraid</td>
<td>Ex</td>
<td>14</td>
<td>8</td>
<td>2.5</td>
<td>1</td>
<td>10</td>
<td>3.5</td>
<td>0.5</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>angry</td>
<td>Ex</td>
<td>17.5</td>
<td>10.5</td>
<td>0.5</td>
<td>4.5</td>
<td>4</td>
<td>1.5</td>
<td>0.5</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>sneeze</td>
<td>Pat</td>
<td>28</td>
<td>7</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>fall</td>
<td>Pat</td>
<td>31</td>
<td>13</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>breathe</td>
<td>Pat/Ag</td>
<td>30.5</td>
<td>5.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>stand</td>
<td>Ag/Pat</td>
<td>30</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>jump</td>
<td>Ag</td>
<td>29</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>fly</td>
<td>Ag</td>
<td>31</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>weep</td>
<td>Pat/Ag</td>
<td>30</td>
<td>6.5</td>
<td>0</td>
<td>1</td>
<td>1.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>laugh</td>
<td>Ag/Par</td>
<td>32.5</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>1.5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>shout</td>
<td>Ag</td>
<td>32.5</td>
<td>5.5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>TOTAL</td>
<td>10175</td>
<td>378</td>
<td>155.5</td>
<td>43</td>
<td>41.5</td>
<td>113</td>
<td>9.5</td>
<td>10.5</td>
<td>18</td>
<td>3</td>
<td>36</td>
</tr>
<tr>
<td>Experiencers</td>
<td>518.5</td>
<td>107.5</td>
<td>90</td>
<td>41</td>
<td>39.5</td>
<td>108</td>
<td>9.5</td>
<td>10.5</td>
<td>18</td>
<td>3</td>
<td>19</td>
</tr>
</tbody>
</table>
6

Losing semantic alignment: from Proto-Yeniseic to Modern Ket

EDWARD J. VAJDA

6.1 The Yeniseic microfamily

Yeniseic languages were once spoken across south-central Siberia, from Western Mongolia and the Altai northward throughout the Yenisei watershed. The family’s six attested languages form three primary nodes: Ket-Yugh, Kott-Assan, and Arin-Pumpokol. Ket is the sole surviving member today, with fewer than 100 fluent speakers. Of its extinct relatives, substantial grammatical descriptions exist only for Yugh, which survived into the 1970s (Werner 1997b), and Kott, which disappeared by 1850 (Castrén 1858; Werner 1997a). Assan, Pumpokol, and Arin are attested by word lists compiled by 18th-century travellers (Khelimskij 1986, Werner 2005). These lists include a modest number of finite verb forms, so that family-internal comparisons yield a fairly clear picture of Proto-Yeniseic verb morphology.

Verb agreement in Proto-Yeniseic is a crucial starting point for understanding the development of typologically unusual features in the daughter languages, particularly in Modern Ket, where subject/object agreement defies overall characterization under any recognized alignment type (cf. Vajda 2001a, 2003, 2004, 2006). Agreement is not conditioned by grammatical alignment, which reveals a mixture of active and nominative traits (Werner 1995). Rather, each verb selects one of several available affix configurations for reasons largely opaque from a synchronic perspective. Ket contains seven productive agreement marker configurations (two transitive and five intransitive), along with a residue of unproductive patterns. Four patterns require subject person to be marked simultaneously in two separate positions, while other verbs mark it only once. The transitive patterns contrast according to the position and shape of their object marker. The two lexically competing object series do not correlate with any parallel contrast in semantic roles. A verb's choice of agreement configuration is linked, however, with an assortment of formal and semantic traits that hint at the origins of each pattern. The extinct Yeniseic languages show a broadly analogous system of verb-internal
agreement, but with important differences that offer further evidence of how this unusual system developed.

By determining which morphological traits represent archaic retentions in the daughter languages and which are innovative, it is possible to explain how and even why Modern Ket lost the earlier Yeniseic system of semantic alignment. The data below are presented inductively. Section 6.2 suggests a reconstruction of Proto-Yeniseic verb morphology, which is argued to follow active/stative agreement, with verb-internal marking of 1st and 2nd person undergoer subjects plus classifiers distinguishing animate- from inanimate-class subjects. Section 6.3 discusses the system that developed from this original prototype in each of the three primary nodes of the Yeniseic family, beginning with Kott/Assan, which preserved the most archaic features, then moving on to Ket/Yugh, the best-described system. The focus is mainly on Kott and Ket. A brief examination of the scant data available from Arin and Pumpokol reinforces the claims being made. Section 6.4 explains why the original active/stative system of Proto-Yeniseic came to be replaced in Ket by a system of typologically anomalous agreement marker configurations. These findings are then considered in light of what is know cross-linguistically about the dynamics of alignment shifts.

6.2 Verb morphology and agreement typology in Proto-Yeniseic

A comparison of the finite verb morphology in the Yeniseic daughter languages suggests the morphological model for Proto-Yeniseic shown in Table 6.1.

Proto-Yeniseic possessed a type of active/stative agreement whereby only undergoer subjects were marked verb-internally. Pronominal NPs correlating with active subjects or objects did not originally form part of the morphological verb, though they may have been able to cliticize to it in certain environments. Pronominal indexes of active intransitive or transitive subjects eventually became prefixes or proclitics in some daughter languages (Ket, Yugh, Arin) and suffixes in others (Kott, Assan, Pumpokol).

Proto-Yeniseic appears to have overtly marked agreement with speech-act-participant subjects only. Agreement with a 3rd-person subject was expressed

<table>
<thead>
<tr>
<th>Table 6.1. The Proto-Yeniseic verb complex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morphemes outside the phonological verb</td>
</tr>
<tr>
<td>Subject NP</td>
</tr>
</tbody>
</table>
indirectly by the presence of animacy classifiers placed farther from the verb root (position 4). The original classifier system is best preserved in Kott.

Some intransitive undergoers in Proto-Yeniseic were expressed as pronouns followed by a postposition, rather than by verb-internal agreement markers. The original pronoun + postposition became a proclitic in Kott, but fused to the morphological verb in Ket-Yugh as a prefix position. In Modern Ket, subject referencing in this prefixal position (position six) became obligatory for most stems where the morpheme expressing the verb’s basic meaning (i.e. the semantic head) stood word-initially rather than word-finally. The shift of semantic head from the rightmost slot (in Proto-Yeniseic) to the leftmost lexical slot in most productive patterns of Ket verb formation appears to have developed through prolonged areal influence from the suffixing-agglutinative languages (Samoyedic, Turkic, Tungusic) spoken by neighbouring groups. This shift, with its concomitant incorporation of new prefixal positions, led to the rise of several different models of verb-internal agreement in Modern Ket.

6.3 Verb agreement across the Yeniseic daughter languages

The contrast between Kott and Ket agreement-marking patterns, discussed in the next two subsections, nicely illustrates the evolution of agreement in Yeniseic. Because Assan closely resembles Kott, and Yugh is similar to Ket, these two languages are not examined separately. Arin and Pumpokol differ more significantly, and what little is known about their agreements systems is briefly considered in a separate subsection. The discussion begins with Kott, which preserves the greatest overall similarity to the original Proto-Yeniseic system.

6.3.1 Agreement marking in Kott

Aside from its innovative subject agreement suffixes, Kott verb morphology appears best to preserve the functional system of the original Proto-Yeniseic core prefix positions. The documented subtypes of Kott agreement marker configurations demonstrate that active/stative and animate/inanimate-class distinctions have retained a more central role than is true of Ket. Table 6.2 shows the position classes of the Kott verb.

Table 6.3 shows the markers attested for each agreement position along with their possible grammatical functions, which depend on the lexical verb in which they appear. The lack of any attested gender distinction in reflexes of masculine and feminine arguments suggests gender marking—restricted as it is to 3rd person terms—was absent in the Kott verb and probably in Proto-Yeniseic as well. The animate/inanimate-class distinction, however, was consistently marked, as was animate-class plurality. Note that the plural subject suffixes contain two
<table>
<thead>
<tr>
<th>Table 6.2: Position classes in Kott finite verb form creation</th>
</tr>
</thead>
<tbody>
<tr>
<td>P7</td>
</tr>
<tr>
<td>Left base (verb root or incorporate)</td>
</tr>
</tbody>
</table>
Edward J. Vajda

Table 6.3. Kott subject/object agreement affix shapes

<table>
<thead>
<tr>
<th>Position and agreement functions</th>
<th>P6 (person/number)</th>
<th>P4 animacy classifier marks all animate Sa in some verbs; in others, only 3p Sa or 3p inan. O</th>
<th>Pt (person/number)</th>
<th>P−1 (person/number)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>most O; some Sa and So</td>
<td></td>
<td>most subjects</td>
<td>most subjects</td>
</tr>
<tr>
<td>1s</td>
<td>aη ~ ey</td>
<td>d1 ~ O</td>
<td>i</td>
<td>η ~ aη</td>
</tr>
<tr>
<td>2s</td>
<td>u</td>
<td>d1 ~ O</td>
<td>i</td>
<td>u</td>
</tr>
<tr>
<td>3m</td>
<td>a</td>
<td>d1</td>
<td>a</td>
<td>Ø</td>
</tr>
<tr>
<td>3f (not attested)</td>
<td>(not attested)</td>
<td>b</td>
<td>a</td>
<td>Ø</td>
</tr>
<tr>
<td>3n (s or pl)</td>
<td>(not attested)</td>
<td>b</td>
<td>a</td>
<td>Ø</td>
</tr>
<tr>
<td>1pl</td>
<td>ηη</td>
<td>d1 ~ O</td>
<td>ηη</td>
<td>antoy</td>
</tr>
<tr>
<td>2pl</td>
<td>ηη</td>
<td>d1 ~ O</td>
<td>ηη</td>
<td>antoy</td>
</tr>
<tr>
<td>3an.pl</td>
<td>ηη</td>
<td>d1</td>
<td>ηη</td>
<td>an</td>
</tr>
</tbody>
</table>

Parts—the plural subject marker -an followed by the person markers -toy for 1st plural, -oη for 2nd plural, and Ø for 3rd plural.

Tables 6.4–6.8 provide sample paradigms illustrating the various transitive and intransitive agreement marker configurations.

Transitive stems in Kott follow two patterns with clear analogues to productive configurations in Ket and Yugh. The first uses intransitive-class marking b, just like Ket Transitive I (see next section). Unfortunately, the documented forms in Kott attest only the object markers a ‘him’ and b ‘it’. Animacy-classifying d1 does not appear in Kott Transitive I verbs, and 3rd person objects trigger the marker a followed by a glottal stricture that is possibly cognate with one of the P5 thematic consonants of Ket. The affix b classifies the object as inanimate. In stems beginning in a thematic consonant (either a shape classifier or a spatial preverb), b elides,

Table 6.4. Kott Transitive Configuration I

<table>
<thead>
<tr>
<th>Subject</th>
<th>‘beat’ (1st, 2nd person object not attested)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>a²atb⁶iyan (I beat him)</td>
</tr>
<tr>
<td>2sg</td>
<td>batb⁶itu (you.s beat it)</td>
</tr>
<tr>
<td>3jan.sg</td>
<td>a²atb⁶ix (he/she beats him)</td>
</tr>
<tr>
<td>3n.sg</td>
<td>batb⁶ix (he/she beats it)</td>
</tr>
</tbody>
</table>

a The forms aatb⁶itnøy (<cetb⁶m-anoy) ‘you.pl. beat him’ and batb⁶itu (< batb⁶m-an) ‘they beat them’ involve dissimilation of n to t before another n, a common occurrence in Kott verb suffixes.
Table 6.5. Kott Transitive Configuration II (inanimate-class subjects or objects not attested)

<table>
<thead>
<tr>
<th>Subject</th>
<th>Intransitive Configuration I</th>
<th>'think'</th>
<th>'thought'</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>anapajakj</td>
<td>anayolok</td>
<td>eneiques</td>
</tr>
<tr>
<td>2sg</td>
<td>anapajakuru</td>
<td>anayoloku</td>
<td>eneiques</td>
</tr>
<tr>
<td>3an.sg</td>
<td>anapajak</td>
<td>anayolox</td>
<td>ena'i美学</td>
</tr>
<tr>
<td>1pl</td>
<td>anapajakantong</td>
<td>anayolokantong</td>
<td>enopique</td>
</tr>
<tr>
<td>2pl</td>
<td>anapajakantony</td>
<td>anayolokany</td>
<td>enopique</td>
</tr>
<tr>
<td>3an.pl</td>
<td>anapajakkan</td>
<td>anayolakan</td>
<td>enapique</td>
</tr>
</tbody>
</table>

labializing the following tense/mood vowel: *d'ok'kiaŋ 'I bend it' (< *d'/b-a-kiyaj). Otherwise o appears only in the past indicative. Instances of phonological-verb initial o in present-tense Kott forms may likewise be due to a verb-initial thematic consonant which later elided: ofulaŋ 'I weave it' (< *C-b-a fujaj). This possibility is supported by the fact that labialization disappears in the imperative, where the marker b normally elides: d'an'kiŋ 'Bend it!', an'fuj 'Weave it!' Otherwise, inanimate-class object b does not disappear in morphological verb-initial position, nor does it labialize the following vowel in this position: b'ath ix 'he/she beats it'. It is possible that object-marking b originally existed only in the 3rd person forms, where it helped rank potency between 3rd person terms, like the obviative system present in Modern Athabaskan, eventually becoming a generic object marker in stems requiring animacy classifiers.

Kott Transitive II also uses the P6 slot to mark the object. This pattern is newer and appears to have arisen with the incorporation into the verb of a new left base. Verbs in which this left base came to express the verb’s semantic head dispensed with animacy classifiers, expressing all objects with pronominal markers followed

Table 6.6. Agreement in Kott intransitives (without animacy classifiers)

<table>
<thead>
<tr>
<th>Subject</th>
<th>Intransitive Configuration I</th>
<th>'forget'</th>
<th>'forgot'</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>anekajakj</td>
<td>aenayolok</td>
<td>enayiques</td>
</tr>
<tr>
<td>2sg</td>
<td>anekajakuru</td>
<td>aenayoloku</td>
<td>enayiques</td>
</tr>
<tr>
<td>3an.sg</td>
<td>anekajak</td>
<td>aenayolox</td>
<td>ena'i美学</td>
</tr>
<tr>
<td>1pl</td>
<td>anekajakantong</td>
<td>aenayolokantong</td>
<td>enopique</td>
</tr>
<tr>
<td>2pl</td>
<td>anekajakantony</td>
<td>aenayolokany</td>
<td>enopique</td>
</tr>
<tr>
<td>3an.pl</td>
<td>anekajakan</td>
<td>aenayolakan</td>
<td>enapique</td>
</tr>
</tbody>
</table>
Edward J. Vajda

Table 6.7. Intransitive Configuration III (animacy classifiers underlined)

<table>
<thead>
<tr>
<th>Subject</th>
<th>'mature'</th>
<th>'matured'</th>
<th>'become'</th>
<th>'became'</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>ḅapijaŋ</td>
<td>anapijaŋ</td>
<td>ḅajaŋ</td>
<td>onajaŋ</td>
</tr>
<tr>
<td>2sg</td>
<td>ḅapiju</td>
<td>anapiu</td>
<td>ḅaja</td>
<td>onau</td>
</tr>
<tr>
<td>3an.sg</td>
<td>ḅapi</td>
<td>anapi</td>
<td>ḅax</td>
<td>onax</td>
</tr>
<tr>
<td>3n.sg</td>
<td>ḅapi</td>
<td>manapi</td>
<td>(not attested)</td>
<td></td>
</tr>
<tr>
<td>1pl</td>
<td>ḅapijaŋoŋ</td>
<td>anapijaŋoŋ</td>
<td>ḅajekantoŋ</td>
<td>on'egoŋantoŋ</td>
</tr>
<tr>
<td>2pl</td>
<td>ḅapijaŋoŋ</td>
<td>anapijaŋoŋ</td>
<td>ḅajekantoŋ</td>
<td>on'egoŋantoŋ</td>
</tr>
<tr>
<td>3an.pl</td>
<td>ḅapijan</td>
<td>anapijan</td>
<td>ḅajekan</td>
<td>on'egoŋan</td>
</tr>
</tbody>
</table>

by a thematic consonant (usually $t^b$). This pattern is restricted to left-headed stems, though a few stems that fill the left base belong to Transitive I instead. Subject person is marked by the same suffix series in P—1, again preceded by the plural morpheme -an.

Again, feminine- and inanimate-class objects are not attested. If the Kott verb did indeed lack a gender agreement distinction, it is possible that a or any 3rd person singular animate-class object. It is also possible that the inanimate-class object marker was zero, as it is in the Modern Ket analogue to this configuration.

Kott intransitives comprise several subtypes with clear parallels to Ket. Intransitives I verbs may or may not contain a left base. They express active events that logically require a sentient or at least an active subject (e.g. 'think', 'live'). They contain no animacy classifier, and express subject agreement exclusively through the suffix in P-1. Intransitive II appears limited to inactive verbs containing a left base.

Intransitive III is confined to inactive verbs without an incorporate. Such verbs can have either sentient or non-sentient subjects and require an animacy classifier in the non-past: $d^j$ for all animate-class subjects, including 1st and 2nd person subjects, and $b$ for a 3rd person inanimate-class subject. The animacy classifier $d^j$ deletes in the past tense and imperative forms. But inanimate-class $b$ remains in

Table 6.8. Kott Intransitive Configuration IV

<table>
<thead>
<tr>
<th>Subject</th>
<th>'lie down'</th>
<th>'lay'</th>
<th>'get born'</th>
<th>'got born'</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>īteŋaŋ</td>
<td>aŋteŋaŋ</td>
<td>ījeŋaŋ</td>
<td>aŋjeŋaŋ</td>
</tr>
<tr>
<td>2sg</td>
<td>īteŋu</td>
<td>aŋteŋu</td>
<td>ījeŋa</td>
<td>aŋjeŋu</td>
</tr>
<tr>
<td>3an sg</td>
<td>ɖateŋ</td>
<td>aŋteŋ</td>
<td>ɖaŋjeŋ</td>
<td>aŋjeŋ</td>
</tr>
<tr>
<td>1pl</td>
<td>ɖɛteŋaŋoŋ</td>
<td>aŋteŋaŋoŋ</td>
<td>ɖaŋjeŋaŋoŋ</td>
<td>aŋjeŋaŋoŋ</td>
</tr>
<tr>
<td>2pl</td>
<td>ɖɛteŋaŋoŋ</td>
<td>aŋteŋaŋoŋ</td>
<td>ɖaŋjeŋaŋoŋ</td>
<td>aŋjeŋaŋoŋ</td>
</tr>
<tr>
<td>3an pl</td>
<td>ɖanteŋaŋan</td>
<td>aŋteŋaŋan</td>
<td>ɖaŋjeŋ</td>
<td>aŋjeŋ</td>
</tr>
</tbody>
</table>
the past tense, where it assimilates to \( m \) before the past tense marker \( n \) (\( \text{bapi} \) ‘it matures’ \( \rightarrow \) \( \text{manapi} \) ‘it matured’) but remains unchanged in forms with \( l \) as the marker of past tense (\( \text{bafuta} \) ‘I paint it’ \( \rightarrow \) \( \text{balafuta} \) ‘I painted it’). Imperatives from inanimate-class subject verbs of this type are not attested, but the inanimate-class object marker \( b \) drops in transitive imperative forms (\( \text{anath} \) ‘Beat him/it!’), just as it does in Ket and Yugh.

Kott Intransitive IV contains change-of-state and resultative verbs. Multi-site subject number marking is highly elaborated, with certain forms containing as many as four indications of plurality. This includes one prefix and two suffixes, with the base in some stems further augmented by the distributive suffix \(-\text{aj}: \text{ontenagantaj} \) ‘we lie down’ [\(1\text{-}\text{ten}/\text{af}^3-\text{an}/\text{ot}^1\) \(1\text{pl} \text{sbj}\)]-lie/distributive\(^3\)-\(1\text{pl} \text{sbj}\)-\(1\text{pl} \text{sbj}\)]. Finally, the animacy classifier \( d \) appears in the non-past indicative, but only in the 3rd person forms. Since no inanimate-class subjects are attested for verbs of this pattern, it is unclear whether the inanimate classifier \( b \) plays a role in this configuration (all verbs attested logically require animate-class subjects). Interestingly, the homologous animacy classifier \( d \) of Ket and Yugh is an object rather than subject classifier. This vestigial morpheme appears only in a few Ket Intransitive I stems, where it classifies 1st and 2nd person as well as 3rd person objects. Unlike Kott, the Ket animacy classifier remains in all past tense and imperative forms: e.g., Ket \( \text{kuld} \text{ldadgigf you.sg} \) washed us [\(\text{kaf}^3-\text{ul}^7-\text{d}^5-\text{aj}^1-\text{kat}^j_2\text{sb}^b\text{f}-\text{water}^2-\text{ac}^5-\text{past}^1/2-\text{isg.obj}^3\)-wash\(^6\)\)]], \( \text{uldalaj} \) ‘Wash him! (you.sg)’ [\(\text{ul}^7-\text{d}^5-\text{a}^4-\text{al}^2-\text{kat}^j_2\text{water}^2-\text{ac}^5-\text{3m.obj}^3\)-imp\(^5\)-wash\(^6\)\)]

Table 6.8 gives two Intransitive IV verbs in Kott, the first of which contains a base with the distributive augment \(-\text{aj}\). In this table, subject agreement prefixes are underlined. Inanimate-class subjects do not seem to occur in this agreement type, which may be connected to the confinement of animacy classifiers to forms with 3rd person subjects. Except for a few anomalous verbs such as ‘get born’, all Kott intransitives require the plural agreement suffix \(-\text{an}\) when used with plural subjects. Given the distinctive singular and plural suffixes that reference 1st and 2nd person subjects (-\(\text{aj}\) ‘I’ vs. -\(\text{to}\) ‘we’, and -\(\text{u}\) ‘you.sg’ vs. -\(\text{oj}\) ‘you.pl’), this suffix is non-redundant only for 3rd person subjects, which are otherwise zero-marked (\(\text{anayajak} \) ‘he/she thinks’ vs. \(\text{anayajatkan} \) ‘they think’).

To summarize, Kott agreement could be described as a type of semantic alignment, with the added lexical idiosyncrasy of subject marking in the object position for one subclass of verbs (derived from impersonal constructions). These verbs lack a subject agreement suffix. All other stems require a subject agreement suffix. Also, the 1st and 2nd person subjects of a minority of unaccusative verbs require an agreement prefix in addition to the generic subject suffix. In this subset of verbs, animacy classifiers appear as 3rd person agreement markers in the non-past tense of verbs logically confined to animate-class subjects (Kott Intransitive IV) and in all persons of verbs capable of taking either animate- or inanimate-class subjects (Kott Intransitive III). Membership in the class of impersonal originally verbs and
their analogues, which represent the only verbs that lack a subject suffix, appears to be a morphological idiosyncrasy standing outside the language’s otherwise general rule of semantic alignment.

6.3.2 Modern Ket agreement marking
In contrast to Kott, the various Ket agreement-marking patterns can no longer be described as conforming to any overall alignment, since none of the seven productive patterns is used by more than 20 per cent of the language’s verb stems. Table 6.9 shows the position class formula followed by all finite verb forms in Ket. Yugh is broadly similar and will not be treated separately here (cf. Werner 1997b). The caption ‘valence’ indicates potential subject or object positions. Which of these positions a stem uses is a lexical idiosyncrasy, though the shape of the markers themselves predictably follows syntactic rules of agreement (cf. Table 6.10). Tense and mood marking uses a similar combination of semantically opaque morpheme shapes in positions 4 and 2, forming six productive tense/mood classes; these will not be discussed here.

In contrast to the attested Kott verbs, the vast majority of Ket verb stems are discontinuous. All productive patterns fill both P7 and P0, and many fill P5 as well, so that these three positions form the basic lexical stem. As in Kott, syllabic morphemes normally occupy P7, the left base, as well as P0, the right base (or simply the base, since this position is present in every verb while P7 is absent in most archaic stem patterns). P5 contains a non-syllabic consonantal morpheme (in a few cases two such morphemes), best glossed as TH for ‘thematic consonant’, the meaning being generally opaque.

Ket verbs possess another category of lexical morphemes in addition to the primary stem-building components occupying positions 7-5-0. Certain agreement slots (P8, P3, P1) may contain an affix that no longer varies to express grammatical agreement but merely signals an increase or decrease in the stem’s semantic valence in some way. There are three such affixes in Modern Ket, all productive at least to some degree. P8 da builds involuntary causatives in stems where the P7 incorporate names the effect produced: daqšlekša ‘it turns yellow’ [da8-qaleq7-ul6-k3-s4-a1t/sh-yellow7-3N.SBF6-TH5-PRES4-process.occurs0]. Adding P1 a and removing the P8 subject marking derives stative resultatives from most types of transitive stem that lexically require object marking in P6. The presence of P1 a (aj before the base -bet ‘make’) conveys stativity and confirms the P6 marker as a subject rather than object: daišlaksivet ‘she breaks it’ [da8-il7-ul6-k3-s4-bet3f.SBF8-small.pieces7-3N.OBJ7-TH5-PRES4-make0] → ilukšajbct ‘it is broken’ [il7-ul6-k3-s4-aj1-bet4small-pieces7-3N.SBF6-TH5-PRES4-RES1-make0]. Non-agreement P3 b can be added to increase or decrease valence in a number of ways. In some stems, it adds intensive meaning: b’γavini ‘I rush out’ [bo6-k5-a4-b1-th/1sg.SBF6-TH5-PRES4-intensive3-walk.MOM0] (cf. b’γaviti ‘I go’ [bo6-k5-a4-th/1sg.SBF6-TH5-PRES4-walk.MOM0]). In
### Table 6.9: Position classes in Ket finite verb form creation

<table>
<thead>
<tr>
<th>Position</th>
<th>Valence</th>
<th>Tense/mood</th>
<th>Thematic</th>
<th>Left base (infinitive or incorporate)</th>
<th>(Right) base (verb root or aspect/voice marker)</th>
</tr>
</thead>
<tbody>
<tr>
<td>P₀</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P₁</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P₂</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P₃</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P₄</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P₅</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P₆</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P₇</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P₈</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Position agreement functions</td>
<td>$P_8$ (P7 left base)</td>
<td>$P_6$ (P5 thematic consonant)</td>
<td>$P_4$ (3 AN class)</td>
<td>$P_3$ (P4/tense/mood)</td>
<td>$P_1$ (right base)</td>
</tr>
<tr>
<td>-----------------------------</td>
<td>---------------------</td>
<td>-------------------------------</td>
<td>--------------------</td>
<td>-----------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>All A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Some Sa and Sp</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Some P</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>Some A, Sa, and Sp</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
</tbody>
</table>

| 1sg                          | $di$ ($d, t, r$)    | $ba\sim bo$                  | $-$                 | $-$                   | $di$ ($d, t, r$) | $-$                  |
| 2sg                          | $ku$ ($k, g, \gamma$) | $ku$ ($gu, \gamma u$)       | $-$                 | $-$                   | $ku$ ($k, g, \gamma$) | $-$                  |
| 3m (sg)                     | $du$ ($d, i, r$)    | $a\sim o\sim bu$            | $-$                 | $-$                   | $a$              | $-$                  |
| 3f (sg)                     | $da$ ($d, i, r$)    | $i\sim u\sim bu$            | $-$                 | $-$                   | $a$              | $-$                  |
| 3n (sg, pl)                 | $da$ ($d, i, r$)    | $\emptyset\sim i\sim u\sim bu$ | $-$                 | $-$                   | $a$              | $-$                  |
| 1pl                          | $di$ ($d, t, r$)    | $d\emptyset (\emptyset)$    | $-$                 | $-$                   | $d\emptyset (\emptyset)$ | $-$                  |
| 2pl                          | $ku$ ($k, g, \gamma$) | $k\emptyset (g\emptyset, y\emptyset)$ | $-$                 | $-$                   | $k\emptyset (g\emptyset, y\emptyset)$ | $-$                  |
| 3an, pl                      | $du$ ($d, t, r$)    | $a\emptyset \sim o\sim bu$  | $a\emptyset \sim o\sim (a\emptyset \sim o\emptyset)$ | $-$ | $a\emptyset \sim o\emptyset (a\emptyset \sim o\emptyset)$ | $-$ |

$A =$ transitive subject, $P =$ object, $S_a =$ active intransitive subject, $S_p =$ inactive intransitive subject.
others, it functions as an instrumental applicative marker: d'əqâd'əviltay'in 'they
dragged him (by conveyance)' [daʰ-a-ii-kî-lî-tijective -tuf'-n-3M.SBJ]-3M.OBJ]-
th5-appl3-pres4-drag2-an.pl.SBJ]-1] (cf. d'əhuay 'he dragged him (neutral with
regard to instrumentality)' [daʰ-a-ii-tuf'-3M.SBJ]-3M.OBJ]-past2-drag)]. P3 b
also appears in a few involuntary causatives: b'əghimun 'I slipped' [buʰ-k3-b3-
ir2-hun0-1sg.SBJ]-th5-ic3-past2-slip]. There is no attested analogue to these
morphemes in the other Yeniseic languages.

Table 6.10 lists the phonemic and allomorphic shapes of the Ket subject/object
markers according to their present position class and potential grammatical
function. The different agreement positions obviously lack a one-to-one corres-
dence with individual semantic roles or syntactic functions, and no single rule
can be given to predict agreement strategy. Also, some stems require the subject
be marked simultaneously in two different positions, once in P8 and again in
either P6 or P1: e.g., dəbutsaq 'she makes a quick round trip' [daʰ-bu-th3-s5-
aq3-3E.SBJ]-3SN.SBJ]-th5-pres7-go.once8], datisgo 'she loads a gun (originally, puts
an arrow to a bowstring)' [daʰ-ti7-s5-a1-qo3-3E.SBJ]-bowstring2-pres4-3sg.SBJ]-
stretch9]. Multi-site subject marking is likewise a fixed lexical idiosyncrasy of the
stem that cannot be varied to convey benefactivity, version, or similar grammatical
categories.

Modern Ket contains two productive transitive configurations (Table 6.11 and
6.12) and five productive intransitive configurations. A residue of unproductive
agreement position configurations exists (cf. Vajda 2004: 70), including two tran-
sitive configurations with multi-slot subject agreement. In productive stem types
agreement is largely predictable based on combinations of semantic and morpho-
logical triggers. In unproductive stem types (e.g. stems without P7), agreement
marking is less predictable, and will be subcategorized according to whether the
given pattern is productive or not.

Let us first examine agreement configurations found among productive tran-
sitive stem types. The marked productive transitive pattern can be found in
morphological causatives made with P5 q mark their objects in P4/3/1, depend-
ing on the object's person and grammatical class. These verbs invariably con-
tain an infinitive in P7, which serves as the verb's semantic head. Causatives-of-
state verbs made with Po sin and a descriptive modifier in P7 also follow this
pattern: dı̆t̆ampin 'I fill it' [dĭ2-ir2-ti5-a3-b3-sin0 1SBJ]-full2-th5-pres4-3M.OBJ]-
cause.to.become10].

Among unproductive stem-building patterns, this agreement pattern represents
the basic, unmarked type and is therefore not associated with any particular mor-
phological or semantic trigger: dĭpipit 'he hits it' [daʰ-b3-tcf0 1SBJ]-3M.OBJ]-hit0],
kirii 'you.s dress me' [kĭ2-di5-ss2-1sg.OBJ]-dress0]. It is basically homolo-
gous with Kott Intransitive I, though stems containing animacy classifiers (an
unproductive remnant in Ket) also follow this pattern: dı̆doldiky 'he washed me'
[daʰ-ud2-d5-a4-l2-di2-kı̆p3 3M.SBJ]-wash0].
Table 6.11. Sample paradigm fragment of Ket Transitive Configuration I

<table>
<thead>
<tr>
<th>Case</th>
<th>Stem</th>
<th>Verb</th>
<th>Case</th>
<th>Form</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>d⁰</td>
<td>dəˈdəq ⁷ ³ q⁵ a⁴ y⁶ ¹ v⁳ ¹ ³ d⁰</td>
<td>1PL</td>
<td>d⁰</td>
<td>dəˈdəq ⁷ ³ q⁵ a⁴ vəq ³ d⁰ n⁻¹</td>
</tr>
<tr>
<td>2SG</td>
<td>k⁰</td>
<td>dəˈdəq ⁷ ³ q⁵ a⁴ d⁰</td>
<td>2PL</td>
<td>k⁰</td>
<td>dəˈdəq ⁷ ³ q⁵ a⁴ ³ vəq ³ d⁰ n⁻¹</td>
</tr>
<tr>
<td>3M</td>
<td>d⁰</td>
<td>dəˈdəq ⁷ ³ q⁵ a⁴ d⁰</td>
<td>3PL</td>
<td>d⁰</td>
<td>dəˈdəq ⁷ ³ q⁵ a⁴ vəq ³ d⁰ n⁻¹</td>
</tr>
<tr>
<td>3F</td>
<td>dəˈdəq ⁷ ³ q⁵ a⁴ vəq ³ d⁰</td>
<td>(she makes him laugh)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 6.12. Sample paradigm fragment of Ket Transitive Configuration II

<table>
<thead>
<tr>
<th>Case</th>
<th>Stem</th>
<th>Verb</th>
<th>Case</th>
<th>Form</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>d⁰</td>
<td>bəˈdɑːk ⁷ ³ ³ q⁵ a⁴ vəq ³ ³ d⁰</td>
<td>1PL</td>
<td>d⁰</td>
<td>bəˈdɑːk ⁷ ³ ³ q⁵ a⁴ vəq ³ ³ d⁰ n⁻¹</td>
</tr>
<tr>
<td>2SG</td>
<td>k⁰</td>
<td>bəˈdɑːk ⁷ ³ ³ q⁵ a⁴ vəq ³ ³ d⁰</td>
<td>2PL</td>
<td>k⁰</td>
<td>bəˈdɑːk ⁷ ³ ³ q⁵ a⁴ vəq ³ ³ d⁰ n⁻¹</td>
</tr>
<tr>
<td>3M</td>
<td>d⁰</td>
<td>bəˈdɑːk ⁷ ³ ³ q⁵ a⁴ vəq ³ ³ d⁰</td>
<td>3PL</td>
<td>d⁰</td>
<td>bəˈdɑːk ⁷ ³ ³ q⁵ a⁴ vəq ³ ³ d⁰ n⁻¹</td>
</tr>
<tr>
<td>3F</td>
<td>dəˈdəq ⁷ ³ ³ q⁵ a⁴ vəq ³ ³ d⁰</td>
<td>(she finds it)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The unmarked transitive agreement pattern in productive stem types (Ket Transitive II) uses the P8 + P-1 circumflex to mark the subject, and P6 to mark the object. This pattern occurs in all productive types of transitive stems except those that contain P3 causative q.

Note that in transitive forms of any type, inanimate-class subjects are rare and take the same morpheme as feminine-class transitive subjects.

The basic nature of this pattern in Modern Ket is underscored by the fact that transitive stems deriving from Russian invariably conform to it, with the borrowed material (normally an infinitive form) appearing in P7. Examples include: 

1. **dalabíthíŋyavat** ‘she loves me’ (from Russian l’ubit’ ‘to love’),
2. **dapl’omzaríŋyavat** ‘he (or it) helps her’ (from Russian pomogat’ ‘to help’).

A few unproductive stem types lacking P7 also require this agreement marker configuration, suggesting that this pattern was originally associated with some sort of morphosemantic trigger. In Modern Ket these are virtually limited to verbs of seeing:

- **dabáťtaŋ’u** ‘she saw me’ [daño-6-β5-t3-o4-t2-ŋ3m.sbj3-1sg.obj6-th6-past4-see6] and verbs in which the instrument role is somehow overtly marked. The latter group includes instrumental applicatives built with P3 b: k’agdáptan ‘you.sg drag him’ (specifically using some conveyance)’ 
  - **ka-no-t3-ta-ŋ3m.obj6-th6-pres4-appl3-drag6**. Compare the more generic d’šmat ‘he dragged him’ [duño-6-β4-t2-taŋ3m.obj6-th6-past4-drag6], which follows Transitive I, as do most transitive stems lacking P7.

Among unproductive patterns, a few Ket transitives require subject person to be marked twice. There are two such double-slot subject agreement configurations. One has a sizeable number of stems and can be called Transitive III. Like be marked twice. There are two such double-slot subject agreement configurations. One has a sizeable number of stems and can be called Transitive III. Like

- **diŋkďátka’‘I sell it’ (etymologically, ‘I extend my hand with it’) [dpré-β6-k5-di-ŋ3m.obj6-3n.obj6-with3-sbj1-extend.hand6]. Another uses this pattern for plural subjects only: 
  - **db’oktaŋ’utn ‘they lead me around’ [dubno-β6-k1-t3-ŋ3m.obj6-with/atalik3-pres5-one.walks5],**
  - **d’oktaŋ’utn ‘they lead me around’ [dubno-β6-k1-t3-ŋ3m.obj6-with/atalik3-pres5-one.walks5],**

Productive intransitive stem types show one unmarked pattern and five patterns associated with specific formal or semantic triggers. The unmarked pattern

- **k’gadíuua, ‘I sell it’ (etymologically, ‘I extend my hand with it’) [dpré-β6-k5-di-ŋ3m.obj6-3n.obj6-with3-sbj1-extend.hand6]. Another uses this pattern for plural subjects only:**
  - **db’oktaŋ’utn ‘they lead me around’ [dubno-β6-k1-t3-ŋ3m.obj6-with/atalik3-pres5-one.walks5],**
Table 6.33. Sample paradigm fragment of Ket Transitive Configuration III

<table>
<thead>
<tr>
<th></th>
<th>-outside³ -sbj⁶ -th³ -obj⁶/3/1 -take⁰ -an.pl -sbj⁻¹ 'lead, take outside'</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>d₈ $\lambda$a₁b₂g₃gu₄hus⁰ (I take you.sg)</td>
</tr>
<tr>
<td>2SG</td>
<td>k₈ $\lambda$d₃b₄g₅d₆hus⁰ (you.sg take me)</td>
</tr>
<tr>
<td>3M</td>
<td>d₈ $\lambda$ak₃b₄g₅d₆hus⁰ (he takes him)</td>
</tr>
<tr>
<td>3F</td>
<td>d₈ $\lambda$d₄b₅g₆h₇(i)hus⁰ (she takes it)</td>
</tr>
<tr>
<td>1PL</td>
<td>d₈ $\lambda$d₅b₆g₇gu₈hus⁻¹ (we take you.pl)</td>
</tr>
<tr>
<td>2PL</td>
<td>k₈ $\lambda$d₉b₁₀g₁₁d₁₂hus⁻¹ (you.pl take us)</td>
</tr>
<tr>
<td>3AN.PL</td>
<td>d₈ $\lambda$d₄b₅g₆h₇a₈g₉hus⁻¹ (they take them)</td>
</tr>
</tbody>
</table>
consists of a subject agreement marker in P₃ for most inanimate-class subjects, and in P₈ + P₋₁ for animate-class (cf. Table 6.14).

Intransitive verbs based on Russian loans normally conform to this pattern, with the borrowed element appearing as the verb’s semantic head in P₇: *dál’-lét-atet' ‘she works’ (*< Russian *rabota ‘work’) [da⁽⁸⁻*lóbet⁽¹⁻*-bet⁽⁶⁽³F,SBF⁾⁻*work⁽⁷⁻*PRES⁽⁴⁽³F⁾⁻*ATELIC⁽³F]*]. Most intransitives without P₇ also follow this pattern, though a significant number do not, as will be shown below.

Each of the remaining five productive intransitive patterns is associated with a particular morphological or semantic feature, at least for stems containing P₇.

The second intransitive pattern is likewise extremely widespread. Several structurally distinct types of inanimate stem cross-reference their subject in P₆. These include intransitives that name the resultant state by using a noun or infinitive, a change-of-state verb that names the resultant state by using a noun or infinitive, a change-of-state verb that names the resultant state by using a noun or infinitive, and a change-of-state verb that names the resultant state by using a noun or infinitive. These intransitives may contain an adjective root in P₈, however, belong to Intransitive I: *dákqáyuan ‘she gets big’ [da⁽⁸⁻*qa⁽¹⁻*⁻*qan⁽³F,SBF⁽⁴⁻*big⁽⁷⁻*PRES⁽⁴⁻*INCEP⁽³F]*], *qáyuan ‘it gets big’ [qa⁽¹⁻*⁻*qan⁽³F,big⁽⁷⁻*PRES⁽⁴⁻*AN.SBF⁽⁴⁻*INCEP⁽³F]*]. This fact underscores the morphological nature of agreement marker selection in Ket.

**Table 6.15. Sample paradigm of Ket Intransitive Configuration II**

<table>
<thead>
<tr>
<th>‘start growing’</th>
<th>‘turn red’</th>
<th>‘come into being’</th>
</tr>
</thead>
<tbody>
<tr>
<td>grow⁽⁶⁻<em>'-SBF⁽⁴⁻</em>'-TH⁽³⁻<em>'-PRES⁽⁴⁻</em>'-INCEP⁽³F*</td>
<td>1C⁽⁴⁻<em>'-RED⁽³⁻</em>'-SBF⁽⁴⁻<em>'-TH⁽³⁻</em>'-PRES⁽⁴⁻<em>'-become⁽³F</em></td>
<td>1C⁽⁴⁻<em>'-RED⁽³⁻</em>'-become⁽³F*</td>
</tr>
<tr>
<td>1SG</td>
<td><em>tiij⁽¹⁻</em>‘a⁽⁴⁻*’⁻<em>uar⁽³F</em></td>
<td><em>da⁽³⁻</em>⁻<em>súle⁽¹⁻</em>‘a⁽⁴⁻*’⁻<em>ka⁽³⁻</em>⁻<em>k⁽³⁻</em>⁻<em>sa⁽³F</em></td>
</tr>
<tr>
<td>2SG</td>
<td><em>tiij⁽¹⁻</em>‘a⁽⁴⁻*’⁻<em>uar⁽³F</em></td>
<td><em>da⁽³⁻</em>⁻<em>súle⁽¹⁻</em>‘a⁽⁴⁻*’⁻<em>ka⁽³⁻</em>⁻<em>k⁽³⁻</em>⁻<em>sa⁽³F</em></td>
</tr>
<tr>
<td>3M</td>
<td><em>di⁽¹⁻</em>‘a⁽⁴⁻*’⁻<em>uar⁽³F</em></td>
<td><em>da⁽³⁻</em>⁻<em>súle⁽¹⁻</em>‘a⁽⁴⁻*’⁻<em>ka⁽³⁻</em>⁻<em>k⁽³⁻</em>⁻<em>sa⁽³F</em></td>
</tr>
<tr>
<td>3N (SG, PL)</td>
<td><em>tiij⁽¹⁻</em>‘a⁽⁴⁻*’⁻<em>uar⁽³F</em></td>
<td><em>da⁽³⁻</em>⁻<em>súle⁽¹⁻</em>‘a⁽⁴⁻*’⁻<em>ka⁽³⁻</em>⁻<em>k⁽³⁻</em>⁻<em>sa⁽³F</em></td>
</tr>
<tr>
<td>1PL</td>
<td><em>tiij⁽¹⁻</em>‘a⁽⁴⁻*’⁻<em>uar⁽³F</em></td>
<td><em>da⁽³⁻</em>⁻<em>súle⁽¹⁻</em>‘a⁽⁴⁻*’⁻<em>ka⁽³⁻</em>⁻<em>k⁽³⁻</em>⁻<em>sa⁽³F</em></td>
</tr>
<tr>
<td>2PL</td>
<td><em>tiij⁽¹⁻</em>‘a⁽⁴⁻*’⁻<em>uar⁽³F</em></td>
<td><em>da⁽³⁻</em>⁻<em>súle⁽¹⁻</em>‘a⁽⁴⁻*’⁻<em>ka⁽³⁻</em>⁻<em>k⁽³⁻</em>⁻<em>sa⁽³F</em></td>
</tr>
<tr>
<td>3AN, PL</td>
<td><em>tiij⁽¹⁻</em>‘a⁽⁴⁻*’⁻<em>uar⁽³F</em></td>
<td><em>da⁽³⁻</em>⁻<em>súle⁽¹⁻</em>‘a⁽⁴⁻*’⁻<em>ka⁽³⁻</em>⁻<em>k⁽³⁻</em>⁻<em>sa⁽³F</em></td>
</tr>
</tbody>
</table>
Table 6.16. Sample paradigm of Ket Intransitive Configuration III

<table>
<thead>
<tr>
<th>Case</th>
<th>Stem</th>
<th>Subject</th>
<th>Object</th>
<th>Process</th>
<th>Resultative</th>
<th>Transitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>č₂č₃₄₅₆₇₈₉₁₀ (I whistle)</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
</tr>
<tr>
<td>1PL</td>
<td>č₂č₃₄₅₆₇₈₉₁₀ (we whistle)</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
</tr>
<tr>
<td>2SG</td>
<td>č₂č₃₄₅₆₇₈₉₁₀ (you whistle)</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
</tr>
<tr>
<td>3M</td>
<td>č₂č₃₄₅₆₇₈₉₁₀ (he/she/they whistle)</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
<td>č₂č₃₄₅₆₇₈₉₁₀</td>
<td>č₂ч₃₄₅₆₇₈₉₁₀</td>
</tr>
</tbody>
</table>

Resultsatives created by adding P₁ a to a basic transitive stem also belong to this type: **diluksjatet** ‘I break it’ ([d₃₄₅₆₇₈₉₁₀-č₂č₃₄₅₆₇₈₉₁₀-č₂č₃₄₅₆₇₈₉₁₀] → **diluksjatet** ‘it is broken’ ([d₃₄₅₆₇₈₉₁₀-č₂č₃₄₅₆₇₈₉₁₀-č₂č₃₄₅₆₇₈₉₁₀] → **diluksjatet** ‘it is broken’). Other verbs of this type have no transitive equivalent. Involuntary causatives containing P₂ follow this pattern: **daliukbutsteq** ‘she twitches’ [č₂č₃₄₅₆₇₈₉₁₀-č₂č₃₄₅₆₇₈₉₁₀-č₂č₃₄₅₆₇₈₉₁₀-č₂č₃₄₅₆₇₈₉₁₀]. Intransitive III motion verbs express unusually quick or intense actions: **daiğdabutsaq** ‘she makes a quick round trip to the river’ [č₂č₃₄₅₆₇₈₉₁₀-č₂č₃₄₅₆₇₈₉₁₀-č₂č₃₄₅₆₇₈₉₁₀]. Synonyms without this nuance follow the basic intransitive pattern: **daiğdakstsat** ‘she makes a round trip to the river lasting several days’ [č₂č₃₄₅₆₇₈₉₁₀-č₂č₃₄₅₆₇₈₉₁₀-č₂č₃₄₅₆₇₈₉₁₀-č₂č₃₄₅₆₇₈₉₁₀]. Most reciprocal verbs also follow this pattern: **thanatanutsadat** ‘they hug’ [č₂č₃₄₅₆₇₈₉₁₀-č₂č₃₄₅₆₇₈₉₁₀]. In all of its various meanings, Transitive III appears restricted to verbs that require a sentient subject.

The fourth intransitive pattern requires multiple-slot subject marking in P₈ and P₇ and can be used with inanimate-class as well as animate-class subjects. Unlike all other patterns that require P₈, the P-1 animate-class plural suffix does not appear, so that subject number is expressed only in P₇. This pattern productively detransitivizes Transitive I stems, as in the verb in Table 6.17, which means ‘to be made to start laughing.’ Other verbs of this pattern lack transitive
Table 6.17: Sample paradigm of Ket Intransitive Configuration IV

<table>
<thead>
<tr>
<th>sbf³-laugh³-caus³-pres³-sbj¹-iter.intr⁰</th>
<th>‘start laughing (iterative meaning)’</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>dafaq q₃ a₄ dgeq₁₀ (I start laughing)</td>
</tr>
<tr>
<td>2sg</td>
<td>kafaq q₃ a₄ xui₁⁰ yeq₁₀ (you.s start laughing)</td>
</tr>
<tr>
<td>3m</td>
<td>dafaq q₃ a₄ dgeq₁₀ (he starts laughing)</td>
</tr>
<tr>
<td>3f</td>
<td>da₃ dafaq q₃ a₄ dgeq₁₀ vaj₁⁰ (she starts laughing)</td>
</tr>
</tbody>
</table>
stem analogues. These include many basic verbs expressing the subject’s active change of state. Etymologically, the bases in such verbs denoted active movements, and the presence of the second subject marker in P7 marked the action as involving a single participant rather than a separate actor and undergoer: datájaq ‘she falls’ [da$^{-}$-aj$^{3}$-a$^{1}$-daq$^{3}$-sib$^{1}$-$\text{TH}^{-}$-pres$^{4}$-3sg.sib$^{1}$-fall$^{0}$], daajatij ‘she grows’ [da$^{3}$-aj$^{3}$-a$^{1}$-ti$^{0}$-3sib$^{1}$-$\text{TH}^{-}$-pres$^{4}$-3sg.sib$^{1}$-grow$^{0}$]. In a few cases, the morphology suggests movement of a specific body part and not the participant as a whole: diranqa ‘we barter, offer things for sale’ (etymologically, ‘extend our hands’) [di$^{-}$-day$^{1}$-qo$^{0}$-sib$^{1}$-$\text{PPL}$.sib$^{1}$-extend.hand$^{0}$], datasawo ‘she loads a gun’ (originally, ‘stretches a bowstring with her hand’) [da$^{3}$-ti$^{-}$-s$^{4}$-a$^{1}$-qo$^{0}$-3sib$^{1}$-bowstring$^{2}$-$\text{PRES}^{4}$-3sg.sib$^{1}$-stretch$^{0}$].

The fifth and last productive intransitive type includes verbs of having that incorporate their possessum in P7 and express their subjects in P4/1. A few intransitives from other semantic groups follow this agreement pattern, as well: irivet ‘I spend the day’ [i$^{3}$-di$^{3}$-bet$^{0}$-day$^{2}$-1sg.sib$^{1}$-have$^{0}$], sitkapi ‘you.pl wake up’ [si$t^{2}$-ka$^{1}$-a$^{0}$-awake$^{1}$-2pl.sib$^{1}$-process.occurs$^{0}$]. Similar to habeo-verbs, these stems logically require sentient subjects and therefore lack forms with inanimate-class subjects.

As might be expected in a language where the agreement-marking strategy is a lexical idiosyncrasy, a handful of Ket intransitives have rare or even unique agreement configurations. A few verbs mark the subject twice only in the plural forms: diraqagaq ‘we fly’ [di$^{-}$-day$^{1}$-doq$^{0}$-j$^{1}$-sib$^{1}$-$\text{PPL}$.sib$^{1}$-fly$^{0}$/pl.sib$^{0}$] (cf. diraq ‘I fly’ [di$^{3}$-(ji)-doq$^{0}$-sib$^{1}$-fly$^{0}$]). The past tense forms of one intransitive verb require subject marking in P6 and P1: tijbagbindo$^{0}$s I jumped up’ [e$^{3}$-ba$^{6}$-k$^{2}$-b$^{1}$-int$^{-}$-di$^{-}$-qo$^{0}$ up$^{1}$-1sg.sib$^{1}$-$\text{TH}^{-}$-intense$^{3}$-past$^{2}$-1sg.sib$^{1}$-jump$^{0}$].

The Ket system is thus much more complicated than that of Kott, though it is possible that additional features of the Kott system simply went unrecorded. Unlike Kott, it is not possible to describe a majority of Ket verbs as following any single type of agreement pattern.

6.3.3 A brief look at agreement in Arin and Pumpokol

The small number of verb forms attested for Arin and Pumpokol seem to represent agreement patterns with clear analogs to those of Ket and Kott. Pumpokol verbs

---

**Table 6.18. Sample paradigm of Ket Intransitive Configuration V**

<table>
<thead>
<tr>
<th></th>
<th>1sg</th>
<th>2pl</th>
<th>3sg</th>
</tr>
</thead>
<tbody>
<tr>
<td>tent$^{11}$-sbj$^{4}$-have$^{0}$ ‘own a tent’</td>
<td>qús$^{t}$-di$^{3}$-vet$^{t}$</td>
<td>qús$^{t}$-day$^{1}$-vet$^{t}$</td>
<td>qús$^{t}$-vet$^{t}$</td>
</tr>
<tr>
<td>1sg</td>
<td>qús$^{t}$-ka$^{1}$-vet$^{t}$</td>
<td>qús$^{t}$-day$^{1}$-vet$^{t}$</td>
<td>qús$^{t}$-ka$^{1}$-vet$^{t}$</td>
</tr>
<tr>
<td>3m</td>
<td>qús$^{t}$-aj$^{3}$-vet$^{t}$</td>
<td>qús$^{t}$-aj$^{3}$-vet$^{t}$</td>
<td>qús$^{t}$-aj$^{3}$-vet$^{t}$</td>
</tr>
<tr>
<td>3f</td>
<td>qús$^{t}$-j$^{1}$-vet$^{t}$</td>
<td>qús$^{t}$-j$^{1}$-vet$^{t}$</td>
<td>qús$^{t}$-j$^{1}$-vet$^{t}$</td>
</tr>
</tbody>
</table>
Losing semantic alignment

appear to use primarily subject person suffixes, like Kott and Assan: ćiin-du ‘he cries’ (du = 3m.sbj). Arin verbs show diverse agreement configurations. Although the number of attested items is too small to draw any firm conclusion, activity and action verbs in Arin seem to require subject prefixes: ani-p’a-man ‘I think’, ba-xatum ‘I go’ (p’a, ba = 1sg.sbj), while stative verbs take suffixes: at’ aj-tay ‘we stand’ (t’aj = 1pl.sbj).

The rise of new subject agreement markers as prefixes in Ket-Yugh, as suffixes in Kott-Assan and Pumpokol, and as both suffixes and prefixes in Arin, demonstrates that verb-internal agreement with active intransitive and transitive subjects is an innovation in the family overall.

6.4 Summary and conclusion

Proto-Yeniseic possessed a type of semantic alignment whereby subject and object NPs were zero marked and undergoer subjects generated verb-internal agreement. Modern Ket, the family’s only surviving member, exhibits a more complex system of verb-internal subject/object marking involving five productive intransitive patterns, two productive transitive patterns, and several unproductive patterns. The choice of agreement-marking strategy is associated with an array of morphological and semantic triggers that defy classification on the basis of generalized semantic categories such as stative or agentive. Using comparative data from extinct Yeniseic languages, this chapter has traced the development of this unusual morphological phenomenon, which no longer can be described as representative of semantic alignment. A comparison of verb morphology across Yeniseic suggests the model of finite verb formation for Common Yeniseic shown in Table 6.19. Several observations can be made regarding the evolution of agreement marking from Common Yeniseic to Modern Ket and other daughter languages. First, the suffix position expressing agreement with a 1st or 2nd person undergoer subject became an object slot in Ket transitive stems which include a causative marker or animacy classifier in position 4, but remained the primary subject-marking slot only in a subgroup of inactive verbs whose semantics limit them to animate subjects (i.e., ‘have’, ‘wake up’, ‘feel sleepy’). Second, the animacy classifiers—d for animate-class undergoers, b for inanimate-class—retained their original functions in certain Kott agreement models. In Ket, however, the animacy classifier became associated with objects

Table 6.19. Position classes in Late Common Yeniseic

<table>
<thead>
<tr>
<th>P6</th>
<th>P5</th>
<th>P4</th>
<th>P3</th>
<th>P2</th>
<th>P1</th>
<th>Po</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incorporate</td>
<td>Object or subject</td>
<td>Thematic consonant</td>
<td>Animacy classifier (d or b)</td>
<td>Tense/mood marker</td>
<td>Undergoer subject (1st, 2nd person)</td>
<td>Verb base</td>
</tr>
</tbody>
</table>
rather than subjects, and dropped out of the productive system of verb stem creation. Third, position 6, which was the original location for objects in the verb phrase, fused with the verb complex as a prefix and expanded its subject-marking functions, which were originally limited to impersonal constructions. This slot therefore came to express subject marking in a variety of patterns, including involuntary causatives and intransitive inceptives that fill the left base with an infinitive form. Finally, each daughter language acquired a new active or transitive subject-marking slot.

Ket and Yugh developed a new person agreement prefix in the leftmost position, found in all transitive verbs and many intransitives as well. Kott and Pumpokol subject agreement came to be expressed as a suffix in most verbs. In Arin, certain verbs require subject prefixes, while others require suffixes, the number of recorded forms being insufficient to determine the conditioning behind this pattern. In Ket and Yugh, retention of the original subject agreement prefix directly before the base in certain stems resulted in a new pattern of multi-site subject marking. A second type of multi-site subject marking developed when the old object slot came to be used as a second subject marker to convey such ideas as quicker than usual action or action performed by the subject’s own body rather than by an external instrument (e.g. ‘carry in one’s arms’ rather than ‘drag by sled’).

In contrast to the rather straightforward morphology of Proto-Yeniseic, verb agreement in the daughter languages displays a much more complex pattern of lexical conditioning. This complication was generated by the incorporation into the morphological verb of the following positions: active or transitive subject, a new lexical position referred to here as the ‘left base’, and an object slot with accompanying postposition. The elements in all of these positions originally represented prosodically independent elements. Their incorporation into the verb also led to the rise of polypersonal agreement, including gender agreement for 3rd person terms. In terms of its non-agreement categories, the verb remained rather simple, conveying only tense (past vs. non-past), and mood (indicative vs. imperative). The internal elaboration of the Modern Yeniseic verb thus derives primarily from the rise of lexically competing subject and object affix positions as the Yeniseic daughter languages gradually accommodated themselves from a prefixing to a suffixing morpheme arrangement. In contrast to Proto-Yeniseic, where all verbs consisted of a root preceded by prefixes, verb stems in the daughter languages can be either right- or left-headed. In Modern Ket, right-headed patterns are archaic, though they remain productive in cases of patient or instrument incorporation. The remaining productive stem types in Modern Ket place the semantic head (an infinitive form) in the leftmost lexical slot, so that all of the original affix positions have become de facto suffixes. This trend appears to have developed under the influence of the suffixing agglutinative languages surrounding the Yeniseic family in Siberia, such as Turkic and Samoyedic, which require the root to be placed verb-initially.
Losing semantic alignment

To summarize, the attested Yeniseic languages show a transition from a monopersonal active/stative role-marking system involving valence-change prefixes, animacy classifiers, and agreement with speech-act-participant undergoer subjects, to a transitivity-based polypersonal agreement system where the choice of subject/object marker configuration serves as a stem-building device. This vividly demonstrates that agreement typology can change not only because of a shift in grammatical alignment but also because of more idiosyncratic morphological factors having nothing to do with alignment, in and of themselves. Yeniseic did not replace its original semantic alignment with either nominative or ergative alignment. Instead, the individual daughter languages developed a series of morphological complications as they shifted their semantic head from the right to the left. In connection with this shift, agreement strategy itself became a lexical idiosyncrasy of each verb stem. Some agreement marker configurations retain direct traces of the original system of semantic alignment; others do not. This has resulted in a bewildering array of patterns that must be regarded as lexical features rather than dictated by the language’s overall alignment type. This morphological feature renders Modern Ket highly unusual among languages with polypersonal agreement systems.
Intransitive split in Tundra Nenets, or how much semantics can hide behind syntactic alignment

OLESYA KHANINA

7.1 Introduction*

7.1.1 Language background

Tundra Nenets, a North-Samoyedic language of the Uralic language family, is reported to be spoken by c.25,000 people along a vast tundra zone of Arctic Russia, from the Kanin Peninsula in the west to the Yenisei River in the east.

Unfortunately, most Tundra Nenets speakers today tend not to use the language in everyday communication nor to transmit it to their children; the sociolinguistic situation is only slightly better among actual reindeer herders living in the tundra. However, Tundra Nenets is still one of the less endangered indigenous languages of Russian Asia (Sulyandziga et al. 2003). Originally, the Tundra Nenets were nomadic reindeer herders. In 1950–60, a forced transition to a settled way of life and to Soviet everyday culture began, weakening ties to the cultural and linguistic traditions.

In its overall typology, Tundra Nenets is a typical Uralic language with accusative alignment, double marking (both head- and dependent-marking) in the clause and in possessive NPs, and extensive use of non-finite verb forms for subordination. Like other Samoyedic languages, Tundra Nenets has rather complex morphophonology in combination with an otherwise agglutinative structure.

* I would like to thank my colleagues Anna Pazelskaya and Andrey Shluinsky for their valuable comments on an earlier version of this chapter. I am also grateful to Edward Vajda for correcting some of my mistakes in English and to Peter Staroverov for his help in verifying spelling of Tundra Nenets examples. Last but not least, I am indebted to the editors of this volume, who first cordially invited me to participate in this collection and then provided a number of useful comments on the first version of this chapter.

1 For example, its closest living relatives, the Samoyedic languages Selkup, Nganasan, and Enets, have considerably fewer speakers: not more than 1,500, 200, and 70 people, respectively (Sulyandziga et al. 2003, Daniel 2005, Khanina and Shluinsky 2005, 2006, to appear).
Major grammatical descriptions of Tundra Nenets are Castrén (1854/1966), Tereschenko (1956, 1973), and Salminen (1997); see also Körtvély (2005). Two extensive dictionaries have been published, Tereschenko (1965) and Salminen (1998).

7.1.2 Scope and aims of this study

This chapter presents new information on the intransitive split in Tundra Nenets, i.e. on the existence of two fully-fledged intransitive paradigms, the so-called 'A paradigm' (1a, 2a) and 'B paradigm' (1b, 2b).

(1) a. ja p'i-nā-Ø.
   flour.soup get.cooked-IPFV.GF-3SGA
   'The flour soup is being cooked (i.e. has been already boiling for some time).'

   b. ja p'i-ni-ʔ.
   flour.soup get.cooked-IPFV.SF-3SGB
   'The flour soup is starting to be cooked (i.e. only starting to boil).'

(2) a. wasja peda-Ø.
   Wasja get.tired.GF-3SGA
   'Wasja has got tired.'
   (definitely tired: after some long-lasting hard work; cannot continue working)

   b. wasja pede-j-ʔ.
   Wasja get.tired.SF-SF-3SGB
   'Wasja has got tired.'
   (somewhat tired, but able to continue working)

Tundra Nenets intransitive verbs can be divided into three lexical classes according to their compatibility with these paradigms: A-only verbs, B-only verbs and A & B verbs (= ‘fluid’ verbs, such as p’i/nā ‘to be cooked’ and peda ‘to get tired’ in (1)–(3)).

The main object of this study is to investigate the semantic principles that govern a verb’s placement in one of the classes as well as the semantic principles governing the choice between the two paradigms in the case of fluid verbs. I claim that these principles have much in common with semantic principles of the semantic alignment as discussed in other chapters of this volume.

Previous research on Tundra Nenets has paid little attention to grammatical semantics; in particular, none of the contributions cited above has discussed the logic of the intransitive split. Moreover, the existence of fluid verbs has been

7.1.3 Structure of the chapter
The rest of the introduction is devoted to the exposition of methodology and data used for this study. Section 7.2 presents a general overview of Tundra Nenets verbal morphology; this auxiliary section describes the formal mechanisms by which the intransitive split is realized. The semantic principles behind the split are discussed in section 7.3. Section 7.4 summarizes the findings of the study and discusses them within the purview of semantically aligned systems.

7.1.4 Methodology and data
This study is based on the results of my fieldwork in the context of Moscow State University Expeditions to a western dialect of Tundra Nenets (the Malozemel’ski dialect spoken in Neneckij Avtonomnyj Okrug, particularly in the village of Nelmin Nos).5

The field trips were devoted principally to a study of aspectual and valence derivations in Tundra Nenets. The main tool for this study was a database initially comprising 110 undervived verbs belonging to the core verbal lexicon. Further, various aspectual and valence-changing derivations from the core verbs were added, such that at the end of our fieldwork the database contained more than 2,500 verbs (the biggest depth of derivation was six consecutive derivations; however, the average depth was only two or three consecutive derivations). For each verb, its possible paradigms (transitive, intransitive A, and intransitive B) were elicited in simple sentences; the telicity of each verbal form was checked by its compatibility with time adverbs ‘in X time’ (telic), ‘for X time’ (atelic); changes in the verb’s meaning depending on a given paradigm were also recorded. The following productive derivations were used for the enlargement of the original set of 110 core verbs: Causatives -tə-, -t̪e-, -ra-, -rə, -pət̪a-, Durative -(m)pə-, Imperfective -nə-, Frequentative -r-, Iterative -tgə-, and Inchoative -l-.

5 Interestingly, verbs which in fact allow both intransitive paradigms are occasionally encountered in both Tundra Nenets dictionaries (e.g. Tereschenko 1965: 55, 531, 759, 776, Salminen 1998: 91, 130, 319, 357). Comments in Tereschenko (1956: 111–12) suggest that when the researcher was faced with intransitive verbs having both intransitive paradigms, the two intransitive usages were considered to be two totally different lexemes having the same phonological form by pure coincidence.

5 The fieldwork was funded by the Russian State Foundation for Humanities (grant no. 03-04-18005e); the field trips were carried out during seven weeks in the summer of 2003, 2004, and 2005. I would like to express my gratitude to Sergey Tatevosov and Ekaterina Lysutikova, who organized and supervised the whole project. I am also very much indebted to the following native speakers of Tundra Nenets for their constant collaboration and invaluable help: Mihail Kiprijanovich Tajbarej, Evgenija Nikolaevna Tajbarej, Mariya Ivanovna Kanyukova, Klara Petrovna Taleeva, Anna Grigorjevna, and Valentina Pavlovna Maryueva.
I collected and analyzed the possible paradigms of the database verbs, and this chapter presents some results of this study. I collected and analyzed the possible paradigms of the database verbs, and this chapter presents some results of this study. Some data on derivations collected by my colleagues were used as well. The overall data used as a background for this study is a sample of more than 700 verbs (a random sample from the 2,500 verbs of the aforementioned database, no Inchoatives included); 282 of these can be used only in intransitive predication.

The present study of the intransitive split in Tundra Nenets, then, is based mainly on a corpus of elicited simple sentences, suggested by native speakers as a reaction to presentation of a verb form. Most of the sentences, though not all, were approved by more than one native speaker. There is a complementary corpus of ungrammatical sentences, containing the verb forms that were rejected by at least one and usually more than one native speaker. Such evidence is far from ideal, given the unnatural context of production, but only this type of procedure allows full information on possible and impossible paradigms of each verb chosen for the study. Moreover, the procedure brought out the semantic differences between various paradigms of the same verbal lexeme.

We should note that the data on verb class membership was obtained via elicitation, in a precarious sociolinguistic situation (see section 7.1.1). This is probably the source of the considerable variation found here, including contradictions from one speaker to another. The tendencies were clear, however: while variation in the possible forms of some individual verbs is possible, all of the verb classes were established without doubt.

### 7.2 Overview of Tundra Nenets verbal morphology

Before going into detail, I will present the general structure of Tundra Nenets verbs. The structure of finite verbs is shown in (3), that of non-finite verbs in (4). There are two obligatory formatives determining the individual paradigms.

---

6 The data was collected by me and Sergey Tatevosov. I am grateful to him for the work done—alone I would never have been able to collect such a volume of data. He also helped me with some insightful comments during our endless discussions of the matter. However, there are a number of substantial differences in our analyses of the Tundra Nenets system, and I am solely responsible for the analysis formulated here.

7 I would like to thank Denis Ivanov, Mihail Ivanov, Ekaterina Lyutikova, Anna Pazelskaya, Andrey Shluinsky, and Sergey Tatevosov, who collected and analyzed most of the data on verbal derivations. My thanks also go to those who helped them, namely, MSU students Pavel Iosad, Alena Khanina, Maria Tsiurupa, Ekaterina Volovich, Nikolay Vorontsov, and Natalia Zevakhina.

8 At least in the dialect studied in the fieldwork, Inchoatives may involve three alternative general formatives and three possible special formatives (see section 7.2.2), so the number of paradigms they have is 12, rather than the usual 4 (semantic differences were attested among most of the 12 paradigms). That is why Inchoatives would require a separate, much more detailed study. Some other derivations may also choose from an array of alternative general and special formatives, but the semantic differences between the paradigms set by these formatives did not prove consistent for any derivation, except Inchoative.
of finite verbs: the so-called ‘inflectional stem’ formative and the cross-reference formative. The latter will be discussed first, as it is the less controversial one.

(3) ROOT + (aspect/valence) + (mood) + ‘inflectional stem’ + (dual object) + cross-reference + (preterite post-fix)

(4) ROOT + (aspect/valence) + participle/infinitive/… + (case) + (possessive (=transitive cross-reference))

7.2.1 Person–number cross-reference formatives

There are four cross-reference paradigms in Tundra Nenets: two intransitive paradigms, A and B, and two transitive paradigms, for singular object and for non-singular object (see Tables 7.1–7.4).

Table 7.1. Subject markers for the Intransitive A paradigm (glossed ‘A’)

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>DU</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-m(ā)˘/d(ā)m</td>
<td>-n’i?</td>
<td>-wa?</td>
</tr>
<tr>
<td>2</td>
<td>-n</td>
<td>-d’i?</td>
<td>-da?</td>
</tr>
<tr>
<td>3</td>
<td>-Ø</td>
<td>-x(ā)?</td>
<td>-?</td>
</tr>
</tbody>
</table>

Table 7.2. Subject markers for the Intransitive B paradigm (‘B’)

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>DU</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-w(ā)?</td>
<td>-n’i?</td>
<td>-na?</td>
</tr>
<tr>
<td>2</td>
<td>-n</td>
<td>-d’i?</td>
<td>-da?</td>
</tr>
<tr>
<td>3</td>
<td>-?</td>
<td>-x(ā)?</td>
<td>-d(ā)?</td>
</tr>
</tbody>
</table>

Table 7.3. Subject markers for the Transitive paradigm, singular object (trs)

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>DU</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-w</td>
<td>-m’i?</td>
<td>-wa?</td>
</tr>
<tr>
<td>2</td>
<td>-r?</td>
<td>-r’i?</td>
<td>-ra?</td>
</tr>
<tr>
<td>3</td>
<td>-da</td>
<td>-d’i?</td>
<td>-da?</td>
</tr>
</tbody>
</table>
As can be seen from Tables 7.1 and 7.2, the subject of an intransitive clause is cross-referenced on the predicate by person–number affixes which distinguish three numbers (singular, dual, and plural) and three persons (1st, 2nd, and 3rd). The verb in (5) provides an example of intransitive A paradigm for all persons in the singular.

(5) a. (mân’) màjäm-la-da-m?.
   (I) become_glad-INCH-TRR.GF-1SGA
   ‘I have become glad.’

b. (pidär) màjäm-la-da-n.
   (you) become_glad-INCH-TRR.GF-2SGA
   ‘You have become glad.’

c. was’a màjäm-la-da-O.
   Wasja become_glad-INCH-TRR.GF-3SGA
   ‘Wasja has become glad.’

The verb in (6) shows an intransitive B paradigm for all persons in the singular.

(6) a. (mân’) màjäm-la-de-j-w?.
   (I) become_glad-INCH-TRR.SF-SF-1SGB
   ‘I have become glad.’

b. (pidär) màjäm-la-de-j-n.
   (you) become_glad-INCH-TRR.SF-SF-2SGB
   ‘You have become glad.’

c. was’a màjäm-la-de-j-?.
   Wasja become_glad-INCH-TRR.SF-SF-3SGB
   ‘Wasja has become glad.’

As can be seen from Table 7.3 and Table 7.4, both the subject and the object of transitive clauses are cross-referenced on the predicate by person–number affixes. These cumulative affixes distinguish three numbers (singular, dual, and plural) and three persons (1st, 2nd, and 3rd) for subject and only two numbers (singular and non-singular) for object. The opposition between dual and plural number

<table>
<thead>
<tr>
<th></th>
<th>SG</th>
<th>DU</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-n</td>
<td>-nɨʔ</td>
<td>-naʔ</td>
</tr>
<tr>
<td>2</td>
<td>-d</td>
<td>-dɨʔ</td>
<td>-dɑʔ</td>
</tr>
<tr>
<td>3</td>
<td>-da</td>
<td>-dɨʔ</td>
<td>-dɑʔ</td>
</tr>
</tbody>
</table>

Table 7.4. Subject markers for the Transitive paradigm, non-singular object (trns)
of the object is realized by presence vs. absence of dual suffix -xäyu-, which goes before the cross-reference affix.

The verb in (7) gives an example of a transitive paradigm for 1st person singular subject.

(7) a. (män¹) wen’eko-m t’or’e-w.
   (I) dog-ACC cry.GF-1SGTRS
      ‘I called the dog.’

b. (män¹) wen’eko-xo? t’or’e-ya-xäyu-n.
   (I) dog-DU.ACC cry-GF-DU-1SGTRNS
      ‘I called two dogs.’

c. (män¹) wen’eku t’or’e-j-n.
   (I) dog.PL.ACC cry-SF-1SGTRNS
      ‘I called the dogs.’

Note also that ‘differential object marking’ (see e.g. Comrie 1979, 1986, Croft 1988, Lazard 1984, Silverstein 1981, Aissen 2003) is attested in Tundra Nenets: when the object of a transitive clause has specific referential properties (usually low; see Körtvély 2005 for more details), the predicate takes the intransitive A paradigm instead of the transitive paradigm. In other words, depending on their referential properties, objects may or may not be cross-referenced in the verb.

Sentences (8a) and (9a) are taken out of the contexts in which the objects mentioned (respectively ‘book’ and ‘knife’) were introduced. Therefore the verbs take the transitive paradigm. In contrast, the object in (8b) is newly introduced and the object in (9b) is generic, so both verbs in these cases take the intransitive A paradigm.

(8) a. kn’iga-m was’a me-da /³me-Ø.
   book-ACC Wasja take.GF-3SGTRS/take.GF-3SGA
      ‘Wasja took the book.’

b. was’a kn’iga-m me-Ø /³me-da /³me-j.?.
   Wasja book-ACC take.GF-3SGA/take.GF-3SGTRS/take-SF-3SGB
   {What did Wasja take?-} ‘Wasja took a book.’
Intransitive split in Tundra Nenets

(9) a. män'i xär-m juxu-pta-w.
   I knife-ACC get.lost-caus.GF-1SG.TRS
   ‘I have lost the knife.’

b. was'a juxu-pta-Ø.
   Wasja get.lost-caus.GF-3SGA
   ‘Wasja (always) loses everything {∼that is in his character}.’

The use of the intransitive A—but not B—paradigm for this type of transitive clause can be explained historically. This kind of ‘differential object marking’ is common for most Uralic languages, and all of them have the intransitive A paradigm, which presumably originates from Proto-Uralic (see Körtvély 2005). The intransitive B paradigm, however, is a recent innovation in North Samoyedic languages,9 and thus did not take over the function already fixed for the other intransitive paradigm.

Summing up, the examples cited above in conjunction with Tables 7.1–7.4 show that there are four distinct sets of cross-reference affixes, none of which can be treated as a subtype of another (even though some endings can in fact be found in more than one paradigm).10

7.2.2 Inflectional formatives

The verbal person–number paradigms are specified not only by cross-reference suffixes but also by the already mentioned ‘inflectional stem’ formative (glossed in the examples as ‘GF’/’sf’). The term was used in Salminen (1997) and Salminen (1998), where the element in question was claimed to be a part of the inflectional stem. As I show below, this formative is better described as contributing exclusively to inflection, without taking part in stem formation. Thus, I call it ‘inflectional formative’.

Figures 3 and 4 above make it clear, first, that inflectional formatives—together with mood markers and cross-reference affixes—only appear in finite verbs and never in non-finite verbs, and secondly, that inflectional formatives always appear next to cross-reference affixes. Inflectional formatives may or may not appear next to the lexical stem depending on the presence or absence of intervening mood markers, as in (10).

9 The innovation concerns the paradigm as a whole: certain person–number forms are found in other Uralic languages, but the whole paradigm is not attested anywhere except in North Samoyedic languages (see Körtvély 2005).

10 A different analysis of intransitive paradigms was suggested in Salminen (1997: 103–4). The A and B paradigms were claimed to represent an ‘almost’ single paradigm. Most differences between ‘surface’ A and B paradigms were attributed to an underlying suffix -t- (realized as -t- before consonants and at the end of words), which appears in paradigm B, but not in paradigm A. This hypothesis, however, is implausible, at least from a synchronic point of view: even if we accept the general rule of -t- insertion, 2 of altogether 9 forms remain unexplained (1SG, 3PL), and the general rule itself follows a logic which is far from being natural, since it only applies to the 1SG, 3SG and 3PL cells of the paradigm.
Table 7.5. Inflectional formatives and sets of cross-reference suffixes

<table>
<thead>
<tr>
<th>Intransitive A</th>
<th>General formative (GF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transitive singular object</td>
<td>GF</td>
</tr>
<tr>
<td>Transitive non-singular object: dual</td>
<td>GF</td>
</tr>
<tr>
<td>Transitive non-singular object: plural</td>
<td>Special formative (SF)</td>
</tr>
<tr>
<td>Intransitive B</td>
<td>SF</td>
</tr>
</tbody>
</table>

(10) a. xada-na-r’ xa-da.
kill-IPFV-APPR.GF-3SGTRS
'It seems to kill it.' (Salminen 1997: 111)\(^{11}\)

b. xada-na-r’xi-da.
kill-IPFV-APPR.SF-3SGTRNS
'It seems to kill them.' (ibid.)

In Salminen (1997: 51–6) it was clearly demonstrated that mood markers belong semantically and structurally to inflection and not derivation. Thus, an inflectional formative cannot be analysed as an element modifying the lexical stem, because its position is between two clearly inflectional elements: a mood affix and a cross-reference affix.

An additional argument for this ‘non-stem-related’ status of inflectional formatives is their interconnectedness with cross-reference affixes. On the one hand, the linear position of inflectional formatives cannot be formulated without reference to cross-reference suffixes (inflectional formatives always immediately precede such suffixes). On the other hand, the structural impact of the formatives on a verbal word cannot be seen as separate from the structural impact of cross-reference suffixes. Indeed, all four paradigms listed in Tables 7.1–7.4 are specified only by combination of inflectional formatives and cross-reference suffixes, as shown in Table 7.5.\(^{12}\) Note that there are only two inflectional formatives, and neither is exclusively associated with an individual cross-reference paradigm.

The examples cited above illustrated three possible types of inflectional formative:

- either the formative represents a phonological unit on its own, as in (7b, c);
- or its presence can be understood only by a phonological change of the segment to which the formative is attached (i.e. a mood affix, as in (10b) or a lexical stem, as in (1));

---

\(^{11}\) When citing other sources I retain the original orthography.

\(^{12}\) In Salminen (1997, 1998) these formatives were called markers of ‘general finite stem’ and ‘special finite stem’, respectively.
Intransitive split in Tundra Nenets

- or the formative represents a phonological unit on its own and at the same time induces a phonological change of the segment the formative is attached to, as in (2) and (6).

The choice of the type of inflectional formative for each individual verb depends on the morphophonology of the verb. Concrete algorithms are too complex to cite here (see Salminen 1997: 99–103, 108–12 for details); however, it is important to note that for any verbal form the choice of inflectional formative is unambiguous and can be predicted.

7.2.3 Transitivity

Many Tundra Nenets verbs are ambitransitive, i.e. they can be used to describe a transitive situation as well as an intransitive one. Consider the following pairs: intransitive A/B ((5) and (6)) vs. transitive (11), and intransitive B (12) vs. transitive (9).

(11) \( (m\text{"am}) \) was\'a-m m\text{"aj}\text{m}-la-da-w.
   (I) Wasja-ACC become.glad-INCH-TRR.GF-1SGTRS
   'I have caused Wasja to become glad.'

(12) \( x\text{"ar} \) juxu-pte-j-?.
   knife get.lost-CAUS.SF-SF-3SGB
   'The knife has got lost.'

In the case of ambitransitive verbs, all combinations of paradigms are attested, i.e. there are:

- ambitransitive verbs which use transitive and intransitive A paradigms;
- ambitransitive verbs which use transitive and intransitive B paradigms;
- ambitransitive verbs which use transitive and both intransitive paradigms.

While the mechanisms of choice between one of the intransitive paradigms, A or B, in the case of ambitransitive verbs are also worth studying, they are outside the scope of this study, which deals only with pure intransitive verbs and not ambitransitive verbs.

7.2.4 Derivational mechanisms

As mentioned in section 7.1.4, the following productive derivations are attested in Tundra Nenets: Causative -pta- (9, 12), Transitive -ta-, -t\text{e}-, -ra-, -r\text{e} (5, 6, 11), Durative -(m)p\text{"a}-, Imperfective -n\text{\text{"a}}- (1, 10), Frequentative -r-, Iterative -g\text{"a}-, and Inchoative -l- (5, 6, 11). The names of the derivations are taken from Salminen (1997, 1998). Except for the more or less ‘classic’ causative and inchoative, these labels are not necessarily meant to be interpreted in a standard typological fashion.
 aspectual and valence derivations must be mentioned in connection with the semantics of intransitive split because they influence verbal semantics substantially and thus participate in the paradigm choice. As can be seen from Appendix 7.1, there are some paradigmatic semantic patterns that appear only in the case of a specific derivation.

7.3 Semantic principles governing the intransitive split in Tundra Nenets

Now that I have introduced the formal mechanisms relating to the two intransitive paradigms in Tundra Nenets, it is time to turn to the question of their distribution.

7.3.1 Evidence from non-finite clauses

We have already seen that the two intransitive paradigms in Tundra Nenets cannot be analysed as having been derived by some formal means from one another. Semantically it is equally impossible to treat one intransitive paradigm as a derivation from the other one.

The key argument here is the semantics of non-finite verbal forms: for a given verbal lexeme, its non-finite forms have as many interpretations as are attested when this lexeme is used finitely. If a lexeme can be used with two intransitive inflectional paradigms and thus have two different interpretations, both interpretations are equally available when this lexeme is used in a non-finite form, i.e. when no inflectional paradigm is chosen. Cf. (13a, b) vs. (13c), (14a, b) vs. (14c), and (15a, b) vs. (15c).

13 In this notation, first appears actional interpretation of the verbal form in the classical terms of Vendler (1957/1967), and second, actional interpretation in the terms of Tatevosov (2002); see section 7.3.2 for details.
Intransitive split in Tundra Nenets

(14) a. sər’ o xāwna-Ø.
   rain be.raining.GF-3SGA
   ‘It is raining.’ <activity/process>

b. sər’ o xəwni-ʔ.
   rain be.raining.sf-3SGB
   ‘It has started to rain.’ <achievement/entry into a process>

c. sər’ o xāwna-ʔma  məliʃəna  mən1 xoni-mâʔ-si.
   [rain be.raining-INF.PF when/during] I sleep-1SGA-PST
   ‘While it was raining, I was sleeping.’ <activity/process>
   ‘When it started raining, I was sleeping’ <achievement/entry into a process>

(15) a. wasə tu-r-ṣa-Ø.
   Wasja come-fr-GF-3SGA
   ‘Wasja comes (regularly).’ <activity/process>

b. wasə tu-r-ʃ-ʔ.
   Wasja come-fr-sf-3SGB
   ‘Wasja has started to come (regularly).’ <achievement/entry into a process>

c. tu-r-ʃ1 tara.
   [come-fr-GER.MOD] necessary
   ‘It is necessary to come (regularly).’ <activity/process>
   ‘It is necessary to start coming (regularly).’ <achievement/entry into a process>

Semantic equivalence between situations rendered by finite and non-finite forms of the same verb suggest that different interpretations observed within different paradigms of the same verb do not originate in the paradigms but in verbal lexical semantics. It allows me to conclude that intransitive paradigms only allow for realizing interpretations contained in the verbal lexical stem, i.e. in the root, with or without aspectual/valence derivations.

7.3.2 Differentiating between two intransitive paradigms

One and the same verb can have different actional (Aktionsart) interpretations when used in different paradigms: cf. the possibilities already exemplified above in (13a) vs. (13b), (14a) vs. (14b), and (15a) vs. (15b). Table 7.6 enumerates the actional interpretations that are available for each intransitive paradigm.

The classification of actional types used here is based on Tatevosov (2002). I have further added two actional types specific to Tundra Nenets: ‘entry into the first stative phase’ for situations consisting of two phases both of which are states and ‘first processive phase’ for situations consisting of two phases both of
<table>
<thead>
<tr>
<th>Actional type</th>
<th>A paradigm</th>
<th>B paradigm</th>
<th>A and B paradigms</th>
<th>A and B paradigms</th>
</tr>
</thead>
<tbody>
<tr>
<td>Telicity</td>
<td>atelic</td>
<td>telic</td>
<td>telic</td>
<td>atelic</td>
</tr>
<tr>
<td>No. of verbs</td>
<td>208</td>
<td>14</td>
<td>14</td>
<td>14</td>
</tr>
<tr>
<td>Examples</td>
<td>(1a), (4a), (5a), (13a), (16)</td>
<td>(1a), (3a), (3b), (18a), (19)</td>
<td>(1a), (4a), (5a), (13a), (16)</td>
<td>(1a), (3a), (3b), (18a), (19)</td>
</tr>
<tr>
<td>Process</td>
<td>telic</td>
<td>atelic</td>
<td>telic</td>
<td>atelic</td>
</tr>
<tr>
<td>Stative phase</td>
<td>73</td>
<td>73</td>
<td>73</td>
<td>73</td>
</tr>
<tr>
<td>Telic phase</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Telic phase</td>
<td>73</td>
<td>73</td>
<td>73</td>
<td>73</td>
</tr>
</tbody>
</table>

Table 7.6. Actional interpretations available for two intransitive paradigms.
which are processes. \textsuperscript{14} ‘Process’ here is understood in exactly the same way as ‘activity’in Vendler (1957, 1967). ‘Entry into a state’ and ‘entry into a process’ both correspond to the union of Vendler’s ‘accomplishment’ and ‘achievement’: such union is often treated together; see, for example, ‘telic processes’ in Chung and Timberlake (1985), ‘changes’ in Dik (1989, 1994), ‘events’ in Mourelatos (1978/1981) and Bach (1986), or ‘transformative’ in Johanson (1996).\textsuperscript{15} Tatevosov (2002: 330–31), however, suggested differentiating between ‘entry into a state’ and ‘entry into a process’: ‘Entry into a state is a label assigned whenever a situation referred to has attained its resultant state . . . [T]he Entry into a process . . . meaning label is assigned if a sentence refers to a situation that terminates producing a new situation.’

Returning to the specific actional interpretations available to Tundra Nenets intransitive verbs, it can easily be noted from Table 7.6 that—except for the problematic ‘entry into a state’, which will be discussed below—all other actional interpretations are expressed by one alternative paradigm only. Moreover, the main principle distinguishing the A-actional interpretations from B-actional interpretations is the well-known opposition ‘atelic vs. telic’.

However, in Tundra Nenets this principle is further refined by a more specific principle. It can be formulated as ‘homogeneity of the situation’, i.e. when any subpart of the situation can be counted for the entire situation.

\textit{Homogeneity of the situation}

While process and state are homogeneous, entry into a state and entry into a process are not. As for the 1st processive phase, it is necessary to know the whole situation, i.e. to have the information about the other subsequent phase taking place, to speak about the 1st phase. If there is no such information, it is actually impossible to speak about the 1st processive phase and the simple process has to be admitted. Therefore the situation ‘1st processive phase’ is not homogeneous in the sense of the term as it is used here.\textsuperscript{16}

To sum up, the A paradigm is used for homogeneous atelic situations, the B paradigm is used for non-homogeneous situations (atelic or telic).\textsuperscript{17}

For verbs with one paradigm only, the choice of the paradigms is thus fairly plain: if a verb describes a homogeneous situation, it has the A paradigm.

\textsuperscript{14} For reasons of simplification I do not treat separately the actional type ‘multiplicative process’ of Tatevosov (2002): it was attested only in a small number of cases, which I have excluded from this study.

\textsuperscript{15} The bibliographical information is from Tatevosov (2002: 320–21).

\textsuperscript{16} I am using the term ‘homogeneity’ in an intuitive, i.e. non-formalized, way; the term may thus invite further specification in a formal aspectual framework (see for discussion of similar problems, among many others, Bennett and Partee 1978, Taylor 1977, Filip 1999). The present study is purely descriptive and aims only to provide evidence for such specification, not to formulate the specification itself.

\textsuperscript{17} While a ‘non-homogeneous’ situation could be called ‘heterogeneous’, I prefer to use the more transparent label ‘non-homogeneous’, since it clearly indicates the binary nature of the opposition.
Table 7.7. Possible pairs of actional types manifested in different paradigms of fluid verbs

<table>
<thead>
<tr>
<th>Actional interpretation in B paradigm</th>
<th>Actional interpretation in A paradigm</th>
<th>No. of fluid verbs with this pair of actional interpretations</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Entry into a process</td>
<td>Process</td>
<td>69 (11 underived)</td>
<td>(14a)-(14b), (15a)-(15b)</td>
</tr>
<tr>
<td>(b) First processive phase</td>
<td>Process</td>
<td>65 (1 underived)</td>
<td>(1a)-(1b)</td>
</tr>
<tr>
<td>(c) Entry into a state</td>
<td>State</td>
<td>8 (4 underived)</td>
<td>(13a)-(13b)</td>
</tr>
<tr>
<td>(d) Entry into first stative phase</td>
<td>Entry into a state</td>
<td>4 (all underived)</td>
<td>(2a)-(2b)</td>
</tr>
<tr>
<td>(e) Entry into a process</td>
<td>Entry into a state</td>
<td>3 (all underived)</td>
<td>(18a)-(18b)</td>
</tr>
</tbody>
</table>

(ex. (16)); if a verb describes a non-homogeneous situation, it has the B paradigm (ex. (17)).

(16) *yuda-w n’o-xana latra-1gái-Ø */latra-1gái-Ø.  
as-1sg.sg door-loc squeeze-iter.GF-3SGA/squeeze-iter.SF-3SGB  
‘I (regularly) squeeze my arm in the door.’

(17) a. *pa tum-j-Ø */tum-a-Ø.  
wood take.fire-GF-3SGB / take.fire-GF-3SGA  
‘The wood has taken fire.’

But what is going on with fluid verbs, i.e. verbs having both A and B paradigms? How can one situation be homogeneous and non-homogeneous at the same time? The answer is: the situation itself cannot but its different phases can. At the same time it can be noted that all generally possible actional interpretations (state, process, entry into a state, entry into a process) and their occasional subtypes attested in Tundra Nenets (first processive phase, entry into first stative phase) refer to situations whose last phase is homogeneous: it is either a state or a process. Thus, Tundra Nenets intransitive paradigms serve as a morphological device for differentiating between two phases of an event: the first, a non-homogeneous one, and the second, a homogeneous one. The former is expressed by paradigm B, the latter is expressed by paradigm A.

Table 7.7 lists all possible pairs of actional interpretations of fluid verbs manifested in paradigm A and in paradigm B.\(^\text{18}\)

(18) a. *jembit yádi-Ø.  
dress be.torn.GF-3SGA  
‘The dress has got torn.’

b. *jembit yádi-Ø.  
dress be.torn.SF-3SGB  
‘The dress has begun to tear.’

\(^\text{18}\) See Appendix 7.1 for lists of verbs belonging to each class.
The existence of seven verbs in (d) and (e) in Table 7 seems to conflict with the foregoing statement regarding homogeneity: the phase of the situation expressed by the A paradigm is not homogeneous. Yet I would claim that it does not contradict the generalization but gives it a more precise shape.

These seven verbs suggest that the exact location of the division between two phases of an event represents a part of lexical semantics. While for all other verbs investigated the event is divided into the first, non-homogeneous dynamic phase and the following, homogeneous phase, these seven verbs are different. Instead of expressing the homogeneous phase of the situation, the paradigm A of these verbs expresses the last non-homogeneous subphase, the one which immediately precedes the homogeneous one. Such quirky behaviour of these verbs can be explained by the complex nature of the event expressed by them in contrast to the simple nature of the event expressed by all other verbs investigated here. Actually, the verbs in (d) and (e) describe a situation with four consequent stages:

(d) entry into first stative phase (B), first stative phase,
   entry into final stative phase (A), final stative phase;

(e) entry into a process (B), process,
   entry into a state (A), state;

while verbs in (a)–(c) describe a situation with two consequent stages only:

(a) entry into a process (B), process (A);
(b) first processive phase (B), final processive phase (A);
(c) entry into a state (B), state (A).

Thus, when involving verbs describing complex events, each intransitive paradigm expresses a respective block of sub-phases, the first (B paradigm) or the second (A paradigm). The first sub-phase of each block stands for the immediate semantic host of the paradigms. It should be stressed, though, that only in this case can the paradigm A denote entry into a state, a non-homogeneous situation it never expresses otherwise.

So the final account of the Tundra Nenets intransitive split can be formulated as follows. Verbal lexical semantics contains information about whether the situation the verb refers to can be seen as having two phases. If it cannot, only one intransitive paradigm is available, A or B, depending on the verb’s actional interpretation: homogeneous situations require the A paradigm, while non-homogeneous situations require the B paradigm. If the situation the verb refers to can be seen as having two phases, it uses both intransitive paradigms: the first non-homogeneous phase is encoded by the B paradigm, while the second homogeneous phase is encoded by the A paradigm. The verbal lexical semantics also contains information about whether two phases of a situation can be further seen as having sub-phases. In the latter case, the same consecutive logic is observed: the first of two resumptive phases is encoded by the B paradigm, and the second of
the two resumptive phases is encoded by the A paradigm. The principal of causal order overrides the homogeneity principle, as the A paradigm encodes a non-homogeneous, albeit final, phase of the situation.

The only thing which remains to be explained is the non-fluid A intransitive verbs denoting entry into a state, i.e. a non-homogeneous situation. Actually, all of them have the same morphonological structure: they are monosyllabic underived verbs, i.e. they follow a synchronically unproductive model.

\[
\begin{align*}
pjî\text{-} & \text{ 'to get cooked'} & ma\text{-} & \text{ 'to say'} \\
tu\text{-} & \text{ 'to enter'} & xa\text{-} & \text{ 'to leave; to become'} \\
to\text{-} & \text{ 'to come'} & xa\text{-} & \text{ 'to die'}
\end{align*}
\]

(19) \(\text{ja pjî-Ø} \quad /^*pjî-ji\text{-} ?.
\)

flour.soup get.cooked.GF-3SGA / get.cooked-SF-3SGB

'The soup has cooked.'

Monosyllabic underived verbs of various languages are generally assumed to be fairly old. Presumably, they have reached their current shape with the help of phonetic reduction conditioned by their high frequency (cf. the meanings of the Tundra Nenets verbs given above). The same high frequency might often be the reason behind their irregular morphosyntax.

Thus for the Tundra Nenets case it can be further assumed that the monosyllabic verbs in question were used as such already in the ancestor of the Tundra Nenets language, which had only one intransitive paradigm, namely the A paradigm.\(^{20}\) When the B paradigm emerged in the language, these verbs did not change the earlier profile, although by their semantics they should have been assigned to the verbs taking the new B paradigm.

For the 22 fluid verbs no differences in actional interpretation and, more generally, in meaning were attested: in both paradigms they denoted either a process (18 verbs) or an entry into a state (4 verbs). Such a small quantity, and the existence of verbs with similar pairs of actional types (see (b) and (d) in Table 7), suggest a descriptive oversight, which will not be significant for the generalization pertaining to 252 verbs.

For seven fluid verbs, the actional interpretations attested in paradigms A and B represent a reverse version of causal order: A paradigm expresses the first phase of situation, while B paradigm expresses the second phase. Once again, the small number of these cases might point to a descriptive oversight, or to occasional confusion by a native speaker commenting on the difference between two paradigms.

---

19 There are also three dissyllabic underived verbs in the list (loxom\text{-} 'to boil up (intr.)'; mîllo\text{-} 'to break down (intr.)'; jaxo\text{-} 'get lost'). I have no explanations, except idiosyncratic ones, for their adherence to the A paradigm.

20 As noted above, the A paradigm existed already in Proto-Uralic, while the B paradigm appeared as an independent paradigm only in Proto-North Samoyedic.
For one fluid verb, *jilje-*-, a lexicalization of the paradigms was attested: the verb means 'to live somewhere' when used with paradigm A, but 'to become alive ('to start living somewhere)' with paradigm B.

### 7.3.3 Evidence from non-verbal predication

In nonverbal predication, the nominal predicate takes the person–number endings of the intransitive A paradigm, if unpossessed, or the transitive singular paradigm if possessed (the possessee is treated as an object).

```
мāнъя тёдъя сава-ва” ханена-ва”.
mān’ja? įeda? sawa-wa? xan’ena-wa?.
I-pl now good-1PLA hunter-1PLA

'We are good hunters now.' (Tereschenko 1965: 919)
```

The exclusive use of intransitive A paradigm suggests another argument for the analysis presented above: nonverbal predication refers to states or properties assigned to the subject, and these are always homogeneous.

### 7.3.4 Typological parallels to semantic alignment systems

The phenomenon of two intransitive paradigms described above is reminiscent, first, of split-S marking (the verbal lexicon is divided into groups defined by possible person–number marking in the intransitive clause) and, secondly, of fluid-S marking (a considerable part of the verbal lexicon allows both types of person–number marking). At the same time, neither of the intransitive markings is identical to the way that arguments of transitive verbs (agents or patients) are treated; thus, we are not dealing with a prototypical semantic alignment system.

However, the differential treatment of arguments of intransitive verbs is governed by the parameter of verbal inner aspect, and this brings the Tundra Nenets system into line with semantically aligned languages treated elsewhere in this volume. Actually, the structural factors operating on intransitive verbs in Tundra Nenets are of the same nature as those found in semantic alignment systems.

Moreover, syntactically accusative Tundra Nenets can be characterized as a language with general event-internal logic typical for semantically aligned languages. Of the four event-internal categories enumerated in Wichmann (2005a) Tundra Nenets makes use of three, i.e. all except semantic alignment.

Thus, its valency-change operations are sensitive to verbal semantics: e.g. there is no purely passive morphology; instead, the intransitive B paradigm of verbs with causative derivation, as well as perfect participles, are used in this function (Tereschenko 1973, Farrell Ackerman, p.c.). Both seem pretty clear cases of patientive resultatives, since they do not modify the structure of an event but rather select an aspect of it for special focus.

Tundra Nenets is definitely an aspect-oriented and not a tense-oriented language. Most frequent in communication is the so-called 'indefinite
tense\textsuperscript{21} which actually has no markers: all the Tundra Nenets sentences cited above (except (14c)) exemplified this. The time reference in the case of absence of any tense markers depends on telicity of the given paradigm of the verb: in (2), (3)–(8), (9a), (11), (12), (13b), (14b), (15b), and (17)–(19) the sentence has a telic interpretation and the whole clause describes a past situation, while in (1), (9b), (13a), (13c), (14a), (15a), and (16) the sentence has an atelic interpretation and the whole clause has present reference. The only tense that has a special marker is ‘preterite’ with its postfix \(-s\hat{\mathring{a}}\), cf (14c), which is characterized not as ‘plain past’ but as ‘relative past, opposed to present’ (Tereschenko 1965: 895). Judged by all structural properties (structure of negative form, co-occurrence with preterite marker), the future marker \(-N\hat{\mathring{o}}-/-t\hat{\mathring{a}}-\) is an aspectual derivation and not a tense marker. Additionally, there are at least five other highly productive aspectual derivations.\textsuperscript{22}

Finally, property concepts, either nouns or true adjectives, are treated as verbs when used in predicating function. As mentioned in section 7.3.3, they take the person–number endings of intransitive A paradigm when unpossessed, or transitive singular paradigm when possessed.

Hence, an intransitive split based on differences in verbal lexical semantics does not represent anything unusual for Tundra Nenets. On the contrary, its whole logic supports a view of this relatively recent structural feature as a continuation of the tendency already existing in the language.

7.4 Summary and conclusion

I have examined the semantic principles underlying the choice between two intransitive paradigms in Tundra Nenets, a North Samoyedic (Uralic) language. According to the evidence presented here, two parameters are at work: event homogeneity and the causal order of subevents. While one intransitive paradigm, the B paradigm, encodes non-homogeneous events or a non-homogeneous first phase of events, the other intransitive paradigm, the A paradigm, encodes homogeneous events or a homogeneous last phase of events. In the rare cases where it is impossible to divide an event into two ontologically equally important phases, one of which is homogeneous and the other of which is non-homogeneous, only the causal order parameter is applied. As mentioned above, the first phase is encoded by the B paradigm and the second phase by the A paradigm.

The nature of the semantic parameters and the lexical domain they act on have allowed a comparison with semantically aligned systems, where intransitives are treated differently depending on their verbal inner aspect, in particular. This comparability has been reinforced by the fact that the choice between the two intransitive paradigms in Tundra Nenets follows the same principles whether it takes place in the lexicon or in the grammar. Identical parameters were shown to

\textsuperscript{21} ‘Past tense forms are met considerably more seldom than indefinite tense forms’ (Tereschenko 1965: 895).

\textsuperscript{22} Compare the insignificance of tense in the system of closely related Enets (North Samoyedic), as discussed in Urmanchieva (2006).
condition the paradigm choice for the one-paradigm verbs (lexical choice) and for the two-paradigm verbs (grammatical choice). The similar phenomena of lexical intransitive split, on the one hand, and fluid intransitive split, i.e. grammatical intransitive split, on the other, are the constituting ones in semantically aligned languages.

These results raise interesting questions with regard to the conceptual distance between semantically and syntactically aligned languages. If the same semantic principles can be found as organizing principles of the lexicon of both types, the main difference between the two types of language may lie in the strength of syntax which either blocks or distorts the semantic principles (in syntactically aligned languages) or leaves them intact (in semantically aligned languages). However, the semantic parameters themselves are chosen from the same pool, and should thus be studied from the same perspective.

The suggested approach unifying syntactically aligned languages with semantically aligned languages seems more promising than a separating treatment. Insisting on the absolute uniqueness of semantically aligned systems might lead us to overlook significant insights from syntactically aligned languages.

Appendix 7.1 Verbs from the database by their paradigms

All translation equivalents are approximate, due to highly specified aspectual shades of meanings which are sometimes difficult to render by English infinitives. The value of the telicity parameter was/is not always visible directly from the (English) translation of the verb: for all such cases, the verb’s compatibility with a time adverbial was checked in order to determine telicity (e.g. ‘for 5 minutes’ (atelic) vs. ‘in 5 minutes’ (telic)).

7.1.1 Only B paradigm verbs

7.1.1.1 Actional interpretation: ‘entry into a state’

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. mājud</td>
<td>to become glad (from something unexpected)</td>
<td></td>
</tr>
<tr>
<td>2. maʔ-nā</td>
<td>to fall down, to pounce on</td>
<td>IPFV</td>
</tr>
<tr>
<td>3. gəndā</td>
<td>to sit down</td>
<td></td>
</tr>
<tr>
<td>4. pańa</td>
<td>to become full</td>
<td></td>
</tr>
<tr>
<td>5. sāl</td>
<td>to return</td>
<td></td>
</tr>
<tr>
<td>6. sara</td>
<td>to burst (of a bulging object)</td>
<td></td>
</tr>
<tr>
<td>7. tāna</td>
<td>to get up</td>
<td></td>
</tr>
<tr>
<td>8. tərpā</td>
<td>to go out</td>
<td></td>
</tr>
<tr>
<td>9. tum</td>
<td>to take fire</td>
<td></td>
</tr>
<tr>
<td>10. xona</td>
<td>to go to sleep, to fall asleep</td>
<td></td>
</tr>
</tbody>
</table>
### 7.1.1.2 Actional interpretation: 'entry into a process'

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>jøxol</td>
<td>to start swimming</td>
<td></td>
</tr>
</tbody>
</table>

### 7.1.1.3 Actional interpretation: 'first processive phase'

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>jil(e)-l-nà</td>
<td>to start living (somewhere)</td>
<td>INCH-IPFV</td>
</tr>
<tr>
<td>pad(à)-ba-nà</td>
<td>to start being written</td>
<td>DUR-IPFV</td>
</tr>
<tr>
<td>peda-nà</td>
<td>to start being tired</td>
<td>IPFV</td>
</tr>
<tr>
<td>p(i)-is(å)-nà</td>
<td>to start laughing</td>
<td>IPFV</td>
</tr>
<tr>
<td>sara-da-nà</td>
<td>to start bursting (of a bulging object)</td>
<td>TRR-IPFV</td>
</tr>
<tr>
<td>tårpå-pta-nà</td>
<td>to be going out, to be about to go out</td>
<td>CAUS-IPFV</td>
</tr>
<tr>
<td>tira-bta-nà</td>
<td>to start getting dry</td>
<td>CAUS-IPFV</td>
</tr>
<tr>
<td>xàwnà-(l)-nà</td>
<td>to start raining</td>
<td>INCH-IPFV</td>
</tr>
</tbody>
</table>

### 7.1.2 Only A paradigm verbs

### 7.1.2.1 Actional interpretation: 'process'

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>jadà</td>
<td>to go, to walk</td>
<td>INCH-ITER</td>
</tr>
<tr>
<td>jadà-(l)-ngà</td>
<td>to start walking (esp. from time to time)</td>
<td>INCH-ITER</td>
</tr>
<tr>
<td>jen'(e)r-bà</td>
<td>to shoot at various targets (esp. from time to time)</td>
<td>DUR</td>
</tr>
<tr>
<td>jewej</td>
<td>to eat (flour) soup</td>
<td>FR-INCH-CAUS-ITER</td>
</tr>
<tr>
<td>jib(e)'-ed-ol-pta-ngà</td>
<td>to understand, to start thinking about</td>
<td>FR-INCH-CAUS-ITER</td>
</tr>
<tr>
<td>jib(e)'-e?</td>
<td>to think (about), to be clever</td>
<td>ITER</td>
</tr>
<tr>
<td>jib(e)'-e?-ngà</td>
<td>to think about (esp. from time to time)</td>
<td>ITER</td>
</tr>
<tr>
<td>jil(e)-l-pà</td>
<td>to live (somewhere)</td>
<td>INCH-DUR</td>
</tr>
<tr>
<td>jil(e)-r-ngà</td>
<td>to live (somewhere)</td>
<td>FR-ITER</td>
</tr>
<tr>
<td>jinz(e)'el-e-ngà</td>
<td>to listen to, to overhear (esp. from time to time)</td>
<td>ITER</td>
</tr>
</tbody>
</table>

(continues)
Continued

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>11. jor⁽⁽e⁾⁾w-or</td>
<td>to deepen (esp. from time to time)</td>
<td>FR</td>
</tr>
<tr>
<td>12. jor-ngä</td>
<td>to deepen (esp. gradually or from time to time)</td>
<td>ITER</td>
</tr>
<tr>
<td>13. joxo-mbä/juxu-mbä</td>
<td>to be lost, to disappear (esp., from time to time)</td>
<td>DUR</td>
</tr>
<tr>
<td>14. lad(ä)-bä-ngä</td>
<td>to tap (e.g. a door)</td>
<td>DUR-ITER</td>
</tr>
<tr>
<td>15. latra-ngä</td>
<td>to squeeze (e.g. in the door) (esp. from time to time)</td>
<td>ITER</td>
</tr>
<tr>
<td>16. läxänä-mbä</td>
<td>to talk (esp. from time to time)</td>
<td>DUR</td>
</tr>
<tr>
<td>17. laxo-(m)ba</td>
<td>to boil</td>
<td>DUR</td>
</tr>
<tr>
<td>18. mäjäm-ngä</td>
<td>to rejoice (esp. from time to time)</td>
<td>ITER</td>
</tr>
<tr>
<td>19. mäj(ä)-mbä</td>
<td>to rejoice</td>
<td>DUR</td>
</tr>
<tr>
<td>20. mäd-nä</td>
<td>to bark</td>
<td>IPFV</td>
</tr>
<tr>
<td>21. men’e-l-ngä</td>
<td>to be falling in love</td>
<td>INCH-ITER</td>
</tr>
<tr>
<td>22. m/i?</td>
<td>to go</td>
<td></td>
</tr>
<tr>
<td>23. gädara-ngä</td>
<td>to get torn (gradually)</td>
<td>ITER</td>
</tr>
<tr>
<td>24. gäd-nä</td>
<td>to be getting torn</td>
<td>IPFV</td>
</tr>
<tr>
<td>25. gädo-mbä</td>
<td>to get torn (esp. gradually or from time to time)</td>
<td>DUR</td>
</tr>
<tr>
<td>26. gämđä-ngä</td>
<td>to be sitting down</td>
<td>ITER</td>
</tr>
<tr>
<td>27. gämđlo-ngä</td>
<td>to take a seat (esp. a number of times)</td>
<td>ITER</td>
</tr>
<tr>
<td>28. gat’e-l-ngä</td>
<td>to start waiting (esp. from time to time)</td>
<td>INCH-ITER</td>
</tr>
<tr>
<td>29. ni/el’e-ngä</td>
<td>to get married (esp. from time to time)</td>
<td>ITER</td>
</tr>
<tr>
<td>30. ni/el’e-pta-ngä</td>
<td>to marry (esp. a number of times)</td>
<td>CAUS-ITER</td>
</tr>
<tr>
<td>31. ni’uc’a-ngu-r-nä</td>
<td>to kiss each other (esp. from time to time)</td>
<td>ITER-FR-IPFV</td>
</tr>
<tr>
<td>32. Ėoxola-ngä</td>
<td>to swim (esp. from time to time)</td>
<td>ITER</td>
</tr>
<tr>
<td>33. padä-ngä</td>
<td>to proceed to the writing (esp. from time to time)</td>
<td>ITER</td>
</tr>
<tr>
<td>34. pad(ä)-nä-ngä</td>
<td>to write (esp. from time to time)</td>
<td>IPFV-ITER</td>
</tr>
<tr>
<td>35. paje-r’e-d/e-tgä</td>
<td>to have abscess/boils appearing</td>
<td>TRR-TRR-ITER</td>
</tr>
<tr>
<td>36. pana-ngä</td>
<td>to fill (intr.) (esp. from time to time)</td>
<td>ITER</td>
</tr>
<tr>
<td>37. peda-mbä</td>
<td>to get tired (esp. from time to time)</td>
<td>DUR</td>
</tr>
<tr>
<td>38. pede-ngä</td>
<td>to get tired (esp. from time to time)</td>
<td>ITER</td>
</tr>
<tr>
<td>39. p’ina-mbä</td>
<td>to take a slight fright of (esp. from time to time)</td>
<td>DUR</td>
</tr>
</tbody>
</table>

(cont.)
Continued

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>40. p’inā-IGā</td>
<td>to take fright of (esp. from time to time)</td>
<td>ITER</td>
</tr>
<tr>
<td>41. p’inā-pta-IGā</td>
<td>to take fright of (esp. from time to time)</td>
<td>CAUS-ITER</td>
</tr>
<tr>
<td>42. p’in(ā)-lā-IGā</td>
<td>to take fright of (esp. from time to time)</td>
<td>INCH-ITER</td>
</tr>
<tr>
<td>43. p’i-IGā</td>
<td>to be cooked, to be boiling (for some time already)</td>
<td>ITER</td>
</tr>
<tr>
<td>44. p’is/ā-mbā</td>
<td>to chuckle, to laugh softly</td>
<td>DUR</td>
</tr>
<tr>
<td>45. sara-mbā</td>
<td>to be bursting out</td>
<td>DUR</td>
</tr>
<tr>
<td>46. sara-IGā</td>
<td>to be almost bursting (of a bulging object), to burst (esp. from time to time)</td>
<td>ITER</td>
</tr>
<tr>
<td>47. so</td>
<td>to be heard, to resound</td>
<td>DUR</td>
</tr>
<tr>
<td>48. tāna-mbā</td>
<td>to ascend (esp. from time to time)</td>
<td>DUR</td>
</tr>
<tr>
<td>49. tārpā-mbā</td>
<td>to go out, to leave</td>
<td>DUR</td>
</tr>
<tr>
<td>50. tārpā-IGā</td>
<td>to go out (esp. from time to time)</td>
<td>ITER</td>
</tr>
<tr>
<td>51. tira-IGā</td>
<td>to be getting dry</td>
<td>ITER</td>
</tr>
<tr>
<td>52. t’i-mbā</td>
<td>to fly (esp. from time to time or without any goal/direction)</td>
<td>DUR</td>
</tr>
<tr>
<td>53. t’im-bā</td>
<td>to rot</td>
<td>DUR</td>
</tr>
<tr>
<td>54. t’im-la-IGā</td>
<td>to get rotten (gradually)</td>
<td>INCH-ITER</td>
</tr>
<tr>
<td>55. t’im-na-mbā</td>
<td>to rot</td>
<td>IPFV-DUR</td>
</tr>
<tr>
<td>56. t’im-na-IGā</td>
<td>to rot, to continue rotting</td>
<td>IPFV-ITER</td>
</tr>
<tr>
<td>57. t’im-IGā</td>
<td>to rot</td>
<td>ITER</td>
</tr>
<tr>
<td>58. t’or’e-IGā</td>
<td>to utter screams (esp. from time to time)</td>
<td>ITER</td>
</tr>
<tr>
<td>59. xāmadum-bā</td>
<td>to flow</td>
<td>DUR</td>
</tr>
<tr>
<td>60. xāda-pt-or</td>
<td>to grasp, to snatch (esp. one after another)</td>
<td>CAUS-FR</td>
</tr>
<tr>
<td>61. xa-mbā</td>
<td>to be sick</td>
<td>DUR</td>
</tr>
<tr>
<td>62. xa-IGā</td>
<td>to die (esp. from time to time)</td>
<td>ITER</td>
</tr>
<tr>
<td>63. xa’aw-ra-IGā</td>
<td>to fall down</td>
<td>TRR-ITER</td>
</tr>
<tr>
<td>64. xāv-da-IGā</td>
<td>to fall down (esp. from time to time)</td>
<td>TRR-ITER</td>
</tr>
<tr>
<td>65. xoda’-pā</td>
<td>to give slight coughs</td>
<td>DUR</td>
</tr>
<tr>
<td>66. xojsa-mbā</td>
<td>to waive</td>
<td>DUR</td>
</tr>
<tr>
<td>67. xon(i-m)bā</td>
<td>to be sleepy, to show signs of being sleepy</td>
<td>DUR</td>
</tr>
<tr>
<td>68. xon’o</td>
<td>to sleep</td>
<td>DUR</td>
</tr>
<tr>
<td>69. xon’o-mbā</td>
<td>to be sleepy (esp. from time to time), to show signs of being sleepy</td>
<td>DUR</td>
</tr>
</tbody>
</table>
### Intransitive split in Tundra Nenets

#### 7.1.2.2 Actional interpretation: ‘state’

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ḡādˈā</td>
<td>to be visible</td>
<td></td>
</tr>
<tr>
<td>2. nu</td>
<td>to stand</td>
<td></td>
</tr>
<tr>
<td>3. sˈārā</td>
<td>to be kept, to be stored</td>
<td></td>
</tr>
<tr>
<td>4. wāndartā</td>
<td>to be crooked</td>
<td></td>
</tr>
<tr>
<td>5. jakā</td>
<td>to itch</td>
<td></td>
</tr>
<tr>
<td>6. jaku-mbā</td>
<td>to itch</td>
<td>DUR</td>
</tr>
</tbody>
</table>

#### 7.1.2.3 Actional interpretation: ‘entry into a state’

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. joxo/juxu</td>
<td>to get lost</td>
<td></td>
</tr>
<tr>
<td>2. laxom</td>
<td>to boil up (intr.)</td>
<td></td>
</tr>
<tr>
<td>3. mālˈjjo</td>
<td>to break down (intr.)</td>
<td></td>
</tr>
<tr>
<td>4. maˈ</td>
<td>to say</td>
<td></td>
</tr>
<tr>
<td>5. pˈi</td>
<td>to get cooked</td>
<td></td>
</tr>
<tr>
<td>6. tˈu</td>
<td>to enter</td>
<td></td>
</tr>
<tr>
<td>7. to</td>
<td>to come</td>
<td></td>
</tr>
<tr>
<td>8. xa</td>
<td>to die</td>
<td></td>
</tr>
<tr>
<td>9. xæ</td>
<td>to leave; to become</td>
<td></td>
</tr>
</tbody>
</table>

#### 7.1.3 A and B paradigm verbs

##### 7.1.3.1 Actional interpretation of A: process (atelic), actional interpretation of B: entry into a process (telic)

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning: A paradigm</th>
<th>Meaning: B paradigm</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. jadā-bt-or</td>
<td>to walk slowly/with difficulty</td>
<td>to start walking slowly/with difficulty</td>
<td>CAUS-FR</td>
</tr>
<tr>
<td>2. jad-er</td>
<td>to walk, to go (esp. from time to time)</td>
<td>to start walking, to start going (esp. walking from time to time)</td>
<td>FR (cont.)</td>
</tr>
<tr>
<td>Verb</td>
<td>Meaning: A paradigm</td>
<td>Meaning: B paradigm</td>
<td>Affixes</td>
</tr>
<tr>
<td>--------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
<td>---------</td>
</tr>
<tr>
<td>3. jad(ā)-n-or</td>
<td>to go, to visit (esp. seldom or irregularly)</td>
<td>to start going, to start visiting (esp. visiting seldom or irregularly)</td>
<td>IPFV-FR</td>
</tr>
<tr>
<td>4. jemb-t-or</td>
<td>to be getting dressed</td>
<td>to start getting dressed</td>
<td>TRR-FR</td>
</tr>
<tr>
<td>5. jen'er</td>
<td>to shoot</td>
<td>to start shooting</td>
<td>FR</td>
</tr>
<tr>
<td>6. jen'er-or</td>
<td>to shoot (esp. from time to time)</td>
<td>to start shooting (esp. shooting from time to time)</td>
<td></td>
</tr>
<tr>
<td>7. jewejn</td>
<td>to eat (flour) soup</td>
<td>to start eating (flour) soup</td>
<td>FR</td>
</tr>
<tr>
<td>8. jewej-or</td>
<td>to eat (flour) soup (often)</td>
<td>to start eating (flour) soup (often)</td>
<td></td>
</tr>
<tr>
<td>9. jib'ed-or</td>
<td>to think (about)</td>
<td>to start thinking (about), to get lost in thoughts</td>
<td>FR</td>
</tr>
<tr>
<td>10. jinz'el/e</td>
<td>to listen to</td>
<td>to start listening to</td>
<td>FR</td>
</tr>
<tr>
<td>11. jor-or</td>
<td>to deepen (esp. gradually or from time to time)</td>
<td>to start deepening (esp. deepening gradually or from time to time)</td>
<td></td>
</tr>
<tr>
<td>12. joxo-r/ juxu-r</td>
<td>to get lost (esp. from time to time), to be lost</td>
<td>to start getting lost (esp. from time to time)</td>
<td>FR</td>
</tr>
<tr>
<td>13. lāxānā</td>
<td>to talk</td>
<td>to start talking</td>
<td>FR</td>
</tr>
<tr>
<td>14. lāxän-or</td>
<td>to talk with</td>
<td>to start talking with (e.g. a door)</td>
<td>CAUS-FR</td>
</tr>
<tr>
<td>15. ladā-pt-or</td>
<td>to tap (e.g. a door)</td>
<td>to start tapping (e.g. a door)</td>
<td></td>
</tr>
<tr>
<td>16. ladā-r</td>
<td>to jerk, to twitch</td>
<td>to start jerking, to start twitching</td>
<td>FR</td>
</tr>
<tr>
<td>17. lamb-or</td>
<td>to ski (esp. from time to time)</td>
<td>to start skiing (esp. skiing from time to time)</td>
<td>FR</td>
</tr>
<tr>
<td>18. latr-or</td>
<td>to force through (intr.)</td>
<td>to start forcing through (intr.)</td>
<td>FR</td>
</tr>
<tr>
<td>19. loxo-r</td>
<td>to purl, to bubble, to murmur (esp. of water)</td>
<td>to start purling, to start bubbling, to start murmuring (esp. of water)</td>
<td>FR</td>
</tr>
<tr>
<td>20. loxow-or</td>
<td>to be almost boiling</td>
<td>to start being almost boiling</td>
<td>FR</td>
</tr>
</tbody>
</table>

(Cont.)
<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning: A paradigm</th>
<th>Meaning: B paradigm</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>21. mǎl’j-ur</td>
<td>to be getting broken</td>
<td>to start getting broken</td>
<td>FR</td>
</tr>
<tr>
<td>22. mànža-ra</td>
<td>to work</td>
<td>to start working</td>
<td>TRR</td>
</tr>
<tr>
<td>23. madr-or</td>
<td>to bark</td>
<td>to start barking</td>
<td>FR</td>
</tr>
<tr>
<td>24. mànza-d-or</td>
<td>to move (intr.)</td>
<td>to start moving (intr.)</td>
<td>TRR-FR</td>
</tr>
<tr>
<td>25. mànza-p’ta-ngā</td>
<td>to move (intr.)</td>
<td>to start moving (intr.)</td>
<td>CAUS-ITER</td>
</tr>
<tr>
<td>26. mànza-p’t-or</td>
<td>to move (intr.)</td>
<td>to start moving (intr.)</td>
<td>CAUS-FR</td>
</tr>
<tr>
<td>27. masta-ngā</td>
<td>to move (intr.)</td>
<td>to start moving (intr.)</td>
<td>ITER</td>
</tr>
<tr>
<td>28. mast-or</td>
<td>to move (intr.)</td>
<td>to start moving (intr.)</td>
<td>TRR-FR</td>
</tr>
<tr>
<td>29. mo’n-or</td>
<td>to fall down, to go and fall down</td>
<td>to start falling down, to start going and falling down</td>
<td>IPFV-FR</td>
</tr>
<tr>
<td>30. ñàd’im</td>
<td>to be visible, to come into view</td>
<td>to become visible, to come into view</td>
<td>FR</td>
</tr>
<tr>
<td>31. ñam’d-or</td>
<td>to sit down (esp. from time to time)</td>
<td>to start sitting down (esp. sitting down from time to time)</td>
<td>FR</td>
</tr>
<tr>
<td>32. ñer-or/ñer-ur</td>
<td>to drink (esp. from time to time), to hit the bottle</td>
<td>to start drinking (esp. drinking from time to time), to start hitting the bottle</td>
<td>FR</td>
</tr>
<tr>
<td>33. n’ā?m-nà</td>
<td>to take (esp. from time to time)</td>
<td>to start taking (esp. taking from time to time)</td>
<td>IPFV</td>
</tr>
<tr>
<td>34. ñoxolà</td>
<td>to swim</td>
<td>to start swimming</td>
<td>FR</td>
</tr>
<tr>
<td>35. ñoxol-or</td>
<td>to swim</td>
<td>to start swimming</td>
<td>IPFV-FR</td>
</tr>
<tr>
<td>36. pad(ā)-n-or</td>
<td>to write (esp. from time to time)</td>
<td>to start writing</td>
<td></td>
</tr>
<tr>
<td>37. paje</td>
<td>to have abscess/boils appearing</td>
<td>to start having abscess/boils appearing</td>
<td></td>
</tr>
<tr>
<td>38. p’inā-n-or</td>
<td>to be slightly frightened of (esp. from time to time)</td>
<td>to be slightly frightened of (esp. from time to time)</td>
<td>IPFV-FR</td>
</tr>
<tr>
<td>39. p’isáʔ</td>
<td>to laugh</td>
<td>to start laughing</td>
<td>INCH-FR</td>
</tr>
<tr>
<td>40. p’is(ā)-l-or</td>
<td>to chuckle, to laugh softly</td>
<td>to start chuckling, to start laughing softly</td>
<td></td>
</tr>
</tbody>
</table>

*(cont.)*
<table>
<thead>
<tr>
<th>VERB</th>
<th>MEANING: A paradigm</th>
<th>MEANING: B paradigm</th>
<th>AFFIXES</th>
</tr>
</thead>
<tbody>
<tr>
<td>41. sara-d-or</td>
<td>to make sth. burst (esp. from time to time)</td>
<td>to start making sth. burst (esp. from time to time)</td>
<td>TRR-FR</td>
</tr>
<tr>
<td>42. ser-ur</td>
<td>to salt: a long-term activity of salting fish, meat, etc. in large quantities</td>
<td>to start salting: to start a long-term activity of salting fish, meat, etc. in large quantities</td>
<td>FR</td>
</tr>
<tr>
<td>43. tān-ir</td>
<td>to ascend (esp. from time to time), to go up and down</td>
<td>to start ascending (esp. from time to time), to start going up and down</td>
<td>FR</td>
</tr>
<tr>
<td>44. tārp-or/ tārp-ur</td>
<td>to be going out</td>
<td>to start going out, to prepare for going out</td>
<td>FR</td>
</tr>
<tr>
<td>45. tį</td>
<td>to fly</td>
<td>to fly up</td>
<td>FR</td>
</tr>
<tr>
<td>46. tį-i-r</td>
<td>to fly (esp. from time to time)</td>
<td>to start flying (esp. flying from time to time)</td>
<td>FR</td>
</tr>
<tr>
<td>47. tįw-or</td>
<td>to rot</td>
<td>to start rotting</td>
<td>FR</td>
</tr>
<tr>
<td>48. tįri-r</td>
<td>to scream</td>
<td>to start screaming</td>
<td>FR</td>
</tr>
<tr>
<td>49. tįu-gāa</td>
<td>to be entering</td>
<td>to start entering</td>
<td>ITER</td>
</tr>
<tr>
<td>50. tol-ir</td>
<td>to count</td>
<td>to start counting</td>
<td>FR</td>
</tr>
<tr>
<td>51. tu-i-or</td>
<td>to come (esp. seldom or from time to time)</td>
<td>to start coming (esp. from time to time)</td>
<td>INCH-FR</td>
</tr>
<tr>
<td>52. tu-r</td>
<td>to come (esp. from time to time)</td>
<td>to start coming (esp. coming from time to time)</td>
<td>FR</td>
</tr>
<tr>
<td>53. xād’et-or</td>
<td>to scratch (intr.)</td>
<td>to start scratching (intr.)</td>
<td>FR</td>
</tr>
<tr>
<td>54. xālt’o</td>
<td>to be washed, to launder (intr.)</td>
<td>to start being washed, to start laundering (intr.)</td>
<td>FR</td>
</tr>
<tr>
<td>55. xālt-ur</td>
<td>to be washed, to lose colour</td>
<td>to start being washed, to start losing colour</td>
<td>FR</td>
</tr>
<tr>
<td>56. xaju-r</td>
<td>to stay, to remain (esp. from time to time)</td>
<td>to start staying, to start remaining (esp. staying from time to time)</td>
<td>FR</td>
</tr>
<tr>
<td>57. xāmda-gā</td>
<td>to spill, to flow (esp. from time to time)</td>
<td>to start spilling, to start flowing (esp. spilling or flowing from time to time)</td>
<td>ITER</td>
</tr>
</tbody>
</table>

(cont.)
Intransitive split in Tundra Nenets

Continued

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning: A paradigm</th>
<th>Meaning: B paradigm</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>58. xâmdu-r/ xâmd-or</td>
<td>to flow (esp. from time to time)</td>
<td>to start flowing (esp. flowing from time to time)</td>
<td>FR</td>
</tr>
<tr>
<td>59. xâmduaw-or</td>
<td>to spill, to flow</td>
<td>to start spilling, to start flowing</td>
<td>FR</td>
</tr>
<tr>
<td>60. xa-nâ</td>
<td>to die</td>
<td>to start dying</td>
<td>IPFV</td>
</tr>
<tr>
<td>61. xân/iw-or</td>
<td>to freeze (esp. from time to time)</td>
<td>to start freezing</td>
<td>FR</td>
</tr>
<tr>
<td>62. xaîw-r-or</td>
<td>to fall (esp. of snow)</td>
<td>to start falling (esp. of snow)</td>
<td>TRR-FR</td>
</tr>
<tr>
<td>63. xod-nâ</td>
<td>to cough</td>
<td>to start coughing</td>
<td>IPFV</td>
</tr>
<tr>
<td>64. xod-or</td>
<td>to cough</td>
<td>to start coughing, to cough once</td>
<td>FR</td>
</tr>
<tr>
<td>65. xod?-pâ</td>
<td>to cough (esp. from time to time)</td>
<td>to start coughing (esp. coughing from time to time)</td>
<td>DUR</td>
</tr>
<tr>
<td>66. xojdi-r</td>
<td>to wave</td>
<td>to start waving</td>
<td>FR</td>
</tr>
<tr>
<td>67. xojsa</td>
<td>to wave</td>
<td>to start waving</td>
<td>ITER</td>
</tr>
<tr>
<td>68. xojsa-ngâ</td>
<td>to wave (esp. from time to time)</td>
<td>to start waving</td>
<td>ITER</td>
</tr>
<tr>
<td>69. xois-or</td>
<td>to wave</td>
<td>to start waving</td>
<td>FR</td>
</tr>
</tbody>
</table>

7.1.3.2 A: process (atelic), B: first processive phase (atelic)

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning: A paradigm</th>
<th>Meaning: B paradigm</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. jemb?-nâ</td>
<td>to get dressed</td>
<td>to start getting dressed</td>
<td>IPFV</td>
</tr>
<tr>
<td>2. jen'er-ngâ</td>
<td>to shoot</td>
<td>to start shooting</td>
<td>ITER</td>
</tr>
<tr>
<td>3. jib'ed-or-ta-nâ</td>
<td>to think (about)</td>
<td>to start thinking (about)</td>
<td>FR-TRR-IPFV</td>
</tr>
<tr>
<td>4. jil'e-mbâ</td>
<td>to live somewhere</td>
<td>to start living   somewhere</td>
<td>DUR</td>
</tr>
<tr>
<td>5. jil'e-mbâ-nâ</td>
<td>to live somewhere</td>
<td>to start living somewhere</td>
<td>DUR-IPFV</td>
</tr>
</tbody>
</table>

(cont.)
<table>
<thead>
<tr>
<th>VERB</th>
<th>MEANING: A paradigm</th>
<th>MEANING: B paradigm</th>
<th>AFFIXES</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. jil'e-pta-nā</td>
<td>to live somewhere, to be alive</td>
<td>to start living somewhere, to revive</td>
<td>CAUS-IPFV</td>
</tr>
<tr>
<td>7. jinζ/el'e-nā</td>
<td>to listen to</td>
<td>to start listening to</td>
<td>IPFV</td>
</tr>
<tr>
<td>8. joxo-nā/juxu-nā</td>
<td>to get lost (e.g. to disappear gradually)</td>
<td>to start getting lost, to try to get lost</td>
<td>IPFV</td>
</tr>
<tr>
<td>9. jut(ā)-nā</td>
<td>to fight (intr.)</td>
<td>to start fighting (intr.)</td>
<td>IPFV</td>
</tr>
<tr>
<td>10. lad-pta-nā</td>
<td>to hit, to tap (e.g. a door)</td>
<td>to start hitting, to start tapping (e.g. a door)</td>
<td>CAUS-IPFV</td>
</tr>
<tr>
<td>11. lad(ā)-nā</td>
<td>to hit</td>
<td>to start hitting</td>
<td>IPFV</td>
</tr>
<tr>
<td>12. laxom-bā-nā</td>
<td>to boil (intr.)</td>
<td>to start boiling</td>
<td>DUR-IPFV</td>
</tr>
<tr>
<td>13. laxom-nā</td>
<td>to be almost boiling</td>
<td>to start being almost boiling</td>
<td>IPFV</td>
</tr>
<tr>
<td>14. mājām-lā-ygā</td>
<td>to rejoice (esp. from time to time)</td>
<td>to start rejoicing (esp. from time to time)</td>
<td>TRR-ITER</td>
</tr>
<tr>
<td>15. māj(ā)m-lāda-nā</td>
<td>to start rejoicing</td>
<td>to start rejoicing (a phase previous to the phase of A paradigm)</td>
<td>INCH-TRR-IPFV</td>
</tr>
<tr>
<td>16. māj(ā)m-lāmbā-nā</td>
<td>to rejoice</td>
<td>to start rejoicing</td>
<td>TRR-DUR-IPFV</td>
</tr>
<tr>
<td>17. māj(ā)m-lānā</td>
<td>to rejoice</td>
<td>to start rejoicing</td>
<td>TRR-IPFV</td>
</tr>
<tr>
<td>18. māl'jo-nā</td>
<td>to break down (intr.)</td>
<td>to start breaking down (intr.)</td>
<td>IPFV</td>
</tr>
<tr>
<td>19. mal'je-nā</td>
<td>to break down (intr.)</td>
<td>to start breaking down (intr.)</td>
<td>IPFV</td>
</tr>
<tr>
<td>20. men'e-nā</td>
<td>to please someone (esp. not definitely yet)</td>
<td>to start pleasing someone</td>
<td>IPFV</td>
</tr>
<tr>
<td>21. ḫāda-nā</td>
<td>to get torn</td>
<td>to start getting torn</td>
<td>IPFV</td>
</tr>
<tr>
<td>22. ḫāda-ra-nā</td>
<td>to get torn</td>
<td>to start getting torn</td>
<td>TRR-IPFV</td>
</tr>
<tr>
<td>23. ḫādo-pta-nā</td>
<td>to get torn</td>
<td>to start getting torn</td>
<td>CAUS-IPFV</td>
</tr>
<tr>
<td>24. ḫut'e-nā</td>
<td>to await</td>
<td>to start awaiting (esp. to start awaiting regularly)</td>
<td>IPFV</td>
</tr>
</tbody>
</table>

(cont.)
<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning: A paradigm</th>
<th>Meaning: B paradigm</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>25. ъамда-нъа</td>
<td>to sit down</td>
<td>to start sitting</td>
<td>IPFV</td>
</tr>
<tr>
<td>26. ъаъва-la-нъа</td>
<td>to eat</td>
<td>to start eating</td>
<td>TRR-IPFV</td>
</tr>
<tr>
<td>27. н’əржа-иягъа</td>
<td>to grow red, to turn red</td>
<td>to start growing red, to start turning red</td>
<td>ITER</td>
</tr>
<tr>
<td>28. н’ел’е-нъа</td>
<td>to get married</td>
<td>to start getting married</td>
<td>IPFV</td>
</tr>
<tr>
<td>29. н’ел’е-p’е-оър</td>
<td>to woo, to ask in marriage</td>
<td>to woo, to ask in marriage (esp. of unsuccessful wooing or of early stage of wooing)</td>
<td>CAUS-FR</td>
</tr>
<tr>
<td>30. н’ук’а-нъа</td>
<td>to kiss</td>
<td>to start kissing</td>
<td>IPFV</td>
</tr>
<tr>
<td>31. н’ук’а-п’а-оър-п’а-нъа</td>
<td>to kiss (esp. slightly, off and on)</td>
<td>to start kissing</td>
<td>ITER-FR-DUR-IPFV</td>
</tr>
<tr>
<td>32. ыъола-нъа</td>
<td>to be going to swim, to start swimming</td>
<td>to be going to swim, to start swimming (a phase previous to the phase of A paradigm)</td>
<td>IPFV</td>
</tr>
<tr>
<td>33. ыъол’-ур-нъа</td>
<td>to swim (esp. from time to time)</td>
<td>to start swimming (esp. swimming from time to time)</td>
<td>FR-IPFV</td>
</tr>
<tr>
<td>34. пад(а)-нъа</td>
<td>to write (as an activity, intr.)</td>
<td>to start writing (as an activity, intr.)</td>
<td>IPFV</td>
</tr>
<tr>
<td>35. паже-нъа</td>
<td>to have abscesses/boils appearing</td>
<td>to start having abscesses/boils appearing</td>
<td>IPFV</td>
</tr>
<tr>
<td>36. паже-иъъа</td>
<td>to have abscesses/boils appearing (esp. regularly)</td>
<td>to start having abscesses/boils appearing (esp. appearing regularly)</td>
<td>ITER</td>
</tr>
<tr>
<td>37. пан(а)-дамъа</td>
<td>to fill (intr.)</td>
<td>to start filling (intr.)</td>
<td>TRR-IPFV</td>
</tr>
<tr>
<td>38. пан(а)-нъа</td>
<td>to fill (intr.)</td>
<td>to start filling (intr.)</td>
<td>IPFV</td>
</tr>
</tbody>
</table>

(continues)
<table>
<thead>
<tr>
<th>VERB</th>
<th>MEANING: A paradigm</th>
<th>MEANING: B paradigm</th>
<th>AFFIXES</th>
</tr>
</thead>
<tbody>
<tr>
<td>p'î-nà</td>
<td>to be cooked, to be boiling</td>
<td>to start being cooked, to start being boiling</td>
<td>IPFV</td>
</tr>
<tr>
<td>39.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p'înà-ŋgà</td>
<td>to take a fright of (esp. from time to time)</td>
<td>to start taking a fright of (esp. taking from time to time)</td>
<td>IPFV-ITER</td>
</tr>
<tr>
<td>40.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p'îs(â)-la-pña</td>
<td>to chuckle, to laugh softly</td>
<td>to start chuckling, to start laughing</td>
<td>INCH-CAUS-IPFV</td>
</tr>
<tr>
<td>41.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>p'îs(â)-ggà</td>
<td>to chuckle, to laugh softly</td>
<td>to start chuckling, to start laughing softly</td>
<td>ITER</td>
</tr>
<tr>
<td>42.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sâl-nà</td>
<td>to return</td>
<td>to start returning, to be going to return</td>
<td>IPFV</td>
</tr>
<tr>
<td>43.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sâl-ra-nà</td>
<td>to return</td>
<td>to start returning, to start bursting out, to start blowing off</td>
<td>TRR-IPFV IPFV</td>
</tr>
<tr>
<td>44.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>sâra-nà</td>
<td>to burst out, to blow off</td>
<td>to start bursting out, to start blowing off</td>
<td>IPFV</td>
</tr>
<tr>
<td>45.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ser-nà</td>
<td>to get salt</td>
<td>to start getting salt</td>
<td>IPFV</td>
</tr>
<tr>
<td>46.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>târpâ-nà</td>
<td>to go out, to prepare for going out</td>
<td>to start going out, to start preparing for going out</td>
<td>IPFV</td>
</tr>
<tr>
<td>47.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tebâlîî</td>
<td>to thin out (intr.)</td>
<td>to start thinning out (intr.)</td>
<td>IPFV</td>
</tr>
<tr>
<td>48.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tîrâ-nà</td>
<td>to dry (intr.)</td>
<td>to start drying (intr.)</td>
<td>IPFV</td>
</tr>
<tr>
<td>49.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tîm-la-nà</td>
<td>to rot, to start rotting</td>
<td>to start rotting (a phase previous to the phase of A paradigm)</td>
<td>INCH-IPFV</td>
</tr>
<tr>
<td>50.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tîm-nà</td>
<td>to rot</td>
<td>to start rotting</td>
<td>IPFV</td>
</tr>
<tr>
<td>51.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tî-nà</td>
<td>to fly up, to fly off</td>
<td>to start flying up, to start flying off</td>
<td>IPFV</td>
</tr>
<tr>
<td>52.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tîr'e-nà</td>
<td>to scream</td>
<td>to start screaming</td>
<td>IPFV</td>
</tr>
<tr>
<td>53.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tî'u-nà</td>
<td>to enter</td>
<td>to start entering (e.g. many people)</td>
<td>IPFV</td>
</tr>
<tr>
<td>54.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tola-nà</td>
<td>to read</td>
<td>to start reading</td>
<td>IPFV</td>
</tr>
<tr>
<td>55.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>weja-ngâ</td>
<td>to bleed</td>
<td>to start bleeding</td>
<td>ITER</td>
</tr>
<tr>
<td>56.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>xăm'da-nâ</td>
<td>to flow</td>
<td>to start flowing, to be going to flow (of rain)</td>
<td>IPFV</td>
</tr>
<tr>
<td>57.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>xâ-da-nà</td>
<td>to scratch</td>
<td>to start scratching</td>
<td>IPFV</td>
</tr>
<tr>
<td>58.</td>
<td></td>
<td></td>
<td>(cont.)</td>
</tr>
</tbody>
</table>
## Intransitive split in Tundra Nenets

### Continued

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning: A paradigm</th>
<th>Meaning: B paradigm</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>59. xàda-pta-nà</td>
<td>to move (intr.) (because of touching)</td>
<td>to start moving (intr.) (because of touching)</td>
<td>CAUS-IPFV</td>
</tr>
<tr>
<td>60. xajo-nà</td>
<td>to stay, to remain (esp. gradually)</td>
<td>to start staying, to start remaining</td>
<td>IPFV</td>
</tr>
<tr>
<td>61. xàlta-nà</td>
<td>to be washed, to launder (intr.), to lose colour</td>
<td>to start being washed, to start laundering (intr.), to start losing colour</td>
<td>IPFV</td>
</tr>
<tr>
<td>62. xàwnû-nà/ xàwnû-dà</td>
<td>to flow</td>
<td>to start flowing, to be going to flow (of rain)</td>
<td>IPFV</td>
</tr>
<tr>
<td>63. xojsa-nà</td>
<td>to wave</td>
<td>to start waving</td>
<td>IPFV</td>
</tr>
<tr>
<td>64. xonà-ngà</td>
<td>to go to sleep (esp. regularly)</td>
<td>to prepare (intr.) for sleeping (esp. regularly)</td>
<td>ITER</td>
</tr>
<tr>
<td>65. xoni-nà</td>
<td>to sleep, to fall asleep</td>
<td>to start falling asleep</td>
<td>IPFV</td>
</tr>
</tbody>
</table>

### 7.1.3.3 A: state (atelic), B: entry into a state (telic)

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning: A paradigm</th>
<th>Meaning: B paradigm</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. jaka-ngà</td>
<td>to itch (esp. from time to time)</td>
<td>to start itching</td>
<td>ITER</td>
</tr>
<tr>
<td>2. màjã</td>
<td>to be glad</td>
<td>to become glad, to rejoice</td>
<td></td>
</tr>
<tr>
<td>3. gamlo</td>
<td>to sit</td>
<td>to sit down</td>
<td></td>
</tr>
<tr>
<td>4. pìnà</td>
<td>to be afraid</td>
<td>to start being afraid, to get frightened</td>
<td></td>
</tr>
<tr>
<td>5. pìnà-þ-o</td>
<td>to be afraid</td>
<td>to get afraid</td>
<td>CAUS-FR</td>
</tr>
<tr>
<td>6. s’ûrû-o</td>
<td>to be hidden</td>
<td>to hide (intr.)</td>
<td></td>
</tr>
<tr>
<td>7. xonô-ngà</td>
<td>to be sleepy</td>
<td>to become sleepy</td>
<td>ITER</td>
</tr>
<tr>
<td>8. xonô-r</td>
<td>to doze, to slumber</td>
<td>to start falling asleep, to start dozing</td>
<td>FR</td>
</tr>
</tbody>
</table>
### 7.3.4 A: entry into a state (telic), B: entry into first stative phase (telic)

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning: A paradigm</th>
<th>Meaning: B paradigm</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. jorim</td>
<td>to become deep (esp. completely)</td>
<td>to start becoming deep, to become deep (esp. incompletely)</td>
<td></td>
</tr>
<tr>
<td>2. peda</td>
<td>to get tired (to a greater extent than for paradigm B)</td>
<td>to get tired</td>
<td></td>
</tr>
<tr>
<td>3. tira</td>
<td>to dry up (esp. completely)</td>
<td>to start drying up, to dry up (esp. incompletely)</td>
<td></td>
</tr>
<tr>
<td>4. xàn’im</td>
<td>to become frozen, to freeze</td>
<td>to start becoming frozen, to start freezing</td>
<td></td>
</tr>
</tbody>
</table>

### 7.3.5 A: entry into a state (telic), B: entry into a process (telic)

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning: A paradigm</th>
<th>Meaning: B paradigm</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. jàdara</td>
<td>to get torn</td>
<td>to start getting torn</td>
<td></td>
</tr>
<tr>
<td>2. jàdo</td>
<td>to get torn</td>
<td>to start getting torn</td>
<td></td>
</tr>
<tr>
<td>3. pajeram</td>
<td>to get covered with abscesses/boils</td>
<td>to start getting covered with abscesses/boils</td>
<td></td>
</tr>
</tbody>
</table>

### 7.3.6 No difference in actional interpretation of A and B paradigms: state (atelic)

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. jambum</td>
<td>to get longer</td>
<td></td>
</tr>
<tr>
<td>2. paje-r</td>
<td>to get covered with abscesses/boils</td>
<td>FR</td>
</tr>
<tr>
<td>3. t’im</td>
<td>to rot, to start rotting</td>
<td></td>
</tr>
<tr>
<td>4. xàmdum</td>
<td>to spill</td>
<td></td>
</tr>
</tbody>
</table>
### 7.1.3.7 No difference in actional interpretation of A and B paradigms: process (atelic)

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>jil’e-ŋgä-nä</td>
<td>to revive</td>
<td>ITER-IPFV</td>
</tr>
<tr>
<td>jorim-da-nä</td>
<td>to deepen (intr.)</td>
<td>TRR-IPFV</td>
</tr>
<tr>
<td>jorj-em-nä</td>
<td>to deepen (intr.)</td>
<td>IPFV</td>
</tr>
<tr>
<td>latra-nä</td>
<td>to start squeezing (intr.)</td>
<td>IPFV</td>
</tr>
<tr>
<td>mäjäm-pta-nä</td>
<td>to start rejoicing</td>
<td>CAUS-IPFV</td>
</tr>
<tr>
<td>mäje-pta-nä</td>
<td>to rejoice, to start rejoicing</td>
<td>CAUS-IPFV</td>
</tr>
<tr>
<td>mal’a-nä</td>
<td>to start getting broken</td>
<td>IPFV</td>
</tr>
<tr>
<td>ṭađa-ŋgä</td>
<td>to tear (intr.)</td>
<td>ITER</td>
</tr>
<tr>
<td>ṭoxol’o-nä</td>
<td>to start swimming</td>
<td>IPFV</td>
</tr>
<tr>
<td>p’inä-mbä-nä</td>
<td>to start getting frightened of</td>
<td>DUR-IPFV</td>
</tr>
<tr>
<td>p’inä-ŋgä-nä</td>
<td>to start getting frightened of</td>
<td>ITER-IPFV</td>
</tr>
<tr>
<td>p’inä-pta-nä</td>
<td>to start getting frightened of</td>
<td>CAUS-IPFV</td>
</tr>
<tr>
<td>tänä-nä</td>
<td>to ascend</td>
<td>IPFV</td>
</tr>
<tr>
<td>t’en’ewa-pta-nä</td>
<td>to start learning</td>
<td>CAUS-IPFV</td>
</tr>
<tr>
<td>t’im-ŋgä-nä</td>
<td>to start rotting</td>
<td>ITER-IPFV</td>
</tr>
<tr>
<td>wamzej-ŋgä</td>
<td>to feel angry, to become angry</td>
<td>ITER</td>
</tr>
<tr>
<td>wandartä-nä</td>
<td>to become crooked</td>
<td>IPFV</td>
</tr>
<tr>
<td>xämnda-pta-nä</td>
<td>to flow, to spill</td>
<td>CAUS-IPFV</td>
</tr>
</tbody>
</table>

### 7.1.3.8 Reverse causal order: A: first processive phase (atelic), B: process (atelic)

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning: A paradigm</th>
<th>Meaning: B paradigm</th>
<th>Affixes</th>
</tr>
</thead>
<tbody>
<tr>
<td>loxom-da-nä</td>
<td>to be almost boiling (a phase previous to the phase of B paradigm)</td>
<td>to be almost boiling</td>
<td>TRR-IPFV</td>
</tr>
<tr>
<td>p’i-r’e-nä</td>
<td>to be cooked (a phase previous to the phase of B paradigm)</td>
<td>to be cooked</td>
<td>TRR-IPFV</td>
</tr>
</tbody>
</table>
### 7.1.3.9 Reverse causal order: A: process (atelic), B: entry into a state (telic)

<table>
<thead>
<tr>
<th>VERB</th>
<th>MEANING: A paradigm</th>
<th>MEANING: B paradigm</th>
<th>AFFIXES</th>
</tr>
</thead>
<tbody>
<tr>
<td>j̤ibje-ptór</td>
<td>to think over, to think about</td>
<td>to understand</td>
<td>CAUS-FR</td>
</tr>
<tr>
<td>n̤e/le-r̤e</td>
<td>to get married</td>
<td>to get married</td>
<td>FR</td>
</tr>
<tr>
<td>pan(à)-do-r</td>
<td>to fill</td>
<td>to get filled</td>
<td>TRR-FR</td>
</tr>
<tr>
<td>t̤i-mb̤-ur</td>
<td>to rot</td>
<td>to get rotten</td>
<td>DUR-FR</td>
</tr>
<tr>
<td>x̤en⁳-tà</td>
<td>to freeze, to start freezing</td>
<td>to get frozen</td>
<td>IPFV</td>
</tr>
</tbody>
</table>

### 7.1.3.10 Lexicalization

<table>
<thead>
<tr>
<th>VERB</th>
<th>MEANING: A paradigm</th>
<th>MEANING: B paradigm</th>
<th>AFFIXES</th>
</tr>
</thead>
<tbody>
<tr>
<td>jilxe</td>
<td>to live somewhere</td>
<td>to revive, to become alive</td>
<td></td>
</tr>
</tbody>
</table>
8

From ergative case marking to semantic case marking:
the case of historical Basque

GONTZAL ALDAI

8.1 Introduction*

In this chapter I address the question whether Basque’s case-marking pattern is ergative or semantically based, focusing on dialectal variation and diachronic change. I first put forward data from Modern Basque to show that, while the Eastern and Central dialects can be considered to have a loose ergative alignment, the Western dialects (Biscayan and Guipuzcoan) have a semantic alignment based on the contrast patient/non-patient. Then, I present historical data which show an evolution from a loose ergative case-marking system in Old Basque (16th–17th centuries) to an even less ergative system in Modern Central Basque, and up to a semantic system in Modern Western Basque (while Modern Eastern Basque has stayed practically unchanged). In other words, the evolutionary path I am proposing is the following:

(1) loose ergative alignment → semantic alignment (patientive/non-patientive)

The reanalysis of the system started in the Western area, and is being extended towards the East. Since this change involves a lexical dimension, the picture that stems from it is rather complex. From a close perspective, the Basque data illustrate a gradual progress of the ergative-to-semantic reanalysis, where not all the lexical verbs involved need undergo the change at the same time. As a matter of fact, the reanalysis had already started by the time of the earliest texts (15th–16th centuries), but it was not completed in the Western area until a few centuries later. From a more distant perspective, however, it seems that, for every stage of the language and for every area or speaker, we might still talk about rather consistent grammars, with a small number of exceptions. Regarding these exceptions, we

* This work has been partly possible thanks to a post-doctoral grant from the Department of Education, Universities and Research of the Basque Government.
need to mention that recent loans are in the main more innovative, while old
high-frequency verbs tend to be more conservative.

The order of presentation is as follows. In section 8.2 I introduce the the-
oretical assumptions in this work. First, I focus on the fundamental distinc-
tion between ergative alignment and semantic alignment (section 8.2.1). Sec-
ond, within the former, I distinguish (along the lines of Harris 1985: 125)
between strict ergative and loose ergative systems (8.2.2). Third, within seman-
tic alignment, I propose a distinction between those systems based on the
contrast agentive/non-agentive and those following a patientive/non-patientive
contrast (8.2.3).

In section 8.3 I offer the basic case-marking pattern of the canonical transitive
clause and the two types of canonical intransitive clause of Basque, namely
unaccusative and unergative. In section 8.4 I examine the issue of Basque
unergatives, including data on lexically simple unergative verbs. In section 8.5 I
elaborate on secondary criteria for the distinction between ergative systems and
semantic systems. I conclude that loose ergative systems may constitute a passage
towards semantic alignment. In section 8.6, I submit historical data, mostly
regarding Basque indirect transitive constructions, which show the evolution
outlined above.

8.2 Definitions and theoretical assumptions

I will employ the following group of sentences, which I assume as pertinent for
the classifications of ergative and semantic systems, regarding case marking:

(2a) The man ate the apple
(2b.i) The man ate fish
(2b.ii) The man ate
(2b.iii) The man followed a/the wolf
(2c) The man danced (for an hour)
(2d) The man went (to Paris)
(2e) The man fell (down)

8.2.1 Ergative case marking vs. semantic case marking

The definitions of ergative marking and semantic marking are usually applied
to canonical transitive and canonical intransitive clauses only, excluding other
‘in-between’ cases such as those in (2b) above. Thus, from the sentences in (2),
we only need to take into account sentences (2a) and (2c)–(2e).

8.2.1.1 Ergative case marking  In this type of marking, the distinction between
subjects is valence-based.

(3a) The man-erg ate the-apple-abs
(3c) The man-abs danced (for an hour)
(3e) The man-abs fell (down)
8.2.1.2 Semantic case marking  In this type of marking, the distinction between subjects is semantically based.

(4a) The man-A ate the-apple-P
(4c) The man-A danced (for an hour)
(4e) The man-P fell (down)

In conclusion, given a language which marks transitive subjects with an overt marker called ERG or A and marks transitive direct objects with a marker called ABS or P, the fundamental criterion for distinguishing whether the system is ergative or semantically based can be drawn from the behaviour of the intransitive class of unergative predicates, e.g. dance, run, work. (For more precise semantic distinctions between unaccusatives and unergatives, see section 8.2.3 below.)

(5) Fundamental criterion: marking of unergative verbs

Ergative system      Semantic system
(a) Peter-P worked yesterday     (b) Peter-A worked yesterday

8.2.2 Strict ergative case marking vs. loose ergative case marking

For a better understanding of ergativity, I will distinguish (along the lines of Harris 1985: 125), two ideal subtypes of ergative case-marking system. I will call them strict vs. loose ergative systems. Moreover, this ideal classification of ergative systems gives us the possibility of putting forward secondary criteria for the distinction between ergative systems and semantic systems. (See more discussion in section 8.5.)

8.2.2.1 Strict ergative case marking  This marks only canonical transitive subjects with the ERG marker. Therefore, it marks all intransitive subjects (including in-between cases) in the same way as canonical transitive direct objects.

(6a) The man-ERG ate the-apple-ABS  
(6b.i) The man-ABS ate fish 
(6b.ii) The man-ABS ate 
(6b.iii) The man-ABS followed a/the wolf-OBL 
(6c) The man-ABS danced (for an hour) 
(6d) The man-ABS went (to Paris) 
(6e) The man-ABS fell (down)

Note that what I am calling strict ergative case marking has the fewest subjects marked with the ERG marker.

8.2.2.2 Loose ergative case marking  This marks only canonical intransitive subjects in the same way as transitive direct objects. Therefore, it marks canonical transitive subjects and in-between cases with the ERG marker.
8.2.3 Semantic case marking: agents/non-agents vs. patients/non-patients

Within semantic case-marking systems, I also propose to distinguish a contrast which turns out to be crucial for the correct analysis of the system of (Western) Basque. This distinction is based on a twofold semantic classification of prototypical intransitive subjects. Traditionally, intransitive subjects (S) are classified as either agents (S_A) or patients (S_P). In the present proposal, intransitive subjects are classified in relation to their agentivity and in relation to their patientivity. Therefore, intransitive subjects can be either agents or non-agents and either patients or non-patients. We can thus have four possible combinations which give the following types of intransitive subject:

(8a) agents + non-patients: work, dance, run, etc.
(8b) agents + patients: go, come, etc.
(8c) non-agents + patients: die, fall, (be + adjective), etc.
(8d) non-agents + non-patients: shine, beep, roll, etc.

Based on the classification of intransitive subjects in (8), I propose to distinguish the two following types of semantic case marking:

8.2.3.1 Semantic case marking based on patientivity This system distinguishes intransitive subjects which are patients regardless of their agentivity, i.e. (8b) and (8c), from intransitive subjects which are non-patients, i.e. (8a) and (8d).\(^1\)

(9a) The man-ERG ate the-apple-ABS\(^2\)
(9b.i) The man-ERG ate fish
(9b.ii) The man-ERG ate
(9b.iii) The man-ERG/ABS followed a/the wolf-OBL
(9c) The man-ABS danced (for an hour)
(9d) The man-ABS went (to Paris)
(7e) The man-ABS fell (down)

1 Split intransitivity based on patientivity seems to be common in Western European languages, although there can still be minor differences from language to language. What is not common, in any event, is for patientivity/non-patientivity to be reflected in case marking.

2 I still use the abbreviations ERG and ABS to gloss the two core cases of this system, following traditional usage in e.g. Western Basque. Bear in mind, however, that the cases should not be considered ergative and absolutive.
8.2.3.2 Semantic case marking based on agentivity  This system distinguishes intransitive subjects which are agents regardless of their patientivity, i.e. (8a) and (8b), from intransitive subjects which are non-agents, i.e. (8c) and (8d).

(8a) The man-A ate the-apple-P
(8b) The man-A ate fish
(8c) The man-A followed a/the wolf-obl
(8d) The man-A went (to Paris)

Henceforth, I will only consider the former type of semantic alignment, namely the one based on patientivity, which is the one that is pertinent for the case marking of Basque.

8.3 The case-marking system of Modern Basque

In this section, I put forward examples illustrating the canonical transitive clause and the two types of canonical intransitive clause of Basque, namely unaccusative (i.e. patientive) and unergative (i.e. non-patientive) clauses. Basque unergatives present a number of issues that will be examined in the following section.

8.3.1 Transitive-clause marking

(9a) Peru-k sagarr-a-Ø jan du
     Peter-erg apple-det-abs eaten has
     'Peter has eaten the apple.'

8.3.2 Marking of unaccusatives (patientive intransitives)

(9b) Peru erori da (non-agentive)
     Peter.abs fallen is
     'Peter has fallen.'
(9c) Peru etxe-ra joan da (agentive)
     Peter.abs house-all gone is
     'Peter has gone home.'

8.3.3 Marking of unergatives (non-patientive intransitives)

Egin-compounds:

(10a) Peru-k dantza-Ø egin du
     Peter-erg dance-abs done has
     'Peter has danced.'
Fully-fledged simple verbs:

(11e.i) **Peru dantzatu da** (Eastern Basque)
Peter.abs danced is
‘Peter has danced.’

(11e.ii) **Peru-k dantzatu du** (Western Basque)
Peter-erg danced has
‘Peter has danced.’

Note that auxiliary selection and subject case marking are always mutually dependent in Basque. That is, an abs subject always triggers so-called intransitive auxiliaries, such as *izan* ‘to be’ (*3s da*), and conversely the intransitive auxiliaries always trigger an abs subject. On the other hand, an erg subject always triggers so-called transitive auxiliaries like *edun* ‘to have’ (*3s du*), and transitive auxiliaries always trigger erg subjects.

### 8.4 Basque unergatives and the case-marking type of Basque

#### 8.4.1 Egin compounds

Although the fundamental criterion in (5) for distinguishing ergative case marking from semantic case marking is absolutely clear, in practice things are not so straightforward. The main problem in Basque (and perhaps also in some other so-called ergative languages from Eurasia) is that most unergative verbs are actually compound verbs consisting of a light verb and a nominal, usually in the (indefinite) abs case.

In the case of Basque, the light verb that functions as the base for unergative verbs is the verb *egin* ‘to do’. Thus, we have *korrika egin* ‘to run’, *igeri egin* ‘to swim’, *dantza egin* ‘to dance’, *lan egin* ‘to work’, *hitz egin* ‘to speak’, etc. As illustrated in (11d) above, the subject of these compound verbs is always in the erg case. It is as if unergative verbs were understood as having a dummy cognate patient, represented by the nominal item in the abs case. The process of lexicalization in *egin* compounds does not appear to be in a very advanced stage, since for the most part the indefinite nominal item can behave in negative and interrogative sentences like other direct object NPs. In any event, in terms of their source of grammaticalization, *egin* compounds should be regarded as transitive.

If *egin* compounds were to be analysed as intransitive verbs, they would provide evidence for considering Basque’s case marking to be semantically based rather than ergative, as they have an erg subject. However, I think it safer to assume that the reason for their erg subject is precisely the fact that they have a transitive structure. In this latter case, we could still consider that Basque has (a rather exceptional) ergative system. (As for the diachronic side of *egin* compounds, within
historical Basque there have been no major changes regarding the behaviour of these verbs.)

8.4.2 Lexically simple (i.e. non-compound) unergative verbs

In order to determine whether Basque has in the main a semantically based or a valence-based (ergative) case-marking system, we need to look at fully-fledged unergatives, i.e. lexically simple unergative verbs. If simple unergative verbs still take an \textit{erg} subject, then Basque is certainly semantically aligned. Yet, if simple unergatives have an \textit{abs} subject, Basque is rather ergative. Unfortunately, there are not too many simple unergatives, especially in the Western dialects.

Below, I present a number of simple unergative verbs in order to see if they have \textit{erg} or \textit{abs} subjects. I indicate the case-marking patterns in, respectively, Western, Central, and Eastern Basque.\footnote{For the phenomenon at issue, I assume the following rough division of Basque dialects: Western (Biscayan and Guipuzcoan), Central (Navarrese and Labourdin), and Eastern (Lower Navarrese and Souletin) (see Figure 8.1.) Note that the Basque dialects correspond, albeit only roughly, to the political divisions.}

\begin{itemize}
\item \textbf{(12a) Non-volitional inanimate activities (of emission etc.)}
\begin{center}
\begin{tabular}{lllll}
English & Basque & W. Basque & C. Basque & E. Basque \\
\hline
flash & \textit{distiratu} & \textit{erg} & \textit{erg} & \\
shine & \textit{argitu} & \textit{erg} & \textit{erg} & \textit{erg} \\
boil & \textit{irakin} & \textit{erg} & \textit{erg} & \textit{erg/(abs)} \\
last & \textit{iraun} & \textit{erg} & \textit{erg} & \textit{erg/(abs)} \\
\end{tabular}
\end{center}
\end{itemize}

\begin{itemize}
\item \textbf{(12b) Agentive verbs of manner of motion}
\begin{center}
\begin{tabular}{lllll}
English & Basque & W. Basque & C. Basque & E. Basque \\
\hline
run & \textit{korritu} & \textit{erg} & \textit{erg} & \textit{abs} \\
dance & \textit{dantzatu} & \textit{erg} & \textit{abs} & \textit{abs} \\
jump & \textit{saltatu} & \textit{erg} & \textit{abs} & \textit{abs} \\
fly & \textit{hegatu}, \textit{hegaldatu} & \textit{erg} & \textit{abs} & \textit{abs} \\
\end{tabular}
\end{center}
\end{itemize}

\begin{itemize}
\item \textbf{(12c) Agentive verbs of human activities}
\begin{center}
\begin{tabular}{lllll}
English & Basque & W. Basque & C. Basque & E. Basque \\
\hline
fight & \textit{borrokatu} & \textit{erg} & \textit{abs} & \textit{abs} \\
enjoy, play & \textit{jostatu} & \textit{abs} & \textit{abs} & \textit{abs} \\
play, enjoy & \textit{jolastu} & \textit{erg} & \textit{abs} & \textit{abs} \\
play games & \textit{jokatu} & \textit{erg} & \textit{erg} & \textit{abs} \\
travel & \textit{bidaiatu}, \textit{biajatu} & \textit{erg} & \textit{abs/erg} & \textit{abs} \\
\end{tabular}
\end{center}
\end{itemize}
Speech verbs

<table>
<thead>
<tr>
<th>English</th>
<th>Basque</th>
<th>W. Basque</th>
<th>C. Basque</th>
<th>E. Basque</th>
</tr>
</thead>
<tbody>
<tr>
<td>speak</td>
<td>mintzatu</td>
<td>ERG</td>
<td>ABS</td>
<td>ABS</td>
</tr>
<tr>
<td>speak with</td>
<td>solastatu</td>
<td></td>
<td>ABS</td>
<td>ABS</td>
</tr>
<tr>
<td>chat</td>
<td>elekatu</td>
<td></td>
<td>ABS</td>
<td>ABS</td>
</tr>
<tr>
<td>converse</td>
<td>hizketatu</td>
<td></td>
<td>ABS</td>
<td>ABS</td>
</tr>
<tr>
<td>argue</td>
<td>iharduki</td>
<td>ERG</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

‘Meal’ verbs

<table>
<thead>
<tr>
<th>English</th>
<th>Basque</th>
<th>W. Basque</th>
<th>C. Basque</th>
<th>E. Basque</th>
</tr>
</thead>
<tbody>
<tr>
<td>have lunch</td>
<td>bazkaldu</td>
<td>ERG</td>
<td>ERG</td>
<td>ABS</td>
</tr>
<tr>
<td>have dinner</td>
<td>afaldu</td>
<td>ERG</td>
<td>ERG</td>
<td>ABS</td>
</tr>
<tr>
<td>have breakfast</td>
<td>gosaldu</td>
<td>ERG</td>
<td>ERG</td>
<td>ABS</td>
</tr>
</tbody>
</table>

The picture that emerges from this simplified list of non-compound unergatives is fairly revealing. Western Basque always has ERG subjects for simple unergative verbs, whereas Eastern Basque mostly has ABS subjects. Central Basque, in turn, presents a transitional position, albeit closer to that of Eastern Basque. Most lexical verbs in the list display ERG subjects in the west but then ABS subjects towards the east. The specific location of the ERG/ABS isogloss corresponding to each verb may vary along a continuum from the Biscayan dialect in the far west to the Souletin dialect in the far east, through Guipuzcoan, Navarrese, Labourdin, and Lower Navarrese, respectively (see Figure 8.1.) For example, the verbs hegatu ‘fly, take off’ and saltatu ‘jump’ present a rather western isogloss: we already find ABS subjects from Eastern Guipuzcoan to the east; the verb mintzatu ‘speak’ has ABS subjects from Navarrese onwards; meal verbs, in turn, only take ABS subjects from eastern Lower Navarrese to the east.

In conclusion, it is safe to consider that Western Basque has a semantic case-marking system. Patientive predicates (i.e. unaccusatives) have ABS subjects, while non-patientive predicates (i.e. unergatives), both verbal compounds and simple verbs, have ERG subjects. As for Eastern Basque, the system is exceptional in that it combines lexical discrimination of many unergatives (i.e. egin compounds) with a valence-based pattern for the rest of unergatives (i.e. simple verbs), excluding inanimate unergatives and some recent loans which may take ERG subjects. This

4 It can be idiosyncratic whether a given dialect has a particular verb. Recall also that most unergatives are not lexically simple, but rather egin compounds. Speech verbs, for instance, are not lexically simple in the Western dialects. A gap in the data, therefore, only indicates that that specific verb is not used in that area.

5 Note that the major exception to the generalization concerning ABS subjects in the simple unergatives of Eastern Basque is represented by inanimate predicates such as flash, shine, or boil (12a). It seems that inanimate unergatives are differentiated from prototypical unergatives (i.e. animates), by being marked with an ERG subject. This marking is, in principle, unexpected. Yet, it is not necessarily in contradiction with the non-patientive nature of Basque unergatives, which is independent from agentivity.
Figure 8.1. Map of the Basque Country (political territories and approximate area where Basque is currently spoken)

case marking in historical Basque

complex system is closer to a genuine ergative pattern than that of Western Basque.

8.4.3 Diachronic change

Diachronically, the system of Eastern Basque is conservative and that of Western Basque innovative. In other words, the evolution which is taking place progresses towards more ERG subjects in unergative predicates. This is an evolutionary trend towards a more semantically based system. (The change may have been aided by the presence of ERG subjects in unergative egin compounds.)

This evolution appears to have started in Western Basque at an early stage. In fact, there is little evidence of ABS subjects in simple unergatives throughout historical Biscayan. Yet it should be acknowledged that there are very few non-compound unergatives in the Western area. We can nevertheless cite the following unergative verbs which have undergone a change from ABS to ERG subject: arremetidu ‘attack’ displays ABS subjects in 16th-century Alavese (Landuchio); its synonym burrustu also has ABS subjects in 16th-century Biscayan (Refranes y Sentencias); jokatu ‘play’ is attested with ABS subjects in 17th-century Biscayan (Capanaga); dantzatu ‘dance’ shows ABS subjects in 19th-century Guipuzcoan (Lardizabal, Soroa). Currently, all these verbs take ERG subjects in Western Basque.
In historical Central Basque, there are also some attested changes. The clearest shift affects ‘meal’ verbs in Labourdin (and western Lower Navarrese). In Old Labourdin, the verb *afaldu* ‘have dinner’ took *abs* subjects, and accordingly its main auxiliary was ‘to be’; but it has *erg* subjects, and ‘to have’ as auxiliary, in Modern Labourdin. The same holds for other meal verbs such as *bazkaldud* ‘have lunch’. (In addition, it seems that a change towards *erg* subjects in other unergative verbs, e.g. the verbs *mintzatu* ‘speak’ or *dantzatu* ‘dance’, might be currently going on among the younger generations of Navarrese and Labourdin speakers.)

It can be seen that, unfortunately, there are few attested changes in the case-marking pattern of lexically simple unergatives. Nevertheless, all the changes are in the same direction: from *abs* to *erg* subjects. In what follows, I try to find further evidence in favour of an evolution from a more ergative system to a more semantic system in historical Basque. As it happens, what I called above the ‘in-between’ cases regarding transitivity, i.e. those sentences that constitute neither canonical transitive clauses nor canonical intransitive clauses, provide interesting and clear-cut evidence to consider that Old Basque (including Old Western Basque) was more ergative.

8.5 Two-place predicates with non-individuated objects (i.e. ‘in-between’ cases)

8.5.1 Strict ergativity vs. loose ergativity revisited

We have already seen that the distinction between ergative and semantic alignments is usually based on the behaviour of canonical transitive and canonical intransitive sentences. However, it turns out that many ergative systems, and certainly the most typical among them (e.g. those in the Australian, Polynesian, or Eskimo languages), present a distinctive characteristic which does not involve canonical transitive or canonical intransitive sentences but rather in-between cases. In section 8.2.2 I called this characteristic ‘strict ergativity’, and provisionally defined it as the marking of only canonical transitive subjects by means of the ergative marker, while all intransitive subjects (including in-between cases) are marked in the same way as direct objects, i.e. absolutive. A more precise definition of strict ergativity would be the following. A strict ergative system is an ergative system where prototypical transitive verbs present antipassive constructions if they have non-individuated objects. These antipassive rules (which I meant to illustrate by the examples (b) in the paradigms of (6) and (7) above; see also examples (13a), (14a), and (15a) below) may in fact include different phenomena, such as: omitted objects,\(^6\) incorporated objects, non-referential

---

\(^6\) Note that the phenomenon of object omission only refers to non-referential omitted objects, as in *The man ate at home*, and not to referential object omission, as in *The man made (it)* in pro-drop
objects, indefinite objects, or oblique objects, i.e. ‘indirect transitive constructions’ (Anderson 1977: 349).

In a strict ergative system, therefore, the agentive non-patientive subject of a prototypical transitive verb is case-marked in different ways according to whether it has an individuated object or a non-individuated object. In the former case, the subject has an ergative case marker, but in the latter it is case-marked as absolutive. I will refer to this scenario as an ‘antipassive alternation.’ The more antipassive alternations an ergative system displays, the more strict ergative I will consider it. In other words, according to the above definition of strict ergativity, an ergative system will be stricter the narrower it establishes the conditioning requirement for transitivity. If the concept of transitivity is very wide, i.e. if transitive clauses include not only canonical transitive sentences but also most or all agentive or non-patientive predicates with non-individuated objects, then the system would be loose ergative. If, on the contrary, the concept of transitivity is very narrow, including only or almost only canonical transitive sentences (which have individuated objects), then the system would be strict ergative.

We may also state that loose ergative systems have a lexically fixed transitivity marking regarding individuation/non-individuation of direct objects. That is, those lexical verbs which are prototypically transitive have ergative subjects even when they have non-individuated objects. Conversely, strict ergative systems never have a fixed transitivity marking in this respect. They always show antipassive alternations in their subject case.

8.5.2 Ergative alignment vs. semantic alignment revisited

If we focus on the choice of the subject’s case marker (whether ERG or ABS), we may say that the main difference between an ergative system and a semantic system is a matter of valence vs. semantics, respectively. Thus, in an ergative system, subject case is selected according to whether there are one or two core arguments in the sentence, whereas in a semantic system the property crucial for case selection is the semantics of the predicate (especially its subject).

Therefore, we could say that a loose ergative system lies between a strict ergative system and a semantic system. This is so because, after all, loose ergative systems mark some agentive or non-patientive clauses that are actually intransitive (albeit not canonical intransitive) by means of the ERG marker. This is an indication of a more semantically based and less valence-based system. Hence, the more agentive or non-patientive sentences a given system marks with the ERG marker, the more semantically based the system will be. On the contrary, the more of these sentences a system marks with the ABS marker, the more ergative it will be; that is, more valence-based.

languages. Not every transitive verbal class may have a non-referential object-omission alternation. (See Levin 1983: 309 for a list of semantic verbal classes which may present this alternation.)
Thus, taking into account different types of agentive or non-patientive sentence with non-individuated objects, we could establish a mini-typology of more ergative marking vs. more semantic marking, according to secondary criteria.

(13) Criterion 2: oblique and indirect objects

<table>
<thead>
<tr>
<th>Type A marking (more ergative)</th>
<th>Type B marking (less ergative, more semantic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a.i) The man-\textsc{erg} followed a/the wolf-\textsc{abs}</td>
<td>(b.i) The man-\textsc{erg} followed a/the wolf-\textsc{abs}</td>
</tr>
<tr>
<td>(a.ii) The man-\textsc{abs} followed a/the wolf-\textsc{obl}</td>
<td>(b.ii) The man-\textsc{erg} followed a/the wolf-\textsc{obl}</td>
</tr>
</tbody>
</table>

(14) Criterion 3: object omission (and incorporation)

<table>
<thead>
<tr>
<th>Type A marking (more ergative)</th>
<th>Type B marking (less ergative, more semantic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a.i) The man-\textsc{erg} ate the-apple-\textsc{abs}</td>
<td>(b.i) The man-\textsc{erg} ate the-apple-\textsc{abs}</td>
</tr>
<tr>
<td>(a.ii) The man-\textsc{abs} ate</td>
<td>(b.ii) The man-\textsc{erg} ate</td>
</tr>
</tbody>
</table>

(15) Criterion 4: generic, non-referential, partitive, or indefinite objects

<table>
<thead>
<tr>
<th>Type A marking (more ergative)</th>
<th>Type B marking (less ergative, more semantic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a.i) The man-\textsc{erg} ate the-apple-\textsc{abs}</td>
<td>(b.i) The man-\textsc{erg} ate the-apple-\textsc{abs}</td>
</tr>
<tr>
<td>(a.ii) The man-\textsc{abs} ate fish</td>
<td>(b.ii) The man-\textsc{erg} ate fish</td>
</tr>
</tbody>
</table>

It should be noted that the antipassive alternations in the type A marking pertain to a system that is certainly valence-based and not semantically based, since the same subject role (that of agent or non-patient) can present different case markers. Conversely, the non-antipassive pattern in the type B marking belongs to a system that is more semantically based and less valence-based, as the case marker of the subject is fixed with the semantics of the predicate, especially the subject.

In conclusion, we have seen that unergative verbs are crucial in the ergative alignment/semantic alignment distinction, since they are one-place predicates but their subjects have the semantic properties of prototypical transitive (i.e. two-place) predicates; namely, they are agents and/or non-patients. However, if unergative verbs fail to make an unambiguous differentiation (as in Basque), it is important not only to look at clear-cut cases of two-place predicates or one-place predicates but also to examine predicates that lie somehow in between the two extremes. These may constitute a significant clue to discerning whether the system abides mostly by semantics or mostly by number of core arguments. The most important predicates to examine are those involving agentive or non-patientive sentences with non-individuated objects. I have proposed above three criteria...
related to alternations with this kind of predicate, by means of which we can then establish whether a given system is (more) ergative or (more) semantic.

8.5.3 Direction of change: from loose ergative to semantic

In the type A pattern of strict ergative systems presented in (13a)–(15a) above, the ERG marker is strictly used for marking canonical transitive clauses (which have individuated objects). Hence, this pattern presents many antipassive alternations in the subject’s case, i.e. ERG/ABS alternations within predicates with agentive or non-patientive subjects. Consequently, the ERG marker cannot be readily identified with a marker of agentivity or non-patientivity. I predict that an extension of the ERG marker to cover unergative predicates can hardly start from strict ergative systems.

On the other hand, in the type B pattern of loose ergative systems in (13b)–(15b), non-canonical sentences with non-individuated objects, which can actually be considered intransitive, are marked as ERG. Hence, this system presents more non-antipassive patterns (ERG/ERG) in the subject’s case. That is, the subject tends to be consistently marked with the ERG marker for all individuated and non-individuated object sentences belonging to a prototypically transitive agentive verb. Consequently, the ERG marker may be more readily identified with, and reanalysed as, a marker of agentivity or non-patientivity, and not necessarily of transitivity. Thus, I predict that an extension of the ERG marker to also mark the subject of unergative verbs should occur more frequently in loose ergative systems than in strict ergative systems. In other words, I propose that loose ergative systems may constitute a passage towards semantic alignment.

8.5.4 Canonical indirect transitive constructions

There may be a number of two-place verbs in a given language which canonically present an indirect or oblique object instead of a direct object. Actually, some of these two-place verbs may never have direct-object transitive constructions, but may be lexically fixed always as ‘indirect transitive’. It is common, not only for perception verbs but also for verbs that convey a metaphorically directional action with partially affected objects (especially animate objects), to display canonical indirect transitive constructions. These may be verbs such as hit, kick, shoot, attack; grasp, seize; follow, pursuit, help, assist; call, wait for, ask for.

These verbs are pertinent for our present discussion because, if they present a canonical indirect construction, they are agentive or non-patientive verbs which have only one core argument. Therefore, even though they are not properly considered unergative verbs, these verbs are similar to unergatives in that their behaviour can contribute to distinguishing whether the system is ergative or semantically based. (Moreover, in the specific case of Basque, these verbs tend to be lexically simple, unlike proper unergatives, which tend to be compound verbs.)
Thus, if agentive two-place verbs such as follow present an ABS-obl case-marking pattern (i.e. if they have ABS subjects and obl objects), they can be considered to be similar to lowly generalized (i.e. lexical) antipassives. This type of verb, therefore, would constitute evidence for an ergative alignment analysis. Conversely, if these verbs present an erg-obl case-marking pattern, they would constitute evidence in favour of a semantic alignment analysis.

8.6 Indirect transitive constructions in historical Basque

In what follows, I present data on two-place agentive non-patientive predicates with oblique or non-individuated objects from historical Basque to show that:

1. Old Western Basque (16th–17th centuries) had a loose ergative case-marking system;
2. this system evolved to a semantic system in Modern Western Basque (19th century onwards);
3. Old Central-Eastern Basque had also a loose ergative system, although apparently more ergative than that of Western Basque;
4. this system evolved to a less ergative system, specially in Modern Central Basque.

8.6.1 Old Western Basque had a loose ergative system

8.6.1.1 Antipassive alternations regarding oblique objects The following examples show that Old Western Basque displayed antipassive alternations of the kind mentioned in Criterion 2 of (13) above. However, these antipassive alternations often implied a change in the meaning of the lexical verb. Thus, the antipassive pattern might have been partly lexicalized.

The verb aditu displayed two different valence patterns, conforming to an antipassive alternation. One of these patterns, (16a), showed ABS objects and (thus) ERG subjects. The meaning of aditu in this case was ’understand’. The other pattern (16b, c) displayed oblique (allative, ALL) objects, and the subject appeared in the ABS case. Its meaning here was ’look at, pay attention to’. Note that all the meanings of aditu involved agentive non-patientive subjects. Thus, this follows an ergative pattern rather than a semantic one. (Examples from Old Biscayan.)

(16a) [ni-k] aditu-ten dot [hori] [I-erg] hear-ing I.have [that.abs]
       ’I understand (it).’ (erg-abs)

(16b) ez nax ni aditu-ten horre-lako gauz-etara
      no am I.abs listen-ing that-like thing-all-pl
      ’I don’t pay attention to those things.’ (abs-obl)

(16c) aurre-ra adi-tzen ez da-n-a atze-rantz jauxtzen da
      forth-all look-ing no is-rel-det back-all falling is
      ’He who doesn’t look forward (abs-obl) goes backward.’
A clearer antipassive alternation is provided by the verb *oratu* ‘catch, capture, seize’. I do not see any change in the meaning of this verb. Thus, the only cause of the antipassive alternation may be the case marker of the object. *abs* objects trigger *erg* subjects (17a); oblique (dative, *dat*) objects trigger *abs* subjects (17b). This follows a straightforward ergative pattern. (Examples from 16th-century Biscayan.)

(17a)  

\[
\text{guen-ak eulli-a ora-tzen [dau]}
\]

cobweb-**erg** fly-**abs** catch-ing has

‘The spider’s web captures the fly.’ (**erg**-**abs**)

(17b)  

\[
\text{egaz ba-neki, [ni] ora ne[n]kikeo txori-ari}
\]

flying if-I.knew [I.**abs**] catch irr.intr.aux.1.3 bird-**dat**

‘If I knew how to fly, I would seize the bird.’ (**abs**-**dat**).

8.6.1.2 *ABS subjects in canonical indirect transitive constructions* Some agentive non-patientive verbs of Old Western Basque which had oblique objects always displayed an *abs* subject. Among them, we can mention *jarrai* ‘follow’ (**abs**-**dat**) or *eskatu* ‘ask for’ (**abs**-partitive). It ought to be mentioned, however, that *eskatu* could also have an *abs*-**abs** pattern. (Examples from 16th-century Biscayan.)

(18)  

\[
\text{inurri-ari harrail}
\]

ant-**dat** follow.imp.intr.2.10.3

‘Follow the ant!’ (**abs**-**dat**)

(19)  

\[
\text{gizon bizi-a bere-en eskatu doa}
\]

man smart-**abs** his-partitive demand hаб.intr.aux.3

‘The smart man claims what belongs to him.’ (**abs**-**obl**)

These verbs also conformed to an ergative rather than a semantic alignment. Yet the system seemed to follow lexicalized patterns rather than grammatical rules.

8.6.1.3 *Conclusion: Old Western Basque* Besides the rather lexicalized non-generalized nature of the antipassive patterns just presented, in Old Western Basque there were no other types of antipassive alternation apart from the one concerning oblique or indirect objects, i.e. (13a) above. That is, there are neither object omission alternations nor incorporated object alternations (13b), nor indefinite or non-referential object alternations (13c). Notice, for instance, that unergative *egiñ* compounds (see 8.4.1) constitute an example of indefinite non-referential object incorporation, and yet these verbs are treated as transitive, their subjects appearing in the **erg** case.

For all these reasons, we might conclude that Old Western Basque followed in the main an ergative alignment system, but one which should be considered loose ergative nonetheless.
8.6.2 Modern Western Basque has a semantic system

8.6.2.1 Non-alternations regarding oblique objects  Within the historical period, and regarding Criterion 2 in (13), a clear-cut change has taken place in the passage from Old Western Basque to Modern Western Basque. At least from the beginning of the 19th century onwards, the verbs previously displaying an antipassive alternation no longer show it. Quite the opposite: two-place predicates with ABS objects have an ERG subject, as expected; but, in two-place predicates with dative or oblique objects, which are rather common in Modern Western Basque, the subject still appears in the ERG case. This constitutes evidence for a semantic rather than an ergative pattern.

For example, the verb aditu ‘understand’, examined in section 8.6.1.1 above, always has ERG subjects in Modern Western Basque, either with ABS inanimate objects meaning ‘understand something’ (20a) or with DAT animate objects meaning ‘understand somebody’ (20b, c). Note that, although we again find a subtle change in meaning when the verb changes its case-marking pattern, the subject’s role remains the same, and so does the subject’s case. (Examples from 19th-century Biscayan.)

(20a) ez-tot aditu-ten euskera hori
no-I.have understand-ing Basque that.ABS
’I don’t understand that (dialect of) Basque.’ (ERG-ABS)

(20b) zu-k aditu-ko deutsazu? Eta mi-k ez?
you-ERG understand-FUT you.have.to.them? And I-ERG no?
’You will understand them (ERG-DAT), and I won’t?’

(20c) Maisu Juan, enzun ta, aditu deutsazu?
Master John, heard and understood you.have.to.him
’Master John, you have listened; have you understood him?’ (ERG-DAT)

8.6.2.2 ERG subjects in canonical indirect transitive constructions Some other agentic non-patientive two-place verbs of Modern Western Basque always show up with oblique objects, especially in the dative case. When this happens, the subject is always in the ERG case. These data constitute additional evidence in favour of considering Modern Western Basque semantically aligned rather than ergative.

Among the verbs with an ERG-DAT valence pattern we can mention oratu ‘catch’ (21a, b), jarraitu ‘follow’ (22a, b), begiratu ‘look at’ (23), and lagundu ‘help’ (24a, b). (Examples from 19th-century Biscayan.) Note that the former two verbs displayed an ABS-DAT pattern in examples (17b) and (18) of Old Western Basque above.

(21a) katu-ak daroa-n sardine-ari, oratu egio
  cat-ERG carries-REL sardine-DAT catch IMP.TR.AUX.2.IO.3
  ’(Go) catch the sardine (ERG-DAT) that the cat holds.’

Unfortunately, it is no longer easy to find the verb aditu with the meaning of ‘look at, pay attention to’. Note that it was with this meaning that aditu displayed ABS subjects in Old Western Basque.
(2b)  \[txakur-ak\] oratu-ten ba-deust?
\[dog\text{-ERG}\] catch-ing if-he.has.to.me
'What if [the dog] catches me?' (ERG-DAT)

(22a)  jangoiko-aren zigorr\-ak jarraitu-ko deutsu
God-gen punishment-ERG follow-FUT it.has.to.you
'The punishment of God will follow you.' (ERG-DAT)

(22b)  bihar jarraitu-ko deutsagu hartu dogu-n lan-ari
tomorrow follow-FUT we.have.to.it taken we.have-REL job-DAT
'Tomorrow, we will pursue the task (ERG-DAT) we have undertaken'

(23)  begira egiizu ondo atso hon-i
look IMP.TR.AUX.2.1O.3 well old.woman this-DAT
'Look thoroughly at this old woman.' (ERG-DAT)

(24a)  lagun-du ze unstan mahai-ko erregu-ak egiten
help-ed you.had.to.me table-GEN prayer-ABS.PL doing
'You helped me (ERG-DAT) saying the table prayers.'

(24b)  ez deutsazu eskatu behar jangoiko-ari lagundu
no you.have.to.him ask need God-DAT help
dagizu-la?
TR.AUX.3.1O.2-COMP
'Aren't you going to ask the Lord to help you?' (ERG-DAT)

8.6.2.3 Conclusion: Modern Western Basque In section 8.6.2 I have put forward data regarding two-place agentive non-patientive predicates with indirect (DAT) objects to show that Modern Western Basque has ERG subjects in these cases. If we add to these data the evidence provided by the handful of simple (non-compound) unergatives which also take ERG subjects (examples (12a-e) in section 8.4.2), we must conclude that Modern Western Basque has a semantically aligned system. Thus, in the (short) passage from the 16th-17th centuries to the 19th, what we might still consider a loose ergative system in Old Western Basque definitely became a semantic system in Modern Western Basque.

8.6.3 Old Central-Eastern Basque had a loose ergative system

8.6.3.1 ABS subjects in canonical indirect transitive constructions In a parallel fashion to Old Western Basque, Old Central and Eastern Basque also displayed antipassive alternations regarding oblique and indirect objects (i.e. of the kind presented in Criterion 2 of (13) above). Similarly, some agentive non-patientive verbs of Old Central-Eastern Basque which had oblique objects always had ABS subjects. Among these verbs we can mention baliatu 'protect' (25), heldu 'help' (26), othoiztu 'beg' (27), and behatu 'look at, pay attention to, understand' (28a, b). The former three verbs always had DAT objects, the latter could have DAT or all objects. (Examples from 16th-century Lower Navarrese.)
Gontzal Aldai

(25) balia zakitzat! [ni-ri]  
protect IMP.INTR.AUX.2.IO.1 [I-DAT]  
'Protect me!' (ABS-DAT)

(26) hel zakitzat! [ni-ri]  
help IMP.INTR.AUX.2.IO.1 [I-DAT]  
'Help me!' (ABS-DAT)

(27) [ni] othoitzen nitzaitzu [zu-ri]  
[LABS] begging I.am.to.you [YOU-DAT]  
'I beg [to] you.' (ABS-DAT)

As with Old Western Basque, the data under consideration appear to reflect a rather lexical non-generalized phenomenon for Old Central-Eastern Basque. In addition, it seems each verb has a given case-marking pattern for any given meaning it conveys. In any event, all the examples presented here have agentive non-patientive subjects in the ABS case.

(28a) ni erhoa, zu ikintsu; beha e-nakidizu  
I fool, you wise; look NO-INTR.AUX.1.IO.2  
'I am foolish, you are learned; I cannot understand [to] you.' (ABS-DAT)

(28b) amoros-ak nahi nuke hon-at beha ba-lite  
lover-ABS.PL desire I.would.have here-ALL look IF-INTR.AUX.3.PL  
'I would like lovers to look at/pay attention to this.' (ABS-ALL)

8.6.3.2 Antipassive (?) alternations regarding omitted objects  Besides the antipassive alternations regarding oblique objects, in Old Central-Eastern Basque (this time unlike in Old Western Basque) we may also find a few examples that may look like object-omission antipassive alternations, of the type presented in Criterion 3 of (14) above. Among the handful of verbs that may display this kind of alternation we may cite dantzatuz to dance; as in (29) below.

(29a) mutil-a-Ø dantzatu da  
boy-DET-ABS danced is  
'The boy danced.' (ABS)

(29b) mutil-a-Ø baso-jantza-Ø dantzatu du  
boy-DET-ERG glass-dance-ABS danced has  
'The boy danced the glass-dance (a specific type of dance).' (ERG-ABS)

The ERG/ABS alternation in (29) provides evidence for considering this system as ergative. However, it ought to be taken into account that there are very few verbs that follow this pattern. Most of the verbs that may be expected to display an omitted-object antipassive alternative have actually consistent ERG subjects either when they are transitive or intransitive. Among the many verbs that do not show this alternation in Old Central-Eastern Basque, we can mention jain 'eat', edan 'drink', irakurri 'read', idatzi 'write', kantiatu 'sing', iraudi 'plough', erein
Case marking in historical Basque

‘sow’, eho ‘grind’, irun ‘spin’, and josí ‘sew’. Furthermore, and perhaps more importantly, in the examples in (29), the canonical and most common use is clearly the intransitive one. Thus, this alternation does not seem to correspond to a genuine antipassive demotion, but rather to the transitivization of a canonical intransitive verb.

8.6.3.3 Conclusion: Old Central-Eastern Basque From the data presented in section 8.6.3, we can conclude that Old Central and Eastern Basque had a case-marking system which was apparently more ergative than that of Old Western Basque. Thus, in addition to antipassive alternations regarding oblique objects, which are common to all areas of Old Basque, we also find in the Central and Eastern areas a small number of alternations that look like omitted-object antipassives. This probably constitutes the most ergative system ever attested for any period or area of historical Basque.

However, if we take into consideration that (i) oblique-object alternations seemed to be partly lexically fixed, (ii) omitted-object alternations were not genuine antipassives and still appeared only in a handful of verbs, and (iii) there was no trace of antipassive alternations regarding either incorporated objects or non-referential objects, then our conclusion should be that Old Central and Eastern Basque still displayed a loose ergative case-marking system. Certainly, the system of Old Central-Eastern Basque is far from presenting a picture similar to the one we find in the typical strict ergative languages, where the presence of antipassive alternations is paramount. Notice again that unergative egin compounds would most probably have an abs subject in a typical strict ergative language, for they are constructions with non-referential (incorporated) objects. In Old Central-Eastern Basque, indeed in historical Basque as a whole, egin compounds always have erg subjects.

8.6.4 The complex case-marking system of Modern Central-Eastern Basque

In the main, the case-marking system of Central-Eastern Basque has remained quite conservative for the past 500 years (a short period of attestation, after all). Yet, although more research is needed regarding this question, the system has progressed in Central Basque towards fewer abs subjects and more erg subjects in unergative verbs. That is, it seems that the system evolves, as in Western Basque, towards a more semantically based pattern. In any event, the Souletin dialect, spoken in the far east, has kept the most conservative case-marking system, i.e. the most ergative.

8.6.4.1 More ERG subjects in canonical indirect transitive constructions For reasons that will be put forward in section 8.6.4.2, it is not easy to find clear-cut examples of specific verbs that have shifted their valence pattern from abs-obli to erg-obli in Central Basque. One of the verbs that might be cited is behatu ‘look at’ (already examined in section 8.6.3.1 for Old Central-Eastern Basque). It seems that this verb changed in Middle Central Basque (17th–18th centuries), so that it lost its
antipassive pattern regarding oblique objects and took on erg subjects. (Examples from 17th-century Labourdin.)

(30a) miserikordia-ren eztitasun-ari beha-tzen diozu
mercy-GEN sweetness-DAT look-ING you.have.to.it
'You look at the sweetness of mercy.' (erg-DAT)

(30b) beha ezazue mundu-ko ienerazino guzti-etara
look TR.AUX.2.PL/(3) world-GEN generation all-ALL
'(You-PL) Look at all the generations in the world!' (erg-OBL)

8.6.4.2 Fewer canonical indirect transitive constructions What seems to be most characteristic of the evolution in Central Basque, and what distinguishes it from the Western Basque system, is that by and large Central dialects avoided two-place predicates with dative or oblique objects, so that a fair number of these verbs took on the canonical transitive pattern erg-abs. Note that this valence pattern is neutral as to whether the system should be analysed as ergative or semantic, since canonically marked transitive sentences do not constitute competing contexts in this contrast. (Examples from 19th-century Labourdin.)

(31a) [ni-k] lagun-du izan ditut hauzo-ak behar-ordu-etan
[I-ERG] help-ed been I.have.them neighbour-ABS.PL need-time-LOC.PL
'I have (sometimes) helped my neighbours in times of necessity.' (erg-abs)

(31b) aro-a ethortzen de-n-ean, lagundu-ko zaitut
time-DET coming is-REL-LOC, help-FUT I.have.you
'When the time comes, I will help you.' (erg-abs)

(32) Jainko-ak begira zaitzala!
God-ERG protect TR.AUX.3/2
'May the Lord protect you!' (erg-abs)

(33) beha nezak ongi, eta orhoit hadi!
look IMP.TR.AUX.2/1 well and remember IMP.INTR.AUX.2
'Look at me thoroughly (erg-abs), and remember!' (erg-abs)

(34) eta segitzen ditu bi jaun hoiek
and following he.has.them two sir those.ABS
'And he follows those two men.' (erg-abs)

(35) otohizten zaitut atzartasun-ekin entzun dezazun gaur-ko solasa
begging I.have.you attention-INSTR listen TR.AUX.2/3 today-GEN talk.ABS
'I ask you (erg-abs) to listen with care to today’s talk.' (erg-abs)

8.6.4.3 More canonical unergatives taking erg subjects As mentioned in section 8.4.3, Modern Central Basque has accepted more erg subjects in a small number of prototypical unergative verbs, such as jokatu ‘play’ and especially
Case marking in historical Basque

in meal verbs (afaldu ‘have dinner’, bazkaldu ‘have lunch’, etc.), and also in recent loans. Thus, these data also constitute a piece of evidence for the proposal that Central Basque is also undergoing an evolution towards a less ergative case-marking system (and consequently towards a more semantically based system).

8.6.4.4 Conclusion: Modern Central-Eastern Basque Taking into account the evidence presented in section 8.6.4, we can say that Modern Central-Eastern Basque partly maintains a loose ergative type of case marking. In addition, we may consider the pattern of Eastern Basque to be (slightly) more ergative than that of Central Basque. At any rate, the system of both Central and Eastern Basque is definitely more ergative than that of Modern Western Basque. The following data corroborate this view: (i) Modern Central-Eastern Basque has more simple (non-compound) unergative verbs; (ii) a fair number of these simple unergatives, including the most typical unergative activities, take ABS subjects; (iii) concerning antipassive alternations, Modern Central-Eastern Basque still conserves a handful of verbs with oblique objects and ABS subjects (e.g. jarraiki ‘follow’ and jazarri ‘attack’, according to Lafitte 1962: 190); (iv) in any case, Modern Eastern Basque does not present many verbs with oblique objects but follows a canonical ERG-ABS transitive pattern most of the time (which is not in contradiction with an ergative system); (v) as for alternations regarding omitted objects, there are still a handful of verbs that may present alternations of this type, although these are not genuine antipassives.

For these reasons, we can still consider that Modern Central-Eastern Basque has basically an ergative case-marking system. However, taking into account that: (i) many unergatives have the form of verbal compounds with ERG subjects, (ii) inanimate unergatives, such as flash, shine, boil, or last, have typically ERG subjects, (iii) recent unergative loans often take ERG subjects, and (iv) there are few antipassive alternations of any type, then we should conclude that the case-marking system of Modern Central-Eastern Basque conforms at best to a loose ergative system, which in some respects lies not too far from a semantically based system.

8.7 Concluding remarks

Although it may be somewhat surprising to find a paper on Basque in a volume about semantically aligned languages, and despite the opinion of some scholars (cf. Trask 2002, Hualde & Ortiz de Urbina 2003: v), I have tried to show in the present work that Western Basque (Biscayan and western Guipuzcoan dialects) actually follows a semantically based case-marking system. Specifically, the system is based on the contrast patientive/non-patientive (for more details on the semantics of this contrast, see Aldai, to appear).
As for Central-Eastern Basque, we may consider it to have an ergative system, but one which differs in several respects from that of the typical ergative languages, and which is anyhow not too dissimilar from a semantic alignment. I have also tried to present here the scenario of evolution that led Western Basque from a loose ergative system to a semantically based system, in the hope that this evolutionary trend, as well as the typology I have used to describe it, may have some cross-linguistic generality and theoretical validity. If so, this may contribute to enhancing our knowledge concerning more detailed classifications of ergative and semantically based systems.
Part III
The Pacific
The semantics of semantic alignment in eastern Indonesia

MARIAN KLAMER

9.1 Introduction*

Over the past few decades, much research has addressed the nature of alignment systems, that is, how core syntactic functions are organized relative to each other. The major patterns of alignment are defined in relation to S (the single argument of a one-place predicate), A (the agent argument of a transitive verb), and P (the patient argument of a two-place transitive verb). Here, I consider languages with alignment systems where S is sometimes treated like a transitive ‘agent’ and sometimes like a transitive ‘patient’ (Mithun 1991: 511), depending on certain semantic features of the argument and/or its predicate. Such systems have been referred to as ‘unaccusative–unergative’ (Perlmutter 1978), ‘split intransitive’ (Merlan 1985, Van Valin 1990), ‘split S’ (Dixon 1979), ‘agentive’, ‘Agent-Patient’, ‘Stative-Active’ (Mithun 1991, Nichols 1987), and, more recently, ‘semantic alignment’ (Wichmann, this volume).

This chapter introduces the semantic alignment systems from nine lesser-known Austronesian and Papuan languages spoken in eastern Indonesia. In some semantic alignment systems, the criterial semantic feature refers to the agentive or patientive characteristics of the participant (resulting in an ‘agent/patient’ system); in others, it is the inherent aspect of the predicate as state vs. event that crucially determines the alignment (resulting in an ‘active/stative’ system); yet other systems are based on the participant’s semantics as well as inherent aspect of the predicate. Despite considerable variation in the grammatical and semantic details, most of the languages discussed here are of the former type, as we will see.

* Acknowledgements: Many thanks to René van den Berg, John Bowden, Mark Donohue, Michael Ewing, Kees Hengeveld, František Kratochvíl, David Mead, Marianne Mithun, and Søren Wichmann for valuable discussions on the analysis of the data and/or comments on previous drafts of this chapter.

1 The terms A and P extend beyond agent and patient to other roles that are treated grammatically in the same way (Blake 2001: 25).
The aim of the present chapter is twofold. First, it illustrates the observation that the grammatical patterns and the semantic parameters of semantic alignment show considerable cross-linguistic variation (Van Valin 1990, Mithun 1991) by introducing data on the alignment systems of some lesser-known languages spoken in the eastern part of Indonesia. Secondly, it presents a first synthesis of the semantic parameters that play a role in the alignment systems found in this part of the world.

In the description of the data, I distinguish between (i) the semantic features of the predicate’s participant, using the proto-Agent and proto-Patient properties introduced by Dowty (1991) to characterize this role; and (ii) the inherent aspect of the predicate, distinguishing between dynamic event predicates and non-dynamic, static ones (cf. Arkadiev, this volume).

As the first contributing property for the proto-Agent role, Dowty mentions ‘volition’—the ‘volitional involvement in the event or state’—while the first contributing property for the proto-Patient role is ‘undergoer of a change of state’ (Dowty 1991:572). In the languages surveyed below, the alignment system is primarily determined by the semantics of the predicate’s participant: in seven languages, the relevant parameter is the proto-Agent feature ‘volition’ (referring to a [+ volitional] or [−volitional] argument); in two languages, it is the proto-Patient feature ‘undergoer of change of state’. The role of inherent predicate aspect in the encoding of S in these languages turns out to be limited; it only plays a role in the alignment system of two of them; but in those languages, argument semantics plays a role as well.

Above, S was defined as the single argument of a one-place predicate, which is taken to include clauses with a nonverbal predicate. Nonverbal predicates are intrinsically stative (non-dynamic), and their argument is typically non-volitional. Apart from the obvious syntactic differences that exist between verbal and nonverbal clauses, the S of a nonverbal clause and the P of a verbal clause are semantically similar because both refer to typically non-volitional participants, and in this respect are the semantic opposites of a prototypical A. In most of the languages in the survey reported here, this semantic parallel is formally reflected: they encode the S of nonverbal clauses identical to P, and unlike A.

2 See also Donohue (2004b), who notes the existence of a number of languages with ‘head-marking split-intransitive alignment’ in eastern Indonesia, and further suggests that this alignment is an areal feature of eastern Indonesian languages. In fact, split intransitivity is proportionally as common in eastern as in western Indonesian languages (Klamer 2006), and in both regions many languages without split intransitivity also occur, which suggests that this alignment type has no special status in the grouping of languages in the region (cf. Ewing, to appear).

3 In the languages discussed here, nonverbal clauses are intransitive and lack a (two-place) copular verb.

4 Although nonverbal predicates can have a volitional argument (‘Don’t be lazy!’), volition appears to be relevant only for certain arguments of ‘adjectival’ predicates: ‘In the domain of non-verbal predication the opposition between controlled and non-controlled states of affairs seems to be relevant only in the case of adjectival predicates . . . with first order arguments’ (Hengeveld 1992:122).
The geographical location of the languages discussed is indicated on Figure 9.1. Indonesia is home to over 742 languages (SIL Ethnologue: Gordon 2005), which belong to many different language families. In this chapter, representatives of the two largest language families in eastern Indonesia are discussed: from the Austronesian (AN) family, we look at Kambera, Kedang, Taba, Larike, Selaru, and Dobel, and from the Trans New Guinea (TNG) family, we look at Klon, Abui, and Tanglapui.5 6 The location of these languages is indicated on Figure 9.1. Table 9.1 gives an alphabetical list of the languages, with their affiliation, the source(s) used, and the number identifying them on Figure 9.1.

The chapter is structured as follows. In section 9.2 I outline the criteria to diagnose a language as having a semantic alignment system, and illustrate how they are used in the analysis of the alignment system of Acehnese, the most cited Austronesian language with ‘active/stative’ alignment. In section 9.3, I present case studies of semantic alignment in nine lesser-known languages in eastern Indonesia, going from west to east on the map in Figure 9.1: Kambera (3.1), Kedang (3.2), Klon (3.3), Abui (3.4), Tanglapui (3.5), Taba (3.6), Larike (3.7), Selaru (3.8), and Dobel (3.9). In section 9.4, a summary of the semantic factors involved in the split in marking is presented, followed by a brief discussion.

9.2 Semantic alignment in the Indonesian area

In the case studies below, I describe how S, A, and P are encoded by pronouns. I will not consider lexical NPs, because languages discussed here are generally head-marking, with pronominals encoding the person, number, and (sometimes) case features of S, A, and/or P as affixes or clitics on the predicate, while the lexical NPs are generally optional adjuncts. Another reason to focus on the pronominals is that cross-linguistically, semantic alignment systems are often restricted to person markers referring to human beings, since proto-A features are more readily attributable to human beings than to inanimate objects (Mithun 1991: 536). Lexical NPs always have 3rd person referents that are often non-human, while pronominal markers on verbs for 1st and 2nd person canonically have human referents. From this perspective it thus makes sense to

---

5 For discussion and references of genetic affiliations of Austronesian languages in Eastern Indonesia, see Blust (1993) and earlier work, and for references on the affiliation of the Trans New Guinea as well as other ‘Papuan’ languages, see Foley (1986, 2000), Pawley (2005), and Ross (2005).

6 The languages in the survey presented here are from a sample of 36 languages (i) about which documentation was available and (ii) which are spoken in Indonesia and East Timor, excluding Borneo and New Guinea. That 36-language sample was collected to study the geographic distribution of Split-S patterns in this part of the archipelago, where a total number of approx. 385 languages are spoken. (For a list of the sample, see the appendix in Klamer 2006.) The sample contains languages with and without Split-S phenomena. According to the definition in section 9.2 below, 12 of the 36 sample languages have semantic alignment. All these are discussed in the present volume: 10 in the present chapter, and 2 (Tobelo and Pagu) by Holton.
Figure 9.1: Indonesian locations of languages discussed: 1 Kambera, 2 Kedang, 3 Klon, 4 Abui, 5 Tanglapui, 6 Tabu, 7 Larike, 8 Selam, 9 Dobel
study semantic alignment systems by focusing on the pronominal encoding of arguments.

In the survey reported below, a language is considered to have semantic alignment when it has an overt split in the marking of S, and when it marks an S with proto-Agent features and/or without proto-Patient features in the same way as an A, and an S with proto-Patient features and/or without proto-Agent features, in the same way as a P. In addition, I will assume that the split marking of S must be found with morphologically underived predicates. This restriction is relevant, because there are a number of Austronesian languages that have variable intransitive patterns depending on the derivational characteristics of the predicate. In such languages, we find intransitive verbs that belong to (at least) two different lexical classes (one with dynamic, 'unergative', or 'event' verbs, the other with non-dynamic, 'unaccusative' or 'state' verbs). In some of them, the semantic contrast between the lexical classes of intransitive verbs is formally expressed by the presence vs. absence of certain derivational affixes, so that it is in fact the derivative prefixes of the verbs which determine the lexical-semantic class they belong to, and (indirectly) also the interpretation of S as more 'agent'-like or 'patient'-like.

Examples of Austronesian languages which have been analysed as split-intransitive on the basis of the morphological potential of their intransitive verbs include Buru (Grimes 1991: 99), Tukang Besi (Donohue 1999a: 482–84), Timugon Murut (Brewis 2002: 42), Balinese (Arka 2003: 33–4), and Begak (Goudswaard 2005: 201). Although they possess intransitive verb classes that are semantically motivated, these languages do not have semantic alignment in the sense defined above, because the semantics of their intransitive predicates (and hence of S) is actually determined by derivational morphology. As derivational morphology (e.g. causative, applicative) interacts in important ways with the encoding of arguments (cf. Mithun 1991: 539), morphologically derived verbs should not be analysed on

---

Table 9.1. Alphabetical list of languages discussed, with affiliation and source(s)

<table>
<thead>
<tr>
<th>Language</th>
<th>Affiliation</th>
<th>Source</th>
<th>No. on Fig. 9.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abui</td>
<td>TNG</td>
<td>Kratochvíl (2007)</td>
<td>4</td>
</tr>
<tr>
<td>Dobel</td>
<td>AN</td>
<td>Hughes (2000)</td>
<td>9</td>
</tr>
<tr>
<td>Kambera</td>
<td>AN</td>
<td>Klamer (1998, 2008)</td>
<td>1</td>
</tr>
<tr>
<td>Kedang</td>
<td>AN</td>
<td>Samely (1991)</td>
<td>2</td>
</tr>
<tr>
<td>Larike</td>
<td>AN</td>
<td>Laidig and Laidig (1991), Laidig (1992)</td>
<td>7</td>
</tr>
<tr>
<td>Selaru</td>
<td>AN</td>
<td>Coward (1990)</td>
<td>8</td>
</tr>
<tr>
<td>Taba</td>
<td>AN</td>
<td>Bowden (2001)</td>
<td>6</td>
</tr>
<tr>
<td>Tanglapui</td>
<td>TNG</td>
<td>Donohue (1996b)</td>
<td>5</td>
</tr>
</tbody>
</table>
a part with underived verbs. In the survey reported here, I have therefore decided
to focus on split intransitive patterns that occur with morphologically underived
predicates only.

According to the definition given above, Acehnese (an Austronesian language
of North Sumatra: Durie 1985, 1987) is diagnosed as a language with semantic
alignment. In Acehnese transitive clauses, A is marked with a proclitic, and P
with an optional enclitic, as in (1). The encoding of S is variable. Sometimes it is
marked like A, sometimes like P, depending on the semantics of S.

One class of intransitives (referred to as ‘controlled verbs’ in Durie 1985: 63
passim) includes motion and posture verbs with an animate and controlling
argument (jak ‘go’, döng ‘stand’, beudôh ‘get up’, iem ‘be still’), verbs of bodily
activity (khêm ‘laugh/smile’, klik ‘cry’, batôk ‘cough’), verbs of speech or mental
activity (marit ‘talk’, kira ‘think’, pham ‘understand’), and some emotion verbs
(chên ‘love/feel sympathy for’, têm ‘want, like’). The S of these verbs is marked like
A because (in the terms of Durie 1985: 63) the ‘more general’, ‘natural’ semantic
characteristics of these verbs involve ‘control’ by the argument. That is, the argu-
ment of these verbs generally has the proto-Agent property of being volitional. An
illustration is (2).

(1) Gopnyan ka lon=ngieng(=geuh)  (2) Geu=jak gopnyan
s/he INCH 1sg=see=3sg  3sg=go s/he

The second class of Acehnese intransitives have an S that need not be animate,
and is always non-volitional. This class includes event and state verbs (rhët ‘fall’,
reubah ‘topple over’, jeuet ‘become’, tròh ‘happen/arrive’), verbs of emotion (ku’eh
‘envy’, seugan ‘not want to’, êk ‘like/feel inclined’), personal attributes (beuhë
‘brave’, caröng ‘clever’, gasien ‘poor’, gasa ‘rude’), and bodily and mental states of
animate arguments (sakêt ‘sick/hurting’, gatay ‘itchy’, mumang ‘confused’, dawôk
‘engrossed’) (Durie 1985: 64–6). The lack of the proto-Agent feature of volitionality
allows the S to be expressed like P:

(3) Gopnyan rhët(=geuh)  
  s/he fall(=3sg)  
  ‘S/he falls.’ (Durie 1987: 369)

8 In his description of Acehnese, Durie (1985: 63) also mentions the problem that: ‘the semantic
component of control—that of the Agent—is not always in itself a sufficient criterion [to account
for the marking of S in Acehnese]: many roots allow this semantic component to be altered by the
application of a derivate prefix....It is significant that the meaning of a derivate verb is usually
rather less general than that of its base, with more restricted connotations.’ For similar reasons, we
focus on the split marking of S with underived verbs here.

9 In the glosses of the examples cited here, I follow the original glosses of the authors as far as
possible. However, the glosses of person, number, and case of pronominals have been standardized
following the Leipzig glossing conventions. In the examples a clitic is separated from its host by [=],
an affix by [-].
The third class overlaps with the other two, and the S of these verbs is ‘fluid’: it is encoded like A when it refers to a ‘wanting’ (Durie 1985: 55) participant, i.e. a volitional one, as in (4), and like P when it refers to the ‘ultimately affected participant’ of an event (Durie 1985: 55, 56, 63), as in (5). In other words, the proto-A feature of volitionality also determines the encoding of S in this verb class.

(4) Rila ji=matê
    ready 3.(familiar)=dead
    ‘He was ready to go to his death.’
(Durie 1985: 57)

(5) ...matê(=jih)
    dead=3.(familiar)
    ‘...he died.’ (Durie 1987: 376)

Finally, the S of non-verbal predicates in Acehnese is always encoded like P (Durie 1985: 126–8), as illustrated in (6). This marks the argument of nonverbal predicates as a non-volitional entity.

(6) Urueung nyan ubê raksasa=geuh
    person that size giant=3sg
    ‘That person is as big as a giant.’ (Durie 1985: 113)

In sum, Acehnese has semantic alignment: the split marking of S depends on the semantics of the argument. From Durie’s (1985) description it is clear that the encoding is in large part based on the lexical class a verb belongs to, i.e. is largely lexically specified. Only the verbs of the third class show alignment that is entirely semantically determined. However, although the distinction between class one and two is now lexicalized, it is transparently based on the distinction volition (or control in Durie’s terms) (class one) vs. the lack of it (class two), the same distinction that still applies regularly in the alignment of the third class, so that the split-S marking found in Acehnese can still be characterized as semantic alignment.

9.3 Case studies of semantic alignment in eastern Indonesia

In this section, nine case studies of semantic alignment in eastern Indonesia are presented, going from west to east: Kambera (3.1), Kedang (3.2), Klon (3.3), Abui (3.4), Tanglapui (3.5), Tab (3.6), Larike (3.7), Selaru (3.8), and Dobel (3.9).

---

10 The third class contains many emotion verbs (cinta ‘love/favour’, galak ‘like’, beungeh ‘angry’), verbs of thought or mental activity (syök ‘suspect’, yakin ‘believe/be sincere’), ability (jœut ‘able’, keuneuk ‘likely to’), personal attributes or attitudes (horeumat ‘polite’, kaya ‘rich’, malee ‘shy’, kiyunut ‘false, treacherous’), but also aspect verbs (mulayi ‘begin’, piyôh ‘stop’), and verbs of motion (teuka ‘arrive’, ilê ‘buzz off’), and the verbs udêp ‘live’ and matê ‘die’ (Durie 1985: 66–7).

11 For example, muntah ‘vomit’ marks S like A, but can S have control on vomiting? Additional examples can be found in Durie (1985).
Marian Klamer

9.3.1 Kambera

Kambera (Klamer 1998, 2008) is spoken in the eastern part of Sumba island. In Kambera, A, S and P are expressed as obligatory clitics on the predicate,\(^{12}\) by clitics from the paradigms in (7). Full pronouns are used for emphasis and disambiguation and are not discussed here as they are not differentiated into separate paradigms according to semantic or syntactic role.

(7) Kambera pronominal clitics

<table>
<thead>
<tr>
<th></th>
<th>NOM</th>
<th>GEN</th>
<th>ACC</th>
<th>DAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>ku=</td>
<td>=nggu =ka</td>
<td>=ngga</td>
<td></td>
</tr>
<tr>
<td>2SG</td>
<td>(m)u=</td>
<td>=mu =kau</td>
<td>=nggau</td>
<td></td>
</tr>
<tr>
<td>3SG</td>
<td>na=</td>
<td>=na =ya</td>
<td>=nya</td>
<td></td>
</tr>
<tr>
<td>1PL.INCL</td>
<td>ta=</td>
<td>=nda =ta</td>
<td>=nda</td>
<td></td>
</tr>
<tr>
<td>1PL.EXCL</td>
<td>ma=</td>
<td>=ma =kama</td>
<td>=nggama</td>
<td></td>
</tr>
<tr>
<td>2PL</td>
<td>(m)i=</td>
<td>=mi =ka(m)i</td>
<td>=ngga(m)i</td>
<td></td>
</tr>
<tr>
<td>3PL</td>
<td>da=</td>
<td>=da =ha</td>
<td>=nja</td>
<td></td>
</tr>
</tbody>
</table>

In a canonical transitive clause, A is marked with a nominative and P with an accusative, as illustrated in (8):

(8) Na=palu=ka

3SG.NOM=watch=1SG.ACC

'He hit me.' (Klamer 1998: 63, 369)

In intransive clauses, the default is to mark S like A, i.e. with a nominative clitic, as illustrated in (9)–(11). As these examples show, S does not need any proto-Agent features in order to be marked like A: the argument of mbana ‘be hot/angry’ and mutung ‘burn’ are not volitional, causing, or moving arguments, nor do they have sentience.

(9) Ba na=luhu=ka weling la pindu uma…

when 3SG.NOM= leave= PFV move.from LOC door house

'When he came out of the house door…' (Klamer 1998: 205)

(10) Na=mbana na tau Java

3SG.NOM=be.hot/angry ART person Java

'The stranger is angry.' (Klamer 1998: 118)

(11) Na=mutung na una jàkà\(^{13}\) u=pajulu wàngu epi

3SG.NOM=burn ART house if 2SG.NOM=play use fire

'The house will burn down if you play with fire.' (Klamer 1998: 152)

---

\(^{12}\) Subject and direct and indirect objects are marked as enclitics with optional additional NPs; however, when objects are indefinite they are not cliticized, but expressed as NPs.

\(^{13}\) In the Kambera examples <à> = [a], and <í> = [i].
The S of nonverbal predicates in Kambera, as a typical non-volitional participant of a non-dynamic state of affairs, is marked with an accusative enclitic, as in (12) and (13).

(12) [Lai nú] =ya
    LOC there=3SG.ACC
    ‘S/he’s over there.’ (Klamer 1998: 162)

(13) [Mbapa=nggu nyungga]=ya
    husband=1SG.GEN I=3SG.ACC
    ‘He is MY husband.’ (Klamer 1998: 156)

Apart from the nonverbal contexts where S is non-volitional and obligatorily marked like P, Kambera also has fluid S marking in verbal clauses. We noted that the default in declarative sentences is to mark S like A, as in (14a), but (14b) shows that S may optionally be marked like P. In the latter sentence, S is presented as explicitly non-volitional, and out of control.

(14) a. ...hi na=hi=ma=a=ka
    and 3SG.NOM=cry= emph=mod=pfv art Mada emph=3SG
    i Mada una...

b. ...hi hi=ma=a=ya=ka
    and cry=emph=mod=3SG.ACC=pfv art Mada emph,3SG
    i Mada una...

Given the appropriate context, all Kambera intransitive verbs allow for such an optionally accusative S. All accusative Ss are interpreted as less volitional than they would canonically be expected to be. Verbs attested with an accusative S include activity verbs (pabànjar ‘chat’), directional verbs (mài ‘come (towards speaker)’), as well as verbs denoting events (meti ‘die’, hi ‘cry’), processes (kalit ‘to grow dark’), or states (hàmu ‘be good’, hangunja ‘sit idly, sit doing nothing’, haledak ‘be clear’). With predicates denoting states or processs, the accusative clitic always has an impersonal referent, referring e.g. to the weather, or to a situation. Personal arguments of such predicates cannot be marked with an accusative (Klamer 1998: 166); compare (15a, b):

(15) a. Lalu haledak=ya
   too be.clear=3SG.ACC
   ‘It’s very clear (weather).’

b. Lalu haledak=na
   too be.clear=3SG.GEN
   ‘He’s very cheerful.’ (Klamer 1998: 168)

In sum, while a Kambera S is marked like A by default, in contexts where S canonically has no proto-A properties, such as when it is the argument of a non-verbal predicate, it is marked like P. S can also be optionally marked like P, and in that case it has a less volitional interpretation.14

14 While I have focused here on describing the contrast between nominative and accusative marking of S, it should be noted that Kambera has three additional ways to mark S: see Klamer (1998 chapter 5, 2008).
9.3.2 Kedang

Kedang (Samely 1991) is an Austronesian language spoken on Lamalera, a small island east of Flores. Kedang has fluid S marking: in principle, one and the same verb allows its S to be expressed like A or like P. Lexical classes of verbs, or verbal aspect, do not play a role.

Like Kambera, Kedang has two distinct paradigms to mark P (henceforth referred to as paradigms I and II). Either paradigm may be used to express S, depending on the semantic factors discussed below. Pronominal arguments in Kedang may be free words and/or attach to the predicate as clitics. Kedang has no case marking on NPs, nor on pronouns—except for the 1sg pronoun which distinguishes S and A from P. Non-first person free pronouns differentiate A/S from P only by their position relative to the verb: S/A pronouns precede the verb, while P pronouns follow it. In (16) the Kedang pronouns are given. An enclitic P may be marked with either of the two paradigms in (17).

(16) Kedang free pronouns (cf. Samely 1991: 70–72)

<table>
<thead>
<tr>
<th></th>
<th>Paradigm I (PI)</th>
<th>Paradigm II (PII)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>ku</td>
<td>=u</td>
</tr>
<tr>
<td>2SG</td>
<td>ko</td>
<td>=o</td>
</tr>
<tr>
<td>3SG</td>
<td>i</td>
<td>=ne</td>
</tr>
<tr>
<td>1PL.INCL</td>
<td>te</td>
<td>=te</td>
</tr>
<tr>
<td>1PL.EXCL</td>
<td>ke</td>
<td>=e</td>
</tr>
<tr>
<td>2PL</td>
<td>me</td>
<td>=me</td>
</tr>
<tr>
<td>3PL</td>
<td>suo</td>
<td>=suo</td>
</tr>
</tbody>
</table>

(17) Pronominal enclitics marking P or S in Kedang (cf. Samely 1991: 70–72)

<table>
<thead>
<tr>
<th></th>
<th>Paradigm I (PI)</th>
<th>Paradigm II (PII)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>=ku</td>
<td>=u</td>
</tr>
<tr>
<td>2SG</td>
<td>=ko</td>
<td>=o</td>
</tr>
<tr>
<td>3SG</td>
<td>=i</td>
<td>=ne</td>
</tr>
<tr>
<td>1PL.INCL</td>
<td>=te</td>
<td>=te</td>
</tr>
<tr>
<td>1PL.EXCL</td>
<td>=ke</td>
<td>=e</td>
</tr>
<tr>
<td>2PL</td>
<td>=me</td>
<td>=me</td>
</tr>
<tr>
<td>3PL</td>
<td>=deq</td>
<td>=ya</td>
</tr>
</tbody>
</table>

Samely (1991: 70) lists both P marking paradigms as synonymous—both have a ‘subjective’ as well as an ‘objective’ function. Since it appears from the source that an A in Kedang is always expressed as a free pronoun (cf. (18)), I interpret

15 There is a set of 59 verbs that obligatorily take subject prefixes (S or A) (Samely 1991: 94–6). The prefixes are single consonants and attach to vowel-initial verbal stems. Such phonotactically triggered inflection is not considered here.

16 These are the unmarked pronouns. The language has other special pronoun paradigms, not considered here.

17 The examples retain Samely’s orthography, where > marks breathy vowels.
this to mean that in ‘objective’ function, pronominal enclitics encode P, and in ‘subjective’ function they encode S.

The transitive clause in (18) illustrates the alignment of A and P. The A of the verbs maqo ‘steal’ and ching ‘deny’ is 3sg nuo ‘s/he’, the P of maqo is doiq ‘money’, the P of ching an enclitic.18 (In the glosses, the numerals I and II refer to P-marking paradigm I and II).

(18)  > Ei > oroq [nuo maqo doiq]  [paq nuo ching=i]
     I    suspect s/he steal money  but s/he deny=3sg.1
     ‘I suspect he steals money but he denies it.’ (Samely 1991: 73)

Turning now to the intransitive clauses, we observe that S is marked like A in (19a), where >ei ‘I’ is a free pronoun, and precedes the verb pan ‘go’. However, S is morphologically P-like in (19b), where it is an enclitic to the predicate phrase. In such constructions, a preverbal pronoun may optionally mark S in preverbal position, as in (19c).

(19)  a.  > Ei pan > owe > ud…
     I    go DEx market
     ‘I go to the market…’ (Samely 1991: 79)

     b.  Pan > oteq=o?
         go DEx=2sg.ii
         ‘Going up, are you?’ (p. 71)

     c.  O    pan > oteq=o?
         you go DEx=2sg.ii
         ‘Going up, are you?’ [slightly more courteous than (b)] (p. 71)

The pattern in (19b) is described as ‘typical for most common, somewhat casual speech’ (Samely 1991: 71), while (19c) is presented as a polite variety of (19b). This suggests that the obligatory item is the clitic, with the additional NP optionally present for pragmatic reasons such as politeness, and/or for emphasis or disambiguation. The analysis presented here focuses on the distribution of the clitics.

Samely (1991) does not discuss the factors that determine the choice to mark S like A or like P. However, Kedang nonverbal predicates align S like P, as in (20)–(22):

(20)  Predicate is a noun:
     > A naq usun têhêq têlê: ‘kusing=ne.’
     child  small speak say  cat=3sg.11
     ‘The children say: “It’s a cat.”’ (Samely 1991: 153)

18 Samely refers to these as ‘suffixes’ (1991: 70) but since their domain of attachment is phrasal rather than morphological, I analyse them as clitics.
Marian Klamer

(21) Predicate is an adjective:
Labur koqo miteng=ne
dress POSS.EMP black=3SG.11
‘My dress is black.’ (p. 77)

(22) Predicate is a location:
Koq lumar >ote bètè wela=ne
1SG.POSS field DEIX interior=3SG.11
‘My field is up there in the interior.’ (p. 75)

As mentioned before, nonverbal predicates typically denote non-dynamic states of affairs, and their argument is typically non-volitional, and the fact that such Ss are marked like P reflects this semantic similarity.

Regarding the fluid S marking in Kedang, this might relate to the interpretation of the argument: when S is expressed like P, it has a less agentive interpretation than when it is marked like A. Thus the S, which is expressed in (23a) and (24a) by verbal enclitics, would be less agentive than the S in (23b) and (24b), which is expressed by the preverbal pronouns suo and nuo. Unfortunately the source provides no further information on the semantics of this distinction.

(23) a. Ebeng boraq bahe nape e bale=ke
   watch look.at compl then 1PL.EXCL return=1PL.EXCL.1
   ‘When we finished watching, we returned.’
   ‘After we will have finished watching, we will return.’ (Samely 1991: 91)

b. Bahe suo bale=dèq,
   then they return=PFV
   ‘Then they returned home.’ (p. 158)

(24) a. Heri, o kua kuoq=ko?
   Heri you why.2SG cry=2SG.1
   ‘Heri, why do you cry?’

b. Nuo kuoq ɔti mawang=i
   s/he cry AGT.FOC 2.harm=3SG.1
   ‘He cries because you harmed him.’

These examples also suggest a relation between the marking of S and other grammatical properties of the clause—for example, irrealis vs. realis, perfective vs. imperfective—but the scarcity of data does not allow more to be said about this. However, it is relevant to note that S=A marking (and not S=P) in Kedang is often found in combination with various kinds of aspect marker (Samely 1991: 92) that give the predicate a more telic interpretation, such as the ‘Inceptive’ dèq mè:

19 Here I follow the classification of Samely (1991: 84–7), where colour terms are included in the class of adjectives.
20 As mentioned above, in constructions where the argument is marked by an enclitic and an additional pronoun, the pronoun is optional.
Table 9.2. Verbs attested in examples in Samely (1991)

<table>
<thead>
<tr>
<th>Verbs with their S marked as PI</th>
<th>Verbs with their S marked as PII</th>
</tr>
</thead>
<tbody>
<tr>
<td>nore 'exist' (‘there are’) (84)*</td>
<td>tawe 'laugh' (90)</td>
</tr>
<tr>
<td>beq 'be here' (72)</td>
<td>pan 'go' (70, 88, 89)</td>
</tr>
<tr>
<td>bale 'return' (91)</td>
<td>hamang 'dance' (93)</td>
</tr>
<tr>
<td>bute 'sleep' (71)</td>
<td>pan &gt;oto* 'go up' (71)</td>
</tr>
<tr>
<td>bikil 'broken' (73)</td>
<td>turu 'come down' (91)</td>
</tr>
<tr>
<td>moruq 'fall' (73)</td>
<td>bèyèng 'run' (91)</td>
</tr>
<tr>
<td></td>
<td>moleng diqen 'be better' (lit. 'healthy good') (89)</td>
</tr>
<tr>
<td></td>
<td>nihon 'be light (of day)' (74)</td>
</tr>
<tr>
<td></td>
<td>mawin 'be wet' (91)</td>
</tr>
<tr>
<td>adaq &gt;alu*</td>
<td>'behave refined' (76)</td>
</tr>
<tr>
<td>mate</td>
<td>'dead' (93)</td>
</tr>
<tr>
<td>bute</td>
<td>'sleep' (73)</td>
</tr>
<tr>
<td>bikil</td>
<td>'broken' (73)</td>
</tr>
<tr>
<td>moruq</td>
<td>'fall' (73)</td>
</tr>
</tbody>
</table>

* Numbers refer to pages in the source.

b See note 17 above.

(25)  >Ei bèq pan dèq mè
       I here go INCEP
       'I am going' / 'I will be leaving now' / 'I am about to go' / 'I will go immediately'

Having addressed the marking of S like A or like P, we continue by studying more details about the marking of S like P. In Kedang, the split in P marking is reflected in a split in the marking of: S is either an enclitic from PI, e.g. =ko '2SG.I' in (24a), or from PII, e.g. =a '2SG.II' in (19b).

When is S marked with PI, and when with PII? Table 9.2 shows some illustrations of intransitive verbs found in examples throughout the sketch.21 Those in the left-hand column mark S with a pronoun from paradigm PI, those in the right-hand column mark S with a pronoun from paradigm PII. Both PI and PII occur with verbs of states, events, and processes, so that lexical aspect does not seem to determine the choice. Neither does it appear to be the case that the marking correlates strictly with certain verbal classes, since the verbs bute, bikil, and moruq occur with both PI and PII. It seems that the split relates to the dynamicity of the predicate, i.e. whether it is a state or an event. In (26), this contrast is illustrated with the verb bute 'sleep'. In the first clause the S is marked with 3SG.II =ne; in the

21 This list gives examples of which P marker is found with which verb. It is neither exhaustive nor definitive; i.e. the source does not tell us that the verbs occurring with PI cannot take PII, or vice versa.
second sentence, it is a 3SG.I = i. The contrast is explained as follows: 'bute=ne
conveys the static nature of the action described, implying that the person is either
sound asleep, or else has slept for a considerable time. Bute=i emphasizes the
dynamic side of the action, in this case that the person has not slept for long but
fell asleep only recently' (Samely 1991: 72).

(26) Nuo bute=ne, doq-doq nuo hoko=i. Eeh, bute=i watiq,
s/he sleep=3SG.II suddenly s/he get.up=3SG.I EXCLAM sleep=3SG.I again
‘He slept, (then) suddenly got up. Why, now he has fallen asleep again!’
(Samely 1991: p. 73)

In a similar way, the contrast between =ne and =i in (27) marks a difference in
dynamicity: (27a) ‘describes the state that the flashlight is presently not usable
because it is broken’, while (27b) ‘draws the listener’s attention to the actual
breaking as the cause for its present state of being unusable’ (Samely 1991: 73),
i.e. bikil gets a more dynamic event reading.

(27) a. Koq senter bikil=ne state
1SG.POSS flashlight broken=3SG.II
‘My flashlight is broken.’ (p. 73)

b. Koq senter bikil=i event
1SG.POSS flashlight broken=3SG.I
‘My flashlight got broken.’

The same distinction applies in (28). (28a) ‘stresses the result of the falling of the
coconuts: they are now lying on the ground, while (28b) focuses on the falling
as the prehistory of the present state’ (p. 73). I interpret this as (28a) describing a
non-dynamic resulting state (‘to have fallen down’), and (28b) as a dynamic event
(‘to be/have been falling down’).

(28) a. Taq muruq=ya state
coconut fall=3PL.II
‘Coconuts fell.’ (or ‘…have fallen down’)

b. Taq muruq=deq event
cocnut fall=3PL.L
‘Coconuts fell.’ (or ‘…are/have been falling down’)

In sum, S is marked like PII when the predicate indicates a (resulting) state, and
like PI when it is an event.22

To conclude, expressed as free pronouns, A is preverbal and P postverbal. S is
marked like A when it is a more agentive participant, and when it is encoded like

22 It is unclear how this alignment of S relates to the alignment of P with Paradigm I or II, though
it seems that Paradigm I is typically used to mark P in contexts where the agentive features of A are
emphasized, (the ‘Agent’ or the ‘Action’ is ‘in focus’ (Samely 1991: 81–3)), while Paradigm II is used in
unmarked contexts.
P it gets a less agentive interpretation. (This needs to be tested further on a richer set of data than is available in the source.) The pronominal enclitics follow an ergative-absolutive alignment system: they mark S and P identically, in contrast to A. Kedang has a split in the marking of P, and the encliticized S goes along in this split. As a result, an enclitic S is sometimes marked with PI and sometimes with PII. In this way, a distinction between stative or more eventive readings predicates are expressed—a classic example of an active/stative split that is marked with two distinct P paradigms in Kedang.

9.3.3 Klon

Klon (Baird 2005, to appear) is a non-Austronesian language spoken on the island of Alor, north of Timor island. A in Klon is marked as a free pronoun that occurs in preverbal position. P is expressed as a prefix or proclitic. The paradigms are given in (29). In general, the choice of which prefix paradigm marks P depends on the lexical specification of the verb. More than 50 per cent of the transitives align P with paradigm II, about 30 per cent align P with paradigm I, and about 4 per cent align P with paradigm IV.23, 24

(29) Klon free pronouns (full and reduced) and pronominal prefixes (Baird 2005: 2, 3)

<table>
<thead>
<tr>
<th></th>
<th>Free pronouns</th>
<th>PI</th>
<th>PII</th>
<th>PIV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>na(n)</td>
<td>n-</td>
<td>no-</td>
<td>ne-</td>
</tr>
<tr>
<td>2SG</td>
<td>a(n)</td>
<td>V/-Ø</td>
<td>o-</td>
<td>e-</td>
</tr>
<tr>
<td>3</td>
<td>ga(n)</td>
<td>g-</td>
<td>go-</td>
<td>ge-</td>
</tr>
<tr>
<td>1PL.INCL</td>
<td>pi</td>
<td>t-</td>
<td>to-</td>
<td>te-</td>
</tr>
<tr>
<td>1PL.EXCL</td>
<td>ngi / ni</td>
<td>ng-</td>
<td>Ngo-</td>
<td>nge-</td>
</tr>
<tr>
<td>2PL</td>
<td>igi / i</td>
<td>V-</td>
<td>ogo-</td>
<td>ege-</td>
</tr>
<tr>
<td>3PL</td>
<td>ini / i</td>
<td>ini g-</td>
<td>ini go-</td>
<td>ini ge-</td>
</tr>
</tbody>
</table>

Agreement in Klon depends to a large extent on the lexical class to which a root verb belongs. Klon has three lexical classes of intransitive root verbs: (i) verbs that mark S like A—with a free pronoun, (ii) verbs that mark S like P—with a prefix, and (iii) verbs that mark S like A or like P, depending on the agentive properties of S. The encoding of the latter type of arguments is thus semantically motivated, see below.

The first class of verbs in Klon is the one that mark S like A. This is the largest class. It contains verbs of various semantic types, including diqiri ‘to think’, hler

23 About 10 per cent of the transitives may be prefixed by a choice between to classes of prefixes—in which case the choice is motivated by the semantics of the context of use (Baird, to appear).
24 Class III of the P marking bound pronouns are not discussed in Baird’s (2005) description, since they are not verbal prefixes, but rather weakly bound clitic-like pronouns that attach to syntactic phrases (see Baird, to appear).
'cut grass', *liir 'to fly', and *mkuun 'be fat' (Baird 2005: 6). (30) and (31) illustrate that the A of *méd 'take' and the S of *waa 'go' are both marked by a free pronoun.²⁵

(30) *Biasa ni balok mé-méd iwi g-gten
Usually ipl.excl beam RED-take house RED-make
'We usually take beams to build houses.'

(31) *Nang ini hok waa nang
NEG 3.pl irr go NEG
'No, they didn’t go.' (Baird 2005: 2)

This class of intransitives marks S like A irrespective of the semantics of the argument or the verb, so that marking S like A can be considered the default pattern.

The second class of intransitive verbs is small. The S of this class is always marked with PII. The S of these verbs is a non-controlling, non-volitional participant; examples include *atak 'rather large', *egel 'tired', and *hrak 'hot'. An illustration is (32), where both P and S are marked with a prefix from class II.

(32) a. *Go-krui 3.ii-scream
'Scream at him.'

b. *Go-hrak 3.ii-hot
'He (is) hot.' (Baird 2004)

The fact that the S of stative verbs like *hrak 'hot' is marked like P has a transparent semantic motivation. However, since the first class also contains stative verbs, but the S of these verbs must be marked like A, we cannot make the generalization that marking of S like P (vs. A) always depends on the semantics of the verb or its argument. In fact, most of the marking of Ss in Klon is determined solely by the class the verb happens to belong to, just as we observed for Acehnese in section 9.2. However, Klon differs from Acehnese in that the semantic motivation for the verbal classes in Klon is much less clear than it is in Acehnese.

The third class of Klon intransitives shows a fluid split in agreement. In this class, the semantic properties of the argument do indeed determine the alignment: S is expressed like P when it is not a volitional and controlling participant, but rather an affected one. This is illustrated in (33b), where S is marked like P with a prefix from paradigm IV. In contrast to (33a), where S is marked like A with a free pronoun, S in (33b) is presented as a more affected participant. Obviously, 'being itchy' always has an argument that is somehow affected. In Klon, even an affected S like this is marked like A, following the default pattern, but the verbs of the third class in Klon allow such an S optionally to be marked like P, in order to draw specific attention to its being affected. For marking of S like P, paradigm IV is used most frequently, although there are some verbs that select paradigm I (Baird 2005: 10).

²⁵ *Biasa and balok are loans from Malay.
Semantic alignment in eastern Indonesia

(33) a. A kaak 2sg itch
b. E-kaak 2sg.iv-itch
   'You’re itchy.'  'You’re itchy (and affected).'</b. Baird 2005: 8)

To conclude, Klon has multiple ways to mark S. In most cases the marking is a fixed property of the lexical class to which the verb belongs: class one always marks S like A, class two always like P. Only the third verbal class has fluid S marking, and the split in the alignment of S in this class is motivated by the affectedness of S. If this property is rephrased in one of Dowty’s (1991) prototypical properties, this is the proto-Patient property ‘undergoer of a change of state’. Note that S need not be a volitional and controlling participant to be aligned like A, since the argument of ‘to be itchy’ in (33a) cannot be considered volitional, nor can it exercise control on the experience of being itchy. Yet it is aligned like A in terms of agreement, which is in line with the analysis that the default alignment of a Klon S is like A. Only diverging from the default pattern needs a semantic motivation in Klon.

Default alignment is also found in Klon nominal clauses, which encode their pronominal argument like A. This is illustrated in (34), where the argument is a 3rd person dual pronoun that refers to actor arguments—if a dual referent refers to an undergoer, it is marked with an additional undergoer prefix on the verb (see Baird, to appear).

(34) Ele 3.dual ool om 1.dual woman man
     'They (dual) were husband and wife.'

9.3.4 Abui

Abui (Kratochvil 2007) is a non-Austronesian language belonging to the Timor–Alor–Pantar subgroup of the Trans New Guinea family, spoken in the west-central part of Alor island. As in Klon, the A in Abui is marked by a free pronoun that precedes the verb. The forms are given in the first column of (35). An A cannot be marked with a prefix; prefixes are used to mark non-controlling/volitional participants (while controlling/volitional participants are always marked like A) (Kratochvil 2007: section 5.1). Abui has three prefix paradigms; they are also given in (35). Unlike in Klon, the choice for any one of the three P paradigms is not lexicalized but based on a set of semantic considerations that is too complex to discuss here in full. They may be summarized as follows. While all prefixes mark non-volitional participants in transitive and intransitive clauses, P.PAT marks the most prototypical patients, P.LOC marks less affected undergoers such as locations, benefactives, and purposes, and P.REC typically marks human/animate recipients or inanimate goals (see Kratochvil 2007: section 5.5, for more details).
Marian Klamer

(35) Abui pronominals (Kratochvíl 2007)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>P.PAT</th>
<th>P.LOC</th>
<th>P.REC</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>na-</td>
<td>na-</td>
<td>ne-</td>
<td>no-</td>
</tr>
<tr>
<td>2SG</td>
<td>a-</td>
<td>a-</td>
<td>e-</td>
<td>a-</td>
</tr>
<tr>
<td>3</td>
<td>ha-</td>
<td>he-</td>
<td>ho-</td>
<td></td>
</tr>
<tr>
<td>1PL.EXCL</td>
<td>ni-</td>
<td>ni-</td>
<td>ni-</td>
<td>nu-</td>
</tr>
<tr>
<td>1PL.INCL</td>
<td>pi-</td>
<td>pi-</td>
<td>pi-</td>
<td>po/pu-</td>
</tr>
<tr>
<td>2PL</td>
<td>ri</td>
<td>ri-</td>
<td>ri-</td>
<td>ro/ru-</td>
</tr>
</tbody>
</table>

An illustration of a transitive clause in Abui is (36), where A is a free pronoun, and P refers to an indefinite patient that is not marked on the verb. In (37), the patient is definite and P is prefixed to the verb.

(36) Na bataa tukong
I cut wood
'I cut wood.' (Kratochvíl 2007: section 5.3)

(37) Fani el ha-wel-i
F. before 3.PAT-pour-PPFV
'Fani washed him.' (Kratochvíl 2007: section 5.4)

Intransitive verbs with a volitional argument express this argument like A, with a free pronoun, as illustrated in (38a) and (39a). Such an S cannot be expressed with any of the prefixes, as the b. examples show.

'I swim.' 1SG.PAT-swim 1SG.LOC-swim 1SG.REC-swim

'I run.' 1SG.PAT-run 1SG.LOC-run 1SG.REC-run

Intransitive clauses with a non-volitional participant always encode it like P, whether it refers to an event, or a state. (40) illustrates the event verb yei ‘fall’ with a non-volitional/controlling argument, which is marked with the P prefix ha- (40a), and which cannot be expressed with a free pronoun (40b).26 (41)–(43) illustrate state verbs with a non-volitional argument. In (41a) the verb indicates a condition, in (42a) an attribute, and in (43a) a bodily experience. To show the parallel with transitive constructions, (41b)–(43b) present transitive clauses, each with a P that is marked with a prefix from the same paradigm as the one used in the (a) examples.

(40) a. Ha-yei b. *Ha yei
3.PAT-fall it/s/he/they fall
'It/s/he/they fall.'

26 Whether the argument in this clause can be marked with any of the other prefixes is irrelevant for the point being made here, since all of the prefixes mark non-volitional arguments (S/P), in contrast to free pronouns that mark volitional arguments (S/A/A).
Semantic alignment in eastern Indonesia

(41) a.  Na-rik (*Na rik)  
isg.pat-be.ill  
'I am ill.'

b.  Trans. with P.PAT:  
Simon  na-wel  
S.  isg.pat-bathe  
'Simon bathes me.'

(42) a.  Ne-do kul (*Na kul)  
isg.loc-hold.punct white  
'I am white.'

b.  Trans. with P.LOC:  
Simon  ne-tatet  
S.  isg.loc-stand  
'Simon waits for me.'

(43) a.  No-lila (*Na lila)  
isg.rec-hot  
'I feel hot.'

b.  Trans. with P.REC:  
Simon  no-dik  
S.  isg.rec-prick  
'Simon tickles me.'

In sum, the alignment of S in Abui depends on its semantics: when it is a volitional participant, it is marked like A; when it is non-volitional, it is marked like P. Which of the P pronouns (PAT, LOC, or REC) is selected for the marking of the non-volitional participant depends on a complex set of other semantic factors that are not relevant for the present discussion. (See Kratochvíl 2007: ch. 5).

In Abui, arguments of nonverbal predicates are typically expressed with P pronouns. This is illustrated in (44a), where the 2nd person addressee is expressed with the prefix e-‘2SG.LOC’, a P prefix on the verb do ‘hold’. Note, however, that in some contexts the argument of a nominal predicate may also be expressed as A, with a free pronoun. This is illustrated in (44b). In such contexts, the S of the nominal predicate is coreferent with the A of the following verbal clause.

(44) a.  E-do  
Ceko  he-ama  kang  
2sg.loc-hold.punct Czech 3.inal-person be.good  
'You are a Czech.'

b.  A  
Ceko  he-ama  kang, hare bir faring buuk-e  
2SG Czech 3.inal-person be.good so beer much consume-IPFV  
'You are a Czech, so you'll drink a lot of beer' [you don't drink enough now].

In general, the argument of a nominal clause in Abui is thus expressed as P, except when it is coreferent with an active, volitional participant in a verbal clause following it.

9.3.5 Tanglapui

Tanglapui is another language belonging to the Timor–Alor–Pantar subgroup of the Trans New Guinea family. It is spoken in the eastern highlands of Alor island. The data presented here are from Donohue (1996b). Tanglapui has two types of transitive verb. One type are the 'transitive non-affective' verbs. These
verbs have a P that is not adversely affected by the event denoted by the predicate. An example is the verb *di* 'see', as in (45) and (46). The paradigms to express A and P of non-affective verbs are given in (47).

(45) \[ \text{Nga-ya-di} \quad \text{1–2-see} \] 'I/we see you.' (Donohue 1966b: 103)

(46) \[ \text{Nga-O-dia} \quad \text{1–3-see} \] 'I/we see him/her/they.'

(47) Person marking on 'non-affective' verbs in Tanglapui (Donohue 1996b: 103–4)

<table>
<thead>
<tr>
<th>A</th>
<th>P</th>
<th>S (‘non-affective’ V)</th>
<th>S (‘affective’ V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>ng(a)–</td>
<td>ng(a)–</td>
<td>ng(a)–</td>
</tr>
<tr>
<td>1pl</td>
<td>ya–</td>
<td>ya–</td>
<td>i–</td>
</tr>
<tr>
<td>2</td>
<td>Ø–</td>
<td>ya–</td>
<td>ya–</td>
</tr>
<tr>
<td>3</td>
<td>Ø–</td>
<td>ya–</td>
<td>ya–</td>
</tr>
</tbody>
</table>

The other type of transitive verb comprises those whose P undergoes a change of state, or is adversely affected by the action denoted by the predicate. These verbs are referred to as 'transitive affective' verbs. An example is *baba* 'hit' in (48).

Unlike non-affective transitives, affective transitives do not always mark both A and P on the verb. In (48a), only A is marked on the verb, in (48b), only P. The pattern underlying this alternation is that the argument indexed on the verb is the one whose referent is ranked highest on the animacy hierarchy (highest: 1st person, lowest: 3rd person). Whenever an action is performed contrary to the expected direction of this hierarchy, an inverse marker (*na-*) must be used. In (48a), the Agent is 1st person, and thus highest on the hierarchy; therefore no inverse morpheme is used on when it is indexed on the verb. In (48b), however, the Agent is 3rd person, which is lower on the hierarchy than the 1st person patient, so that the highest person on the hierarchy is not the Agent. In such cases, the inverse marker must be used when this argument is indexed on the verb.

(48) a. \[ \text{Nga-baba} \quad \text{1sg-hit} \] 'I hit her/him/it.'

b. \[ \text{Nga-na-baba} \quad \text{1sg-inv-hit} \] 'He/she hit me.' (Donohue 1996b: 106)

Like the transitive verbs, Tanglapui intransitive verbs are divided into non-affective and affective verbs. The non-affective intransitives include 'most of the verbs which have been referred to in the literature as “active” … verbs' (Donohue 1996b: 101), but they also include 'non-agentive verbs' — the four examples mentioned in the source are *ve* 'go', *mitti* 'sit', *yi* 'go up', *te* 'sleep'. The S of non-affective intransitives uses the S paradigm given in (47).

(49) \[ \text{Ng-ve} \quad \text{1sg-go ‘I go.’} \] (50) \[ \text{Ya-mitti} \quad \text{2/3-sit ‘You/they sit.’} \] (Donohue 1996b: 102)
Examples of affective intransitives are mata ‘sick’, ima ‘fever’, loki ‘wet’, and tansi ‘fall’, the latter two are illustrated in (51) and (52). These verbs use a similar paradigm to the non-affective paradigm, except that 1st person number is not marked (i.e. nga- is used for 1st person singular and plural). The reason why affective intransitives are considered a separate verbal class is that the S of such verbs can only be marked on a verb with an inverse morpheme, as shown in (51)–(52).

\[(51) \text{ Nga-na-loki} \quad \text{(52) Y-a-na-tansi} \]
\[1\text{sg-INV-wet} \quad 2\text{-INV-fall} \]
\['I'm/we're wet.' \quad 'You fall.' \]

In sum, in Tanglapui, intransitive verbs with an affected argument encode S like P. They use a construction that is formally identical to the inverse construction with affective transitive verbs, where P is marked on the verb and not A, as in (48b). The S of the other intransitive verbs is non-affective and expressed like A, with a prefix, and no inverse marker on the verb.\(^27\) Assuming that it is possible to rephrase ‘affectedness’ in terms of Dowty’s (1991) proto-Patient properties, the relevant property of the affected argument in Tanglapui will be the property ‘undergoer of a change of state’—whereas (lack of) volition is not a relevant notion in the alignment found in this language.

9.3.6 Taba

Taba (Bowden 2001) is an Austronesian language spoken on Makian island, west of Halmahera in north Maluku. In Taba, A is marked with proclitics, accompanied by optional free pronouns. The forms are given in (53).

\[(53) \text{ Taba free pronouns and proclitics marking A (Bowden 2001: 189–190)} \]
\[
\begin{array}{ll}
\text{Free} & \text{Proclitic to mark A} \\
1\text{sg} & \text{yak} k= \\
2\text{sg} & \text{au} m= \\
3\text{sg} & \text{i} n= \\
1\text{pl.incl} & \text{tit} t= \\
1\text{pl.excl} & \text{am} a= \\
2\text{pl} & \text{meu} h= \\
3\text{pl} & \text{si} l= \\
\end{array}
\]

Taba has various ways to mark P, but for the present discussion only two characteristics shared by all of them are relevant: Unlike an A, P is never cross-referenced on the verb, and unlike a preverbal A, P normally follows the verb, whether the referent is definite, as in (54), or not, as in (55).

\(^{27}\) If the analysis is correct that the -na morpheme derives inverse verb forms in Tanglapui, the alignment interacts with a verbal voice form and is thus less of a canonical example of semantic alignment given the definition in section 9.2 above.
In Taba, intransitive verbs with a human argument always mark S like A, as in (56), while the argument of non-verbal predicates is always marked like P, as in (57) (Bowden 2001: 161). (If A is additionally expressed with a pronoun, this appears before the predicate, as in (58).)

(56) $N=amlih$
    $3SG=laugh$
    'She’s laughing.' (Bowden 2001: 206)

(57) $Australia \text{ si}$
    Australia they
    'They’re Australian.' (p. 139)

(58) $Si \text{ l=wom}$
    they $3PL=come$
    'They’ve come.' (p. 188)

There is a split in the marking of non-human arguments of intransitives: they are marked like A when they are ‘effectors’ and like P when they are ‘non-effectors’ (Bowden 2001: 164). An effector is the dynamic participant doing something in an event, which differs from an agent in that an effector need be neither volitional nor even animate (Bowden 2001: 106, referring to Van Valin and Wilkins 1996: 289). In (59) and (60) S has a non-human referent that is an effector, and marked like A, with a proclitic.

(59) $Motor \text{ n=han do}$
    motor.boat $3SG=go \text{ real}$
    'The motor boat has gone.' (Bowden 2001: 107)

(60) $Mat \text{ n=giat te} \text{. Karna wah Tabu ni dad-doba kaklida}$
    but $3SG=shake \text{ neg be.tall/long}$
    'But it didn’t shake. Because Makian island has hard earth.' (Bowden 2001: 407)

In (61), the non-human referents of S is not an effector, but rather the non-volitional argument of a stative predicate. Such Ss are encoded as P in Taba, postverbally with a free pronoun.28

---

28 When S is a lexical NP, it is preverbal: $Wola \text{ ne mlongan \ 'rope \ prox be.tall/long} ‘This rope is long’ (‘\ldots \text{n=mlongan}’) (Bowden 2001: 119).
Semantic alignment in eastern Indonesia

(61) Mapot i (*n=mapot)
    heavy 3SG
    'It's heavy.' (Bowden 2001: 102)

In sum, Taba encodes the human argument of intransitives always like A, and (any) argument of a nonverbal predicates always like P. Semantic alignment referring to the stative/dynamic distinction only applies in the domain of non-human arguments, when the (non-volitional) non-human S of a dynamic predicate is marked like A, and the (also non-volitional) non-human argument of a stative predicate is marked like P.29

9.3.7 Larike


(62) Larike free pronouns and pronominal affixes (Laidig and Laidig 1991: 30, 37)31

<table>
<thead>
<tr>
<th></th>
<th>Free</th>
<th>Prefix</th>
<th>Suffix</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>a’u</td>
<td>au-</td>
<td>-a’u</td>
</tr>
<tr>
<td>2SG</td>
<td>ari</td>
<td>ari-</td>
<td>-ane</td>
</tr>
<tr>
<td>3SG</td>
<td>mane</td>
<td>me-</td>
<td>-ma</td>
</tr>
<tr>
<td>3SG.NH</td>
<td>-</td>
<td>i-</td>
<td>-a (ya, wa)</td>
</tr>
<tr>
<td>1PL.EXCL</td>
<td>ami</td>
<td>ami-</td>
<td>-ami</td>
</tr>
<tr>
<td>1PL.INCL</td>
<td>ite</td>
<td>ite-</td>
<td>-ite</td>
</tr>
<tr>
<td>2PL</td>
<td>imi</td>
<td>imi-</td>
<td>-imi</td>
</tr>
<tr>
<td>3PL</td>
<td>mati</td>
<td>mati-</td>
<td>-mati</td>
</tr>
<tr>
<td>3PL.NH</td>
<td>-</td>
<td>iri-</td>
<td>-ri</td>
</tr>
</tbody>
</table>

In Larike, A is indexed on the verb by a prefix, and P by a suffix, as illustrated in (63).

(63) Ai-tuhe-ya
    2SG-cut.open-3SG.NH
    'You cut it open.' (Laidig and Laidig 1991: 33)

29 Foley (2005: 409) claims that the class of event verbs (‘unergatives’) in Taba marks S like A, while the state verbs (‘unaccusatives’) mark S like P. However, this only applies to Ss with a non-human referent, since human arguments of both state and event verbs are always encoded like A in Taba. As the semantic properties of the argument (being human or not) also play a role in the encoding, the Taba system cannot be described by referring to lexical classes of verbs alone.

30 The language Allang is another variety of the same Allang–Waksi–Larike language group. For an overview of agentive alignment in Allang and related Central Maluku languages, see Ewing (to appear).

31 Only the singular and plural forms are listed in this survey; in addition the language has dual and trial paradigms; see Laidig and Laidig (1990).
Larike has two classes of intransitive verbs: one class that marks S like A, and another that marks S like P. Most Larike intransitives belong to the verb class that marks S like A. This class includes activity verbs such as *du'i* 'crawl', *lawa* 'run', *nanu* 'swim', *pese* 'work' and motion and event verbs like *wela* 'go home', *ra'a* 'climb', *ken* 'to go', and *piku* 'to burn'. It also contains verbs expressing property concepts such as *'ata* 'be tall', *'ida* 'be big', *'ko'i* 'be small', *'nala* 'be named' (Laidig and Laidig 1991: 32, 60, 66, 88). In other words, both events and states can have an argument that is marked like A. Illustrations are (64) and (65).

(64)  

\[
\text{Ai-du'i} \\
\text{2sg-crawl} \\
\text{You are crawling.'}
\]

(65)  

\[
\text{Ai-'ida} \\
\text{2sg-be.big} \\
\text{You are big.'}
\]

Examples where S is marked like P are shown in (66) and (67), taken from Larike narratives. In both cases the referent is non-human (NH).

(66)  

\[
\text{Tanei-u hise duma hilale pe'a-ri tahi sasa} \\
duS exist house inside finish-3PL.NH not anything \\
lohana si'u. \\
little also \\
'His belongings inside the house were totally gone.' (Laidig and Laidig 1991: 69–70)
\]

(67)  

\[
\text{Mei-hete mise ma-ta dupu ao ri'a pusu-a.} \\
3SG-say mentioned 3SG-NEG build fire for hot-3SG.NH \\
'He said he won't ever again start a fire during the dry season.' \\
(lit....for [when] it’s hot) (p. 74)
\]

S is marked like P when it is non-volitional (Laidig and Laidig 1991: 32), with verbs indicating states, such as *pe'a* 'be finished', *pehe* 'be tired', or *lopo* 'be wet' (68), or bodily experiences like *duarene* 'be hungry', (pp. 32, 69). There are also event verbs like *hanahu* 'fall' with an S marked like P (p. 32), as in (69).

(68)  

\[
\text{Lopo-ne} \\
wet-2SG \\
'You are wet.'
\]

(69)  

\[
\text{Hanahu-ne} \\
fall-2SG \\
'You fell.'
\]

Since the class of verbs that marks S like P includes both event and state verbs, the Larike system cannot be described by referring to ‘dynamic’ vs. ‘stative’ verbs—both types occur with an A-like S, as well as with a P-like S. The generalization is thus that in Larike, an S marked like P will never have a volitional referent. The reverse is not true: an S lacking volition need not be marked like P.\footnote{32}

These classes are referred to as ‘unergative’ and ‘unaccusative’ verbs by Laidig and Laidig (1991: 31–2) and in Foley (2005).

\footnote{33}{This conclusion is supported by Ewing (to appear), who argues that the split in the Allang variety of Larike is broadly based on agentivity and affectedness, and differs from the one that Foley (2005).}
Semantic alignment in eastern Indonesia

The S of Larike nonverbal predicates is marked with free pronouns, and is thus neither marked like A nor like P, as the following example illustrates:

(70) A’u putri, ane ma maka-pese-ta.
    I princess you DET AGENT-work-NOMINALIZATION
    ‘I am a princess, you are the servant.’

9.3.8 Selaru

Selaru is an Austronesian language, spoken in Selaru island, in the Tanimbar archipelago between Timor and New Guinea. Its pronominal forms are given in (71).

(71) (Selaru pronominal prefixes and pronouns (Coward 1990: 14–15))

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th></th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>k(u)-</td>
<td>1sg</td>
<td>yaw</td>
</tr>
<tr>
<td>2SG</td>
<td>m(u)-</td>
<td>2sg</td>
<td>o</td>
</tr>
<tr>
<td>3SG.AN</td>
<td>i-</td>
<td>3sg</td>
<td>ia</td>
</tr>
<tr>
<td>3SG.INAN</td>
<td>k-</td>
<td>3sg</td>
<td>Ø</td>
</tr>
<tr>
<td>1PL.INCL</td>
<td>t(a)-</td>
<td>1pl</td>
<td>iti</td>
</tr>
<tr>
<td>1PL.EXCL</td>
<td>arami-</td>
<td>1pl</td>
<td>arami</td>
</tr>
<tr>
<td>2PL</td>
<td>mi-</td>
<td>2pl</td>
<td>e</td>
</tr>
<tr>
<td>3PL</td>
<td>r(a)-</td>
<td>3pl</td>
<td>sir</td>
</tr>
</tbody>
</table>

In a transitive construction, A is expressed with an obligatory prefix and P with a postverbal pronoun from the P marking paradigm. In (72), A is prefixed, and P is a (resumptive) free pronoun if following the verb.

(72) Enw-ne-ke ra-ketya i ne i-tesu36 inatw
    turtle-this-ART 3PL butcher him this 3SG-eggs lots
    ‘This turtle they are butchering here has lots of eggs.’ (Coward 1990: 80)

draws for Larike. Foley claims that the Maluku languages have two verbal subclasses, one for ‘states’ and one for ‘performed events’ (2005: 409), the former marking S like P, the latter marking S like A, and he concludes that the Maluku languages ‘lean towards’ a split that is based on the aspectual contrast between states and events (p. 426). However, since both state and event verbs mark S like P or like A in Larike (Laidig and Laidig 1991: 31), there must be other factors involved than just aspect.

34 The orthography of Selaru used here diverges from Coward’s when high vowels in the pronominal prefixes are spelled consistently as such.

35 As a rule, the C-prefix form attaches to vowel-initial verbs, and the CV-prefix to consonant-initial verbs. When the onset of the verb is simple, the high vowel of the pronominal prefix and the verbal onset metathesize (though there appear to be some exceptions to this rule). For example, i-tabahunwa ‘3SG-kill’ becomes t-i-abahunwa (Coward 1990: 53; see below). The low vowel /a/ in the 1st inclusive and 3rd plural prefix does not metathesize; in such contexts the consonantal form of the prefix is used, e.g. t-mades ‘1PL.INCL-sweat’ (and not t-m-a-aslyes) (see Coward 1990: 15).

36 In Selaru, prefix vowels are phonologically incorporated into the verb through metathesis; for expository reasons, I added morpheme boundaries in verbs with such a metathesized prefix vowel.
Intransitive verbs always mark S like A. These include actions, (73), events (74), and mental states or bodily experiences, (75).

(73) *T-karia lan*
   1pl.incl-work hard
   ‘We work hard.’ (Coward 1990: 43)

(74) a. *R-sukar*
   3pl-enter
   ‘They enter.’ (ibid., p. 27)

b. *I-maty bony-o mu-hait i…*
   3sg-dead just-tense 2sg-drag him
   ‘Once he was dead, you dragged him…’ (p. 142)

(75) a. *Ete mu-mai*
   don’t 2sg-shy
   ‘Don’t be shy.’ (p. 72)

b. *… de asu-Vre r-aka i nini i-nkol*
   and dog-pl 3pl-howl him until 3sg-tired
   ‘…and the dogs howled at him until he was tired.’ (p. 127)

The only type of predicate that encodes S like P are the nonverbal predicates; (76) illustrates a nominal predicate, (77) an adjectival one. The S in these clauses is animate; when it is inanimate it is not overtly expressed, as in (78).

(76) *Guru i*
   teacher him
   ‘He is a teacher.’

(77) *Hahy-ke lan i*
   pig-art big him
   ‘The pig is big.’ (Coward 1990: 57)

(78) *Batbatak-ke lan Ø*
   chest-art big 3sg.inan
   ‘The chest is big.’ (p. 57)

In sum, in Selaru, the S of verbal predicates is marked like A, and the S of nonverbal predicates like P. The latter predicates are typically non-dynamic, with a non-volitional argument.

9.3.9 *Dobel*

Dobel (Hughes 2000) is an Austronesian language spoken in the Aru islands, located in the southeast of the Maluku province. In transitive clauses, A and P are marked by clitics, as illustrated in (79) and (80). In (81), the clitic paradigms are given.
Semantic alignment in eastern Indonesia

(79) \( ?A=\text{dayar}=ni \) 
\[ 3PL=\text{hit}=\text{SG.AN} \]
'He is hitting him.' (Hughes 2000: 143)

(80) \( ?A=yokwa=ni \) 
\[ 1SG=\text{see}=\text{SG.AN} \]
'He sees it.' (p. 148)

(81) Pronominal clitics in Dobel (Hughes 2000: 140)

<table>
<thead>
<tr>
<th>Case</th>
<th>Pronominal Clitic</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>( ?u = / ?o = 37 )</td>
<td>yu</td>
</tr>
<tr>
<td>2SG</td>
<td>m=/mo=</td>
<td>( ?a )</td>
</tr>
<tr>
<td>3SG.AN</td>
<td>( ?a=/na= )</td>
<td>ni</td>
</tr>
<tr>
<td>3SG.INAN</td>
<td>( ?a=/na= )</td>
<td>O/V#&gt;i</td>
</tr>
<tr>
<td>1PL.INCL</td>
<td>ma=/ma=</td>
<td>( ?a\text{ma} )</td>
</tr>
<tr>
<td>1PL.EXCL</td>
<td>ta=/ta=</td>
<td>da</td>
</tr>
<tr>
<td>2PL</td>
<td>mi=/mina=</td>
<td>( ?a\text{mi} )</td>
</tr>
<tr>
<td>3PL</td>
<td>da=/da=</td>
<td>( ?a\text{di} )</td>
</tr>
</tbody>
</table>

Intransitives are divided into two classes in Dobel. One class marks S like A, with a proclitic. This class is semantically characterized as encoding events, and illustrated in (82)–(84).

(82) \( ?A=num \) 
\[ 3SG=\text{dive} \]
'He dives' (Hughes 2000: 151)

(83) \( ?A=lesi \) 
\[ 3SG=\text{raise} \]
'He raises.' (p. 151)

(84) \( ?A=bana \text{ ti} \) 
\[ 3SG=\text{leave PFV} \]
'He has left.' (p. 148)

The argument of such event predicates does not need to be an agent. For example, the non-volitional argument of 'to sink' and 'to die' is encoded like A, as in (85) and (86):

(85) \( Na=ba?arum \) 
\[ 3SG=\text{sink} \]
'He sinks.' (Hughes 2000: 142)

(86) \( \text{Tamatu s-soba}=ni \) 
\[ \text{person RED-good-3SG.AN dem 3SG=die PFV} \]
'He has left.' (p. 148)

The other class of intransitives mark S like P. This class encodes states (Hughes 2000: 153), and is illustrated in (87)–(88).

37 The allomorphy is irrelevant for the present context.
Marian Klamer

(87)  Tumatu ne soba yu?i=ni
      person dem good INTENS=3SG.AN
   'That person is very good.' (p. 143)

(88)  Ne?pan=ni
      heavy=3SG.AN
   'He is heavy.' (p. 148)

Nonverbal predicates have an argument that is encoded like P:

(89)  Nor wadi ?a?ni=ye
      coconut.tree dem 3SG.POSS=3PL
   'These coconut trees are his.' (p. 146)

Clearly, the distinction between 'dynamic' and 'stative' predicates is pervasive in the semantic alignment of this language, but the encodings of S may cross the lexical class boundaries. For example, the argument of the state verb 'to be seasick', normally encoded like P, is marked like A in (90):

(90)  Mayas?i ?a=sula ma?del
      perhaps 3SG=drunk wave
   'Perhaps she is seasick.' (p. 162)

Furthermore, event verbs, such as dopalu?i 'appear' and koytul 'dive/sink', can have an argument that is marked like P, with an enclitic, as shown in (91) and (92) (Hughes 2000: 153). Hughes notes that this unexpected encoding entails that the participant is an 'undergoes' in the event (p. 154), i.e. S is explicitly non-volitional here.

(91)  Kwoyar ne dopalu?i=ni
      dog dem appear=3SG.AN
   'That dog appeared.' (p. 154)

(92)  Yiram ne ?om=ni ?a=f-fan re,
      axe rel 1SG.CAUSE=3SG.AN 3SG=RED-fall LOC
      koytul=ni ?ona'lay
   dive/sink=3SG.AN indeed
   'The axe, which I dropped then, did indeed sink.' (p. 177)

In sum, while semantic alignment in Dobel is mainly based on the dichotomy between state and event verbs, the encodings of S do not always obey the lexical aspect patterns of state versus event verbs. The source mentions in particular that non-volitional arguments of events may be encoded like P.

9.4 Summary and discussion

In all the languages considered here, S is encoded with a dependent pronoun (affix or clitic) attached to the predicate. The majority of them also use dependent
pronouns to mark A and P (Kambera, Tanglapui, Tobelo, Larike, Dobel), two use dependent forms only to mark A (Taba, Selaru), three use dependent forms only to mark P (Kedang, Klon, Abui). In none of the languages is semantic alignment expressed with independent pronouns only. This is in line with Mithun’s (1991: 542) observation that ‘active/agentive patterns appear especially frequently in pronominal affixes within verbs’, and Siewierska’s (2004: 54–5) finding that ‘active alignment with independent pronouns is extremely rare, while with dependent pronouns, it is [more] common’. As both of these authors explain, this is no accident, since semantic alignment systems represent the grammaticalization of semantic relations between verbs and their arguments.

Some of the languages studied have a lexicon with separate classes of intransitive verbs: one class has an S that is encoded like A, another class has an S encoded like P, and a third class has an S encoded like either A or P. Examples of such languages are Klon, Tabo, and Dobel. In Klon, the semantic motivation for the verb classes is unclear; in Tabo and Dobel, the verb class distinction is based on lexical aspect: event verbs pattern differently from state verbs. Despite the existence of such verb classes, however, in Tabo and Dobel, semantic features of the verbal argument (+/−volitional, +/−undergoing a change of state, +/−human) are also relevant parameters for its encoding.

In a number of languages verb classes do not play any role in the encoding of S. Examples are Kambera, Kedang, and Selaru, where the alignment seems entirely dependent on a semantic feature of the argument.

Most of the semantic alignment patterns we observed can be described using the proto-Agent feature ‘volition’, referring to a [+volitional] or [−volitional] argument. The proto-Patient feature ‘undergoer of change of state’ is crucial in Tanglapui and Klon. In Tabo, volition is relevant only for the distinct encoding of human and non-human arguments.

Kambera, Larike, and Klon use a default encoding for S, and the default is to mark S like A. In these languages only diverging from the default has a semantic motivation: in Kambera and Larike, a [−volitional] S may be marked like P; in Klon, an S that undergoes a change of state may be so marked.

Depending on the role the semantic feature of the argument plays in the SA, the following four types of system can thus be distinguished:

(i) [+volitional] S = A, [−volitional] S = P (Kedang, Abui, Selaru, Dobel);
(ii) [+volitional] S = A, [−volitional] S = A or S = P, depending on other factors (Taba);
(iii) [+undergoes change of state] S = P, [−undergoes change of state] S = A (Tanglapui);
(iv) Default marking of S = A (Kambera, Larike, Klon). S = P when it is [−volitional], as in (i) (Kambera, Larike); or [+undergoes change of state], as in (iii) (Klon).

Table 9.3 summarizes some of the conclusions.
Marian Klamer

Table 9.3. The encoding of S like A or P, according to the proto-Agent or proto-Patient feature of the argument ([+/−VOLitional] and [+/−undergoer of Change Of State (COS)])

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Proto-Agent</td>
<td>+VOL</td>
<td>n/a</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>n/a</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A(n/hu.)</td>
</tr>
<tr>
<td></td>
<td>−VOL</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>A(hu.)</td>
<td>P</td>
<td>A</td>
<td>P</td>
<td>P(n/hu.)</td>
</tr>
<tr>
<td>Proto-Patient</td>
<td>+COS</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>P</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td></td>
<td>−COS</td>
<td>n/a</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
</tbody>
</table>

Regarding the encoding of S according to the predicate semantics, the generalization emerged that none of the languages discussed here has a semantic alignment system based solely on a distinction between dynamic and static verbs. Only in Taba and Dobel do we find that verbal semantics plays a role—but note that in Taba the split only pertains to non-human arguments, and that in Dobel the volition of the argument is also relevant in the split. A summary of the patterns of S marking according to the aspectual semantics of the verbs is given in Table 9.4. Observe that there is an asymmetry in the encoding of S of dynamic and stative verbs: in all the languages, dynamic verbs are allowed to have an S that is encoded like A (as well as like P, in most cases), while the stative verbs cannot always have such an S: in three of the languages it can only be marked like P.

Turning now to the argument of non-verbal predicates, in Kambera, Kedang, Taba, Selaru, and Dobel this argument is always encoded like P; in Abui this is the prototypical pattern. In Larike is it encoded neither like A nor like P, and in Klon it is encoded like A, the default marking of any S. This is summarized in Table 9.5.³⁸ Apart from demonstrating that predicates of different syntactic categories use different marking systems, I suggest that those patterns where the S

Table 9.4. The encoding of S like A or P according to aspectual semantics of the predicate

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Stative</td>
<td></td>
<td>A/P</td>
<td>A/P</td>
<td>A/P</td>
<td>P</td>
<td>P</td>
<td>A</td>
<td>A/P</td>
<td>A/P</td>
</tr>
<tr>
<td>Dynamic</td>
<td></td>
<td>A/P</td>
<td>A/P</td>
<td>A/P</td>
<td>A/P</td>
<td>A/P</td>
<td>P</td>
<td>A/P</td>
<td>A/P</td>
</tr>
</tbody>
</table>

³⁸ Note that Kedang does not fit this table well, because the only alignment where semantics is involved is achieved by using distinct P clitics, while the pronouns in general follow a nominative-accusative system, and the clitics an absolutive-ergative system.
Semantic alignment in eastern Indonesia

Table 9.5. The encoding of the argument of nonverbal predicates

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>argument type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>−VOL</td>
<td>nonverbal</td>
<td>P</td>
<td>P</td>
<td>A</td>
<td>P (A)</td>
<td>[no data]</td>
<td>P</td>
<td>[other]</td>
<td>P</td>
</tr>
</tbody>
</table>

of nonverbal predicates is marked like P, and unlike A, are also a formal reflection of the semantic parallel that exists between these two types of argument. Like P, the S of a nonverbal predicate is typically a non-volitional argument, and the semantic opposite of a prototypical A. (Note that S of nonverbal clauses is not a prototypical P; it does not undergo a change of state.) In other words, in most of the languages of the survey, the non-volitional character of the S of nonverbal clauses is in harmony with how it is morphosyntactically encoded: as P, unlike a typically volitional A.

In sum, the semantic parameters of alignment in the languages of eastern Indonesia show considerable variation. They refer to the semantic features of the predicate’s participant as well as to the inherent aspect of the predicate, and often it is not easy to tease the two types apart. The proto-Agent feature of ‘volitional involvement in the event or state’ plays an important role in the semantic alignment of seven languages, and the proto-Patient role ‘undergoer of change of state’ is relevant for the semantic alignment in two.

39 See note 4 above.
The rise and fall of semantic alignment in North Halmahera, Indonesia

GARY HOLTON

10.1 Introduction∗

Among the non-Austronesian languages of Eastern Indonesia one commonly finds pronominal systems which exhibit more than one morphological pattern for indexing single arguments of intransitive verbs. In such languages core arguments are semantically aligned. That is, the two patterns for marking the single argument of an intransitive verb are distinguished not only formally, but also semantically. Active or agentive arguments follow one pattern, while stative or patientive arguments follow another. As the chapters in this volume attest, such semantically aligned systems are clearly not restricted to Eastern Indonesia or even to non-Austronesian languages. However, they are common enough in the region to lead Donohue (2004b) to propose semantic alignment as an areal tendency in Eastern Indonesia. Within this region semantic alignment cuts across genetic boundaries, occurring in genetically unrelated languages and realized to differing degrees within closely related languages. This distributional pattern begs the question of how semantically aligned systems arise, and vice versa, how non-semantically aligned systems arise from originally semantically aligned ones. Evidence of both types of evolutionary development is presented here.

The North Halmaheran languages of North Maluku provide an interesting laboratory in which to examine the evolution of semantic alignment. Although

∗ Where not indicated otherwise, North Halmaheran data cited in this chapter derive from the author’s field work in Halmahera in 1995, supported in part by the Henry Luce Foundation, grant P95280F348864. The author was assisted by many people in Halmahera, but special thanks are due to Yohanis Labi, Matias Oga, Paltiel Oga and family, Bapak Guru Kukihi and family, Domingus Diba, Frans Diba, Paulina Tindagi, Tobias Tjileni, and Jason Moloku. The author also wishes to thank the participants in the 2005 workshop on the typology of stative-active languages for extremely productive comments on an earlier version of this chapter. All responsibility for any remaining errors of fact or interpretation remains with the author.
Semantic alignment in North Halmahera, Indonesia

the North Halmaheran languages form a closely related genetic subgroup within West Papuan, semantic alignment is formally realized to varying degrees among members of this group. Some North Halmaheran languages show robust formal patterns of semantic alignment based on a distinct pronoun choice for active and stative intransitive verbs; others show little or no evidence for semantic alignment. Still other languages lie somewhere in between, reflecting systems which exhibit semantic similarities with semantically aligned systems but which lack formal instantiation of semantic alignment. This intra-family variation can provide insight as to the origin and evolution of semantic alignment. Further insights can be gleaned from the wealth of historical documentation of North Halmaheran languages, dating from the late 19th century. Combined with modern documentation, historical records can provide evidence of recent changes in alignment patterns. This evidence may help us to understand whether semantic alignment is a genetic feature of West Papuan or an areal feature acquired after North Halmaheran migration. More broadly, knowledge of the evolution of semantic alignment within North Halmaheran may contribute to a better understanding of the semantic motivations for semantically aligned systems and the diachronic pathways by which these systems arise.

In this brief survey I suggest that the formal realization of grammatical relations vis-à-vis the alignment of semantico-syntactic macro-roles may be less relevant to understanding semantic alignment than the underlying categorization of intransitive predicates into active and stative classes. The formal realization of semantic alignment in terms of the assignment of distinct pronominal prefixes for active and stative intransitive verbs is likely to be a recent feature of North Halmaheran languages. Indeed, some of the North Halmaheran languages are already morphologically impoverished to the point that this type of formal semantic alignment cannot be realized. I argue here that the formal realization of semantic alignment seen in some modern North Halmaheran languages is an epiphenomenon resulting from: (i) our pre-theoretical insistence on using macro-roles to analyse grammatical relations; and (ii) an underlying semantic categorization of verbs based on lexical aspect which pervades the North Halmaheran languages.

The North Halmaheran family comprises some ten languages spoken on the northern and eastern peninsulas of Halmahera, North Maluku, Indonesia, and some of the surrounding islands: Ternate, Tidore, Sahu, West Makian, Tobelo, Galela, Tabaru, Modole, Loloda, and Pagu. These languages form a closely related group first recognized by Robide van der Aa (1872) (see Figure 10.1). The pronominal prefix systems in North Halmaheran languages have been traditionally characterized as having nominative-accusative alignment (cf. van der Veen 1915). However, many North Halmaheran languages contain a distinct subclass of 'stative'

1 Semantically aligned systems are also reported in the non-Austronesian Yapen languages, which may be much more closely related to North Halmaheran than has been thought previously (Donohue 2005c).
or ‘objective’ intransitive verbs which index their single argument via the ‘object’ paradigm. In at least some North Halmaheran languages there is no syntactic evidence on which to distinguish a formal object relation which aligns P as opposed to S and A roles. This is perhaps most clear in Galela, in which both A and S_A roles are aligned in opposition to the P and S_P roles. Thus, the first person singular to- is found with both transitive to-ni-doto ‘I teach you’ and intransitive to-tagai ‘I go’; while the form i- is found with both transitive no-i-doto ‘you teach me’ and intransitive i-bole ‘I am tired’. By this evidence Galela pronominal prefixes provide a clear example of semantic alignment in North Halmaheran languages.

From a functional perspective other North Halmaheran languages can also be argued to be semantically aligned, even though they lack the clear morphological evidence of semantic alignment evidenced by the Galela examples above. For example, Tobelo, though superficially accusative displays characteristic properties of semantically aligned systems (Holton 2003: 55ff.). The difference has to due with the structure of the person-marking paradigms in each language, as discussed in the following section. Furthermore, even in Galela the picture is not quite as clear as the previous examples would indicate, since the occurrence of person-marking prefixes is governed by pragmatic factors related to topicality (Shelden 1986). Other languages exhibit similar variation in constraints on the occurrence of pronominal prefixes which necessarily complicate the interpretation of alignment patterns encoded by pronominal prefixes.

These differences in the morphosyntactic constraints on pronominal prefixes lead to huge differences in the formal realization of grammatical relations. In some North Halmaheran languages the pronominal prefixes can be argued to be semantically aligned: stative intransitive verbs cross-reference their single argument via an undergoer pronominal prefix, while active intransitive verbs cross-reference their single argument via an actor pronominal prefix. In other North Halmaheran languages actor prefixes must be present even with stative verbs, so that such languages reflect a formally nominative-accusative system. Nevertheless, almost all of the North Halmaheran languages exhibit a clear semantic distinction between active and stative verbs, and it is precisely this distinction which underlies the Galela ‘split’ within the intransitive verbs. That is not to say...
that all intransitive verb roots may be categorized as either active or stative; in fact, many verb roots may occur as either active intransitive or stative intransitive (or indeed transitive) forms, with corresponding pronominal morphology. Moreover, while stative-active distinction such as that in Galela is found in many North Halmaheran languages, the precise semantic criteria on which the split is based vary significantly from language to language. In spite of the close correspondence between lexical forms, active intransitive verbs in one language may be stative in another.

The remainder of this chapter is organized as follows. Section 10.2 describes the distribution of pronominal prefixes in the North Halmaheran languages and the formal realization of alignment patterns. Section 10.3 examines the constraints governing the occurrence of pronominal prefixes. The underlying semantics governing the distinction between active and stative verbs are discussed in section 10.4, while section 10.5 discusses possible pathways for semantic realignment through the reanalysis of these semantic distinctions. Finally, section 10.6 presents a possible pathway by which semantic alignment may have arisen—and eventually faded—in North Halmaheran languages.

10.2 Pronominal prefixes

All of the North Halmaheran languages exhibit some form of a person-marking prefix system which cross-references nominal arguments on verbs. However, a major distinction can be made between the person-marking system in the ‘insular’ languages Ternate, Tidore, and West Makian and that in the ‘mainland’ languages Sahu, Tobelo, Galela, Tabaru, Modole, Loloda, and Pagu.\(^2\) The mainland languages contain two distinct paradigms of verbal person-marking prefixes, which I refer to here as ‘actor’ and ‘undergoer’. The insular languages exhibit only the actor set.\(^3\) When both prefixes are present on the verb, the actor precedes the undergoer according to the basic verb template shown in Figure 10.2. While the basic verb template is the same across the North Halmaheran languages, the constraints on the occurrence of pronominal prefixes vary across the languages. First, as noted above, some North Halmaheran languages lack undergoer prefixes entirely. Even among those which do have undergoer prefixes, only some languages permit undergoer prefixes to be used with semantically intransitive (i.e. ‘stative’) verbs. Finally, those languages which index stative verbs via the undergoer paradigm may be in turn divided into those in which stative verbs are formally intransitive and those in which stative verbs are formally transitive.

\(^2\) Ternate, Tidore, and West Makian communities do in fact exist on the Halmaheran mainland, though these generally represent migrant communities. Similarly, significant Tobelo and Galela communities exist on islands surrounding Halmahera. The terms ‘insular’ and ‘mainland’ are used here as heuristic labels rather than precise geographic descriptions.

\(^3\) The terms ‘actor’ and ‘undergoer’ are used here as purely structural labels to distinguish two distinct paradigms of person-marking prefixes.
The latter employ a pleonastic actor prefix to mark the presence of an ‘experiencer’ object (cf. contributions by Malchukov and Mithun to this volume for typological discussion).

On the basis of the distribution of actor and undergoer prefixes it is thus possible to recognize four types of North Halmaheran pronominal prefix systems, corresponding to a primary division between those languages without undergoer pronominal prefixes and those with undergoer pronominal prefixes, as in Table 10.1. This table lists examples of languages for each of the four categories in this typology; but not all languages fit so neatly into these categories. For example, Tabaru verbs sometimes behave like Galela (category B.b.ii) and sometimes like Tobelo (category B.b.i).

10.2.1 Argument status

The North Halmaheran actor and undergoer person-marking prefixes are pronominal arguments, not cross-reference prefixes. In particular, these prefixes are bound forms which cannot occur independently. They function as pronouns so that a single verb may stand alone as a complete utterance, as in the following Tobelo example.

(1) Tobelo

\[
\text{no-mi-hi-honenge-oka} \\
2A-3SG\text{.FEM} \text{.U-CAUS-die-PERF} \\
\text{‘You have killed her.’ (Hueting 1936: 329)}
\]

While they may co-occur with full nominal arguments, person-marking prefixes do not co-occur with independent pronouns except with a contrastive or emphatic

Table 10.1. Typology of North Halmaheran pronominal prefix systems

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.</td>
<td>Languages without an undergoer pronominal prefix (Tidore, W. Makian)</td>
</tr>
<tr>
<td>B.</td>
<td>Languages with an undergoer prefix:</td>
</tr>
<tr>
<td>a.</td>
<td>those which index stative verbs via actor paradigm (Sahu);</td>
</tr>
<tr>
<td>b.</td>
<td>those which index stative verbs via undergoer paradigm:</td>
</tr>
<tr>
<td>i.</td>
<td>stative verbs formally transitive (Tobelo);</td>
</tr>
<tr>
<td>ii.</td>
<td>stative verbs formally intransitive (Galela).</td>
</tr>
</tbody>
</table>
reading. The Tobelo example in (2) shows a typical contrastive usage. The third singular masculine independent pronoun *una* in the second line contrasts that referent with the one in the first line.

(2) Tobelo

```
i-mi-bole
ma mi-ruba-ua
`we were tired but we didn’t trip,’
ma una i-wi-timono jadi i-wi-bole de wo-ruba
but 3M.PRO 3A-3SG.M.U-old thus 3A-3SG.M.U-tired and 3SG.M.A-trip
’but HE was old so he was tired and he tripped.’
```

Van Staden (2001) offers further arguments in favor of viewing actor and under-goer prefixes as pronominal arguments in Tidore.

10.2.2 Pronominal prefixes as markers of grammatical relations

Pronominal arguments are the sole locus of grammatical relations in North Halmaheran languages. In particular, North Halmaheran nouns are not marked for core case roles, though nouns and pronouns may be marked for non-core oblique roles via adpositions (3) or directional suffixes (4).

(3) Tobelo

```
de ma-kakatama n-a-lye-ino
with ART-tongs 2SG.A-3SG.U-roll-all
’Roll it up with the tongs.’
```

(4) Tobelo

```
o-lyoku-iha to-oiki
ART-mountains-LANDWARD 1SG.A-go
’I’m going inland to the mountains.’
```

Core arguments of both intransitive and transitive verbs are unmarked for case. The same independent pronouns are used to reference the $S_A$, $S_P$, $A$, and $P$ macro-roles. This is exemplified by the first singular pronoun *ngohi* in the following Tobelo examples.

(5) Tobelo

```
ngohi to-tagi
1SG 1SG.A-go
’I am going.’
```

(6) Tobelo

```
ngohi i-hi-tagi
1SG 3A-1SG.U-tired
’I am tired.’
```

(7) Tobelo

```
ngohi to-ni-gohara
1SG 1SG.A-2SG.U-hit
’I hit you.’
```

(8) Tobelo

```
ngohi no-hi-gohara
1SG 2SG.A-1SG.U-hit
’You hit me.’
```
Full nominals may occur with articles, which are obligatory in some North Halmaheran languages. However, these articles do not mark case: the choice of article is independent of macro-role. For example, the Tobelo article o- occurs with both SA (9) and Sp (10) arguments.

(9) Tobelo
   o-nyawa wo-boa
   ART-man 3SG.M.A-arrive
   'A/The man arrived.'

(10) Tobelo
   o-nyawa to-wi-gohara
   ART-man 1SG.A-3SG.M.U-hit
   'I hit a/the man.'

Word order, another cross-linguistically common device for encoding grammatical relations, is also unexploited by North Halmaheran languages for this purpose. While the unmarked word order in the mainland languages Tobelo, Galela, Tabaru, Modole, Loloda, and Pagu is verb-final, alternate orders are frequent enough to cause this to be an unreliable criterion by which to assign grammatical relations. For example, in Tobelo both actor and undergoer nominal arguments may occur following the verb, with both intransitive (11) and transitive (12) verbs.

(11) Tobelo
    i-wi-magawe una
    3A-3SG.M.U-diligent 3M.PRO
    'He was diligent.'

(12) Tobelo
    wo-hi-tulung-oka una
    3SG.M.A-1SG.U-help-perf 3MASC.PRO
    'he helped me.'

In Tidore, Ternate, Sahu, and West Makian, word order is typically more strictly SVO. Nevertheless, alternate word orders are frequently attested.

Given the lack of nominal case marking or rigid word order, pronominal prefixes are the sole domain for encoding grammatical relations in North Halmaheran languages. However, the patterning of pronominal prefixes with respect to S, A, and P macro-roles derives largely from distributional morphological constraints on the occurrence of particular prefixes. As a consequence, seemingly small differences in morphology across the languages lead to radical differences in alignment. Some languages exhibit nominative-accusative patterns; others exhibit clear stative-active patterns. Still other languages exhibit split patterns with nominative-accusative and stative-active subsystems occurring in different domains. These three types of North Halmaheran alignment system are discussed in more detail in the following subsections.

10.2.2.1 Syntactically aligned systems

North Halmaheran languages which lack undergoer pronominal prefixes exhibit a nominative-accusative system of grammatical relations. These can be said to be syntactically aligned, since they organize pronominal prefixes on syntactic rather than semantic principles. The actor prefix indexes A arguments and S arguments of both active and stative verbs (SA and Sp). The following examples are from Tidore, but Ternate behaves
similarly. Intransitive verbs all employ the same pronominal prefix.

(13) a. Tidore
   mo-tag
g3s.pdf-go
   ‘She is going.’ (van Staden 2001)

   b. mo-rohe
   g3s.pdf-pregnant
   ‘She is pregnant.’

   c. mina mo-jang lau
   3f g3s.pdf-beautiful too
   ‘She is very beautiful.’

The same set of prefixes indexes A arguments of Tidore verbs, in this case the 3rd person singular feminine mo-. 

(14) Tidore
   muna mo-fuu nyao toma pasar
   3f.pron g3s.pdf-sell fish loc market
   ‘She sells fish on the market.’ (van Staden 2001)

As will be discussed in the following section, pronominal prefixes are not obligatory in Tidore. Hence, this nominative-accusative alignment is not apparent when the pronominal prefixes are omitted.

West Makian exhibits a similar but slightly different system. Pronominal prefixes are obligatory, and the same prefix is used for both active intransitive verbs and for the A argument of transitive verbs. The vowel of pronominal prefixes of shape CV harmonizes with the following vowel; however, with stative intransitive verb roots the vowel of pronominal prefix is always a high front vowel (Voorhoeve 1982).

(15) West Makian
   a. to-toba
   1sg.a-bathe
   ‘I bathe.’

   b. to-co eme
   1sg.a-see 3pl
   ‘I see them.’

   c. ti-bele
   1sg.a-hungry
   ‘I am hungry.’

Ternate resembles Tidore in having a single pronominal prefix paradigm referencing the actor. However, the precise characteristics of the pronominal systems may differ in these two very closely related languages (cf. van Staden 2001: 17). See Watuseke (1991) for additional details.
The non-harmonizing prefixes are also associated with other classes of verbs, including what Voorhoeve (1982) refers to as directional verbs, those expressing movement in a certain direction. Consequently the category of stative intransitive verbs is not uniquely delimited by morphological properties of the pronominal prefixes. Moreover, as seen in (15), there is no distinct undergoer prefix paradigm and hence no pronominal prefix indexing the P argument of transitive verbs. The West Makian system can be described as nominative-accusative with a subclass of verbs which includes but is not limited to stative verbs.

Nominative-accusative patterns are found also in North Halmaheran languages with distinct actor and undergoer person-marking paradigms. In Sahu, for example, both active and stative intransitive verbs employ the same person-marking prefixes as used for the more agent-like argument of a transitive verb. Thus, the Sahu prefix to-indexes first person singular S₁, S₂, and A arguments in examples (16), (17), and (18), respectively (from Visser and Voorhoeve 1987).

(16) Sahu
to-sapolo
1sg. a-arrive
‘I arrived.’

(17) Sahu
to-malata
1sg. a-cold
‘I am cold.’

(18) Sahu
to-ni-putulu
1sg. a-2sg. u-beat
‘I beat you.’

A distinct prefix ri-indexes the P macro-role.

(19) Sahu
no-ri-putulu
2sg. a-1sg. u-beat
‘You beat me.’

Sahu thus behaves formally like Tidore, except that the P macro-role is explicitly marked via a pronominal prefix.

A formally nominative-accusative alignment is found in Tobelo as well. Tobelo also has a distinct set of undergoer pronominal prefixes which index the P macro-role. However, Tobelo differs from Sahu in the treatment of stative verbs. A large class of stative verbs which are coded as intransitive in Sahu, such as

---

5 As hinted at by Voorhoeve (1982: 13), the high front vowel may have different historical sources in the directional and stative verbs. However, the morphology is synchronically indistinguishable.
Semantically aligned systems

Galela behaves very much like Tobelo except that stative verbs occur without the pleonastic i- prefix. Like Tobelo, Galela has a distinct set of actor and undergoer pronominal prefixes. Active intransitive verbs index $S_A$ arguments via the same paradigm used to index the A argument of transitive verbs.

<table>
<thead>
<tr>
<th>Galela</th>
<th>Galela</th>
</tr>
</thead>
<tbody>
<tr>
<td>no-tagi</td>
<td>no-wi-doto</td>
</tr>
<tr>
<td>‘You are going.’</td>
<td>‘You teach him.’</td>
</tr>
</tbody>
</table>

Stative intransitive verbs index $S_P$ arguments via the same paradigm used to index the P argument of transitive verbs.

<table>
<thead>
<tr>
<th>Galela</th>
<th>Galela</th>
</tr>
</thead>
<tbody>
<tr>
<td>ni-kiolo</td>
<td>wo-ni-doto</td>
</tr>
<tr>
<td>25G.U-asleep</td>
<td>3m.8G.A-25G.U-teach</td>
</tr>
<tr>
<td>‘You are asleep.’</td>
<td>‘He teaches you.’</td>
</tr>
</tbody>
</table>
Galela thus provides what might be called a textbook example of semantic alignment. The S role is split in that the pronominal prefix system does not treat all intransitive verbs in the same way. Rather, arguments of both intransitive and transitive verbs are coded as actor or undergoer according to the underlying verbal semantics, namely lexical aspect. There are certain complexities to the Galela system (to be discussed in the following section), but the basic pattern is quite clear and reflects what might be called the prototypical instance of a semantically aligned system. However, such clear morphological evidence of semantic alignment is difficult to find in the other North Halmaheran languages.

10.2.2.3 Mixed systems Other North Halmaheran languages show a split between syntactic and semantic alignment in different domains. For example, Pagu combines features of both Galela and Sahu. Some Pagu stative intransitive verbs may occur with a single undergoer prefix and no actor prefix. Wimbish (1991: 43) cites the following two examples.

(29) Pagu
oras gena ngoi ni-dogoli i-togu-togum-uwa
time that 1PRO 1SG.P-labour 3A-RED-stop-NEG
'At that time I was having labour pains that didn’t stop.'

(30) Pagu
o-panyakit wi-daen
ART-disease 3SG.M.P-experience
'He had a disease.'

Other apparently stative verbs behave as in Sahu, requiring an actor prefix.

(31) Pagu
ai-lokat wo-sawin
3SG.M.POSS-husband 3SG.M.A-hungry
'My husband is hungry.' (Wimbish 1991: 32)

In Tabaru we find a mixture of Galela and Tobelo systems. Tabaru pronominal prefix alignment is split based on person. First person plural inclusive arguments of stative intransitive verbs behave as in Tobelo, requiring a pleonastic i- prefix.

(32) Tabaru
po-okere po-maka-sano asa i-na-surugogo
1INCH.A-drink 1INCH.A-RECIP-ask FUT 3A-1INCH.U-choke
'If we drink and ask each other, then we'll choke.'

Other persons behave as in Galela, employing only the undergoer paradigm.

(33) Tabaru
i-punusu-okau so to-odomo-uwau
1SG.U-satiated-PERF thus 1SG.A-EAT-NEG
'I am satiated so will eat no more.'
These examples from Pagu and Tabaru demonstrate only a few of the many ways in which mixed alignment systems can arise in North Halmaheran languages. From a purely synchronic point of view, Pagu and Tabaru can be viewed as split systems with semantic alignment localized in one part of the pronominal prefix system. But such a synchronic analysis ignores a fundamental lexical semantics which underlies the system. Significantly, the existence of mixed systems as in Pagu and Tabaru may help to delineate a diachronic pathway by which semantically aligned systems as in Galela may have evolved from syntactically aligned systems as in Tobelo. In order to understand this evolution it is necessary to examine more closely the constraints on the omission of pronominal prefixes.

10.3 Omission of pronominal prefixes

The discussion so far has tacitly assumed that pronominal prefixes are obligatory. In fact, in most North Halmaheran languages the occurrence of pronominal prefixes is governed by discourse constraints including topicality, referentiality, and definiteness. However, the precise constraints governing the occurrence of pronominal prefixes differ significantly among the languages, and these differences in the obligatoriness of pronominal prefixes help to explain the apparent differences in the formal realization of alignment patterns.

Among those languages which index the P argument of transitive verbs via an undergoer prefix, the indexing of non-human third person undergoer arguments is governed by referentiality. In the following Tobelo examples the referential definite argument gota ‘wood’ in (34) is referenced via a pronominal prefix on the verb, while the non-referential argument hilo ‘resin’ in (35) is not.

(34) Tobelo
   o-gota mi-a-tobiki de mi-a-bela-belaka
   ART-wood 1INCH.A-3U-break and 1INCH.A-3-RDP-split
   ‘We bucked and split the wood.’

(35) Tobelo
   jadi ngohi to-lye-ua o-hilo
   therefore 1PRO 1SG.A-get-NEG ART-resin
   ‘Therefore I didn’t get any resin.’

Discourse factors conditioning the omission of undergoer prefixes in Galela and Pagu are discussed in more detail by Shelden (1986) and Wimbish (1991), respectively.

More relevant to the discussion of alignment are the constraints on the occurrence of actor prefixes. In Tobelo, the actor prefix is obligatory with all verbs. This constraint holds even when an independent pronoun or nominal argument is present, as the following Tobelo examples demonstrate.
In Tobelo, actor pronominal prefixes are obligatory even in imperative constructions.

Furthermore, Tobelo actor prefixes cannot be omitted even when an undergoer prefix is present.

Crucially, actor prefixes are required in Tobelo even with semantically intransitive stative verbs which index their single argument via an undergoer pronominal prefix. In this case the actor prefix position is filled by the 3rd person non-human actor prefix i- (or at least a morpheme homophonous with this one) which functions as a pleonastic marker lacking any explicit argument.

Because of the presence of this pleonastic actor prefix, Tobelo stative verbs are formally transitive.

Among the North Halmaheran languages Tobelo is particularly strict in enforcing the constraint against omission of the actor prefix. In other languages it is possible to omit actor prefixes under certain conditions. As was noted above, Galela stative intransitive verbs index their single argument via an undergoer prefix, and no actor prefix is present.

Galela

ni-kiole
2sg.u-asleep
‘You are asleep.’
Galela actor prefixes may be also omitted with transitive verbs under certain conditions. Deidre Shelden (1986) proposes that Galela actor prefixes may be omitted when the referent is non-topical, or at least less topical than the undergoer referent.

(42) Galela
  o-nyawa ni-sano
  ART-person 2SG.U-ask
  ‘Someone asked you.’ (Shelden 1986: 235)

The constraints on actor prefixes are similar in Modole. Certain stative intransitive verbs, such as the nominal predicate -to'ata ‘be a witch’ in (43), may occur without an actor prefix.

(43) Modole
  ma-ngoa'a ge wo-temo, apu, bote ni-to'ata
  ART-child that 3SG.M.A-say granny surely 2SG.U-witch
  ‘The child said: “Granny, you must be a witch.”’ (Ellen 1916)

Transitive verbs, such as -dahe ‘to court’ in (44) and -dihiwa ‘shine on’ in (45), may also occur without an actor prefix when the actor is non-topical.

(44) Modole
  o-nyawa moi w-a-ino mi-dahe
  ART-person one 3SG.M.A-VP-ABL 3SG.FEM.U-court
  ‘Someone came to court her.’ (Ellen 1916)

(45) Modole
  de gen'a ade wi-dihiwa o-wange
  and then 3SG.M.U-shine.on ART-sun
  ‘And then the sun shone on him.’ (Ellen 1916)

In Tabaru, actor prefixes other than the pleonastic i- actor prefix cannot be omitted. However, as noted in the previous section, the i- actor prefix can be omitted with stative verbs when the undergoer argument is other than 1st person plural inclusive.

(46) Tabaru
  ni-tootasa
  2PL.U-angry
  ‘You (pl.) are angry.’ (Fortgens 1928: 362)

A similar situation arises in Pagu, where continued reference to less important A referents may be omitted. For example, in the Pagu text excerpt in (47) the first token of the verb -ao ‘carry’ occurs with an actor prefix i-, while the second token, in the immediately following line, occurs without i-.
Gary Holton

(47) Pagu
dina-isagram inland-landward

‘They carried her inland (to the house).’

(48) Tidore
muna wo-sari wo-wako
be.about.to be.return

‘She is about to go home.’

(49) Tidore
muna sari wako
be.about.to return

‘She is about to go home.’

Actor prefixes in Tidore thus appear to be entirely optional, provided the referent can be sufficiently identified via accompanying pronoun or nominal. Thus, Tidore differs from both Tobelo and Galela in freely permitting bare stem verbs. In fact, in some cases Tidore verbs may not admit actor pronominal prefixes, as in some serial verb constructions.

(50) Tidore
ngofa ngge yo-pekak (*yo-*)tora
cild fall (*3A-*)downwards

‘The child fell down.’

In contrast, actor prefixes are never omitted in serial verb constructions in Tobelo.
Table 10.2. Constraints on person-marking prefixes in North Halmaheran languages

<table>
<thead>
<tr>
<th>Language</th>
<th>A obligatory</th>
<th>U paradigm</th>
<th>U obligatory</th>
<th>U with stative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tidore</td>
<td>−</td>
<td>−</td>
<td>n. a.</td>
<td>n. a.</td>
</tr>
<tr>
<td>W. Makian</td>
<td>+</td>
<td>−</td>
<td>n. a.</td>
<td>n. a.</td>
</tr>
<tr>
<td>Sahu</td>
<td>(+)</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Tobelo</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Galela</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>Tabaru</td>
<td>(+)</td>
<td>+</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>Modole</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Pagu</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>(+)</td>
</tr>
</tbody>
</table>

(51) Tobelo

\[
\begin{array}{ccc}
\text{o-gaharu} & \text{yo-ija} & \text{yo-uti} \\
\text{ART-sandalwood} & \text{3PL.A-buy} & \text{3PL.A-descend} \\
\end{array}
\]

‘They came down to buy sandalwood.’

\(\ast o\text{-gaharu yo-ija uti}\)

The omission of actor prefixes with active intransitive verbs has also been reported in some other North Halmaheran languages. For example, Visser and Voorhoeve (1987) report that 3rd person singular and plural prefixes may be omitted in modern Sahu by younger speakers. However, 1st person singular actor prefixes are never omitted in Sahu, even by younger speakers.\(^6\)

As is evident from the preceding discussion, the differences in the behaviour of pronominal prefixes in the North Halmaheran languages derive largely from the status of the prefix as optional or obligatory in the various languages, as well as the treatment of stative verbs. The behaviour of pronominal prefixes can be summarized along four parameters: obligatoriness of the actor prefix; presence of an undergoer paradigm; the obligatoriness of the undergoer prefix; and the use of undergoer prefixes with semantically stative verbs. Table 10.2 categorizes the eight North Halmaheran languages for which the best data are available according to these three parameters. Those entries listed in parentheses indicate a feature which has been reported to be inconsistent. For example, younger speakers of Sahu reportedly omit the actor prefix. As it happens, each of these eight North Halmaheran languages exhibits a different set of values for these four parameters. Cross-language variation in the obligatory nature of verbal person-marking prefixes has profound implications for differences in the nature of grammatical relations in the various North Halmaheran languages.

\(^6\) It is not entirely clear whether one can conclude that the omission of 3rd person actor prefixes in modern Sahu results in bare stem verbs, as Visser and Voorhoeve offer no explicit examples beyond the statement on the bottom of p. 30, which could be interpreted as applying equally to transitive as intransitive verbs.
The variation in constraints on the occurrence of pronominal prefixes tends to obscure underlying alignment patterns. Arguably, this is due to the fact that the locus of semantic alignment lies not in the prefix system, but in the lexicon.

10.4 Semantic alignment and the lexicon

We must be careful not to paint all semantically aligned systems with the same broad brush. The distinction between ‘agentive’ systems, which base case marking on semantic properties of the participant, and ‘active’ systems, which base case marking on the lexical aspect or Aktionsart of the situation, has been well established. But semantically aligned systems may also be sensitive to much subtler distinctions (cf. Mithun 1991). Such subtleties are apparent across the North Halmaheran languages.

While North Halmaheran languages code actor as distinct from undergoer, the choice of what counts as an actor is highly lexicalized, based on the dynamicity of the verb itself. North Halmaheran semantic alignment thus reflects a stative-active pattern based on lexical aspect rather than an agent-patient system based on agency of the actor. The notion of actor as ‘the participant which performs, effects, instigates, controls the situation denoted by the predicate’ (Foley and Van Valin 1984: 29) is not relevant to the choice of pronominal prefix. Both controlling (52) and non-controlling (53) agents of dynamic intransitive verbs are coded with actor prefixes.

(52) Tobelo controlling agents requiring actor prefix

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>bata</td>
<td>‘jump’</td>
</tr>
<tr>
<td>hioru</td>
<td>‘paddle’</td>
</tr>
<tr>
<td>hoho</td>
<td>‘fly’</td>
</tr>
<tr>
<td>oara</td>
<td>‘run’</td>
</tr>
<tr>
<td>ohiki</td>
<td>‘bathe’</td>
</tr>
<tr>
<td>olyomo</td>
<td>‘eat’</td>
</tr>
<tr>
<td>temo</td>
<td>‘speak’</td>
</tr>
<tr>
<td>dumunu</td>
<td>‘dive’</td>
</tr>
<tr>
<td>toimi</td>
<td>‘shoot’</td>
</tr>
</tbody>
</table>

(53) Tobelo non-controlling agents requiring actor prefix

<table>
<thead>
<tr>
<th>Verb</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>adono</td>
<td>‘reach’</td>
</tr>
<tr>
<td>ari</td>
<td>‘cry’</td>
</tr>
<tr>
<td>gegoto</td>
<td>‘worry’</td>
</tr>
<tr>
<td>gehanga</td>
<td>‘yawn’</td>
</tr>
<tr>
<td>gogere</td>
<td>‘live’, ‘dwell’</td>
</tr>
<tr>
<td>guroko</td>
<td>‘snore’</td>
</tr>
<tr>
<td>hangeru</td>
<td>‘sneeze’</td>
</tr>
<tr>
<td>iete</td>
<td>‘laugh’</td>
</tr>
<tr>
<td>wunenge</td>
<td>‘vomit’</td>
</tr>
</tbody>
</table>
Even non-agents may be coded as actors in Tobelo provided they are participants in a dynamic event.

(54) Tobelo non-agentive verbs requiring actor prefix

\begin{itemize}
  \item \textit{honenge} ‘die’
  \item \textit{lyahini} ‘drift’
\end{itemize}

On the other hand, participants can be grammatically encoded as undergoers even when they are semantically performing, effecting or instigating (in the sense of Mithun 1991), as with \textit{gogama} ‘shivering’ and \textit{kioko} ‘asleep’ above, and \textit{magawe ‘diligent’} in (55).

(55) Tobelo

\begin{itemize}
  \item[\textit{i-wi-magawe} una \textit{3a-3sg.m.u-diligent 3m.pro}]
  \item[\textit{He was diligent,’}]
  \item[\textit{de ai-hininga i-rahai}]
  \item[\textit{3sg.m.poss 3a-good}]
  \item[\textit{‘and his heart was good.’}]
\end{itemize}

The crucial factor for determining choice of pronominal prefix with Tobelo intransitive verbs is the dynamicity of the verb itself. Yet just which situations are construed as dynamic and which are construed as static can vary across the North Halmaheran languages. While Tobelo and Galela behave nearly identically in this respect, other languages differ.

For example, there are a number of Tabaru intransitive verbs which take actor prefixes but whose cognates in Tobelo and Galela take undergoer prefixes. Some Tobelo forms of these verbs are listed in (56) below.

(56) Tobelo stative intransitive verbs whose Tabaru cognates take actor prefixes

\begin{itemize}
  \item[\textit{siri}] ‘sick’
  \item[\textit{amoko}] ‘large’
  \item[\textit{piloko}] ‘blind’
  \item[\textit{tingoono}] ‘small’
  \item[\textit{potoono}] ‘healthy’
  \item[\textit{pesaka}] ‘wet’
  \item[\textit{faro}] ‘feverish’
  \item[\textit{tiikiti}] ‘cough’
  \item[\textit{kuata}] ‘strong’
  \item[\textit{patilanga}] ‘stiff’
  \item[\textit{kakuru}] ‘long’
  \item[\textit{morene}] ‘happy’
  \item[\textit{tebini}] ‘beautiful’
  \item[\textit{pelelaka}] ‘thin’
  \item[\textit{timono}] ‘aged’
  \item[\textit{pululunu}] ‘thick’
  \item[\textit{pereki}] ‘old’
  \item[\textit{burere}] ‘fat’
  \item[\textit{kudai}] ‘numerous’
\end{itemize}

Thus, the verb ‘sick’ is coded as a stative verb with undergoer prefix in Tobelo and Galela, while the cognate form in Tabaru is coded as an active verb with actor prefix. On the other hand, there exist a number of other stative intransitive verbs in Tobelo and Galela whose Tabaru cognates do require undergoer prefixes, as in (57).
Gary Holton

(57) Tobelo stative intransitive verbs whose Tabaru cognates take undergoer prefixes

<table>
<thead>
<tr>
<th>Tabaru</th>
<th>Tobelo</th>
</tr>
</thead>
<tbody>
<tr>
<td>tootasa</td>
<td>'angry'</td>
</tr>
<tr>
<td>eto</td>
<td>'drunk'</td>
</tr>
<tr>
<td>tirine</td>
<td>'tremble'</td>
</tr>
<tr>
<td>punusu</td>
<td>'satisfied'</td>
</tr>
<tr>
<td>pelesoko</td>
<td>'brave'</td>
</tr>
<tr>
<td>sowono</td>
<td>'sob'</td>
</tr>
<tr>
<td>tuuduku</td>
<td>'bored'</td>
</tr>
</tbody>
</table>

Thus, while Tabaru is semantically aligned (in at least part of the person marking paradigm), the semantics governing the distinction between the S_A and S_P roles differ from those in Galela. Tabaru and Galela differ in terms of what counts as ‘dynamic’. That is, dynamicity is lexicalized.

Nevertheless, semantics still plays an active role in the lexicon of North Halmaheran languages. Nearly all intransitive verb roots may occur in either ‘active’ or ‘stative’ form. That is, they may cross-reference their single argument by either the actor or undergoer paradigm, with concomitant change in lexical semantics. For example, the Tobelo verb eluku ‘lie, deceive’ may index a single core argument via the actor (58) or undergoer (59) paradigm.

(58) Tobelo

\[ \text{Tobelo} \]

\[ \text{wo-eluku-oka} \quad 3\text{sg.m.a-lie-perf} \]

\[ \text{He lied. ‘He is a liar.’} \]

(59) Tobelo

\[ \text{Tobelo} \]

\[ \text{i-wi-eluku} \quad 3\text{a-3sg.m.u-lie} \]

\[ \text{He lied. ‘He is a liar.’} \]

The division here is based on lexical aspect. The actor intransitive verb eluku is construed as a single telic action; the undergoer intransitive verb is eluku is construed as an atelic state, without a well-defined end point. Thus, (59) might be equally well glossed as ‘He continuously lies’ or ‘He is in a state of lying’. Additional examples from Tobelo are given in (60).

(60) Tobelo intransitive verbs which take actor or undergoer prefixes

<table>
<thead>
<tr>
<th>Actor</th>
<th>Undergoer</th>
</tr>
</thead>
<tbody>
<tr>
<td>birahi</td>
<td>‘rejoice’</td>
</tr>
<tr>
<td>hitanga</td>
<td>‘go astray’</td>
</tr>
<tr>
<td>kioko</td>
<td>‘go to sleep’</td>
</tr>
<tr>
<td>lihiti</td>
<td>‘sprain’</td>
</tr>
<tr>
<td>modongo</td>
<td>‘fear’</td>
</tr>
<tr>
<td>ngamo</td>
<td>‘quarrel’</td>
</tr>
<tr>
<td>tikiti</td>
<td>‘cough’</td>
</tr>
<tr>
<td>tohata</td>
<td>‘angry’</td>
</tr>
</tbody>
</table>

Whether or not one conceives of verbs such as those in (60) as having two distinct lexical entries corresponding to different lexical semantics, Tobelo alignment is ‘fluid’ in the sense that most intransitive verb roots can occur with either actor
or undergoer morphology, with concomitant difference in lexical semantics. In Tobelo, stative intransitive verbs are formally encoded as transitive, with a pleonastic i- prefix. In Galela, both active and stative intransitive verbs are coded as intransitive. There may be subtle variations in the semantics of the pronominal prefixes across the North Halmaheran languages, but the choice of actor vs. undergoer pronominal prefix is ultimately governed by semantics.

Indeed, the very fact that the split between $S_A$ and $S_P$ can be made in different ways by closely related languages can be said to be a hallmark of semantically aligned systems, which are much freer to make subtle adjustments to the semantic parameters which govern them. Such phenomena are not unknown in syntactically aligned systems, witness dative subject in European languages, for example; however, such variation in the formal realization of argument marking is arguably more common in semantically aligned systems. So it is that Tobelo and Galela treat stative verbs very differently from Tidore and Sahu, while Tabaru lies somewhere in between, with some stative intransitive verbs taking undergoer prefixes as in Tobelo and Galela, and others taking actor prefixes as in Tidore and Sahu.

10.5 Semantic realignment

While they may differ formally, semantically aligned systems like Galela are not so different from syntactically aligned systems such as Tobelo. Both types of system index the single argument of lexically stative, non-dynamic, atelic verbs using the undergoer paradigm. The difference lies in the presence of the pleonastic i- actor prefix in Tobelo. The formal realization of alignment depends crucially on the status of the i- prefix.

Most existing descriptions analyse stative verbs as a kind of passive (cf. van Baarda 1891). If i- were to be analysed as a passive marker, then this would be a non-promotional passive, since the passive ‘subject’ remains in situ as an undergoer. One might consider i- to be a detransitivizing morpheme which neutralizes the actor argument. Yet in Tobelo there is no formal difference between semantically intransitive (single-participant) undergoer constructions such as (61) and semantically transitive (two-participant) constructions with 3rd person singular non-human actors, as in (62).

(61) Tobelo
   i-hi-bole
   3a-1sg.u-tired
   ‘I am tired.’

(62) Tobelo
   i-hi-goli
   3a-1sg.u-bite
   ‘It bit me.’

That is, there is no evidence that the construction in (61) has been passivized. Some Tobelo verb roots can indeed occur as either formally intransitive or formally transitive, with either one or two pronominal prefixes, respectively. Yet most stative intransitive verbs do not occur as transitives. In particular, there is no corresponding transitive verb root bole. So if (61) is a passive then it has no
active transitive counterpart. Ultimately, the analysis of Tobelo stative intransitive constructions with $i$- as passives relies on an analogy with European languages. The $i$- prefix may indeed have a historic role as an actor marker, but it no longer functions synchronically in this way.

Tobelo and Galela represent two poles on a continuum of possibilities for undergoer cross-referencing in North Halmaheran languages. Tobelo requires the $i$- prefix with all semantically intransitive stative verbs; in Galela, the $i$- prefix is never present. Other languages fall somewhere in between. For example, in Tabaru the $i$- prefix occurs with some persons but not others. Comparative and historical evidence suggests that constraints on the occurrence of the pleonastic $i$- prefix have changed significantly in the recent history of North Halmaheran languages, leading to realignment of the pronominal prefix system in some languages.

The formal semantic alignment attested in Galela is clearly a recent phenomenon resulting from aphaeresis of the $i$- prefix. While modern Galela lacks the $i$- prefix in undergoer intransitive verbs, data from the late 19th century show a change in progress. Most Galela stative verbs cited by van Baarda (1891, 1895) and Kern (1891, 1892) include the $i$- prefix as in modern Tobelo. For example, the $i$- prefix is present throughout the paradigm for toosa ‘angry’ (Kern 1892: 118).

\[
\begin{align*}
\text{Galela} \\
i-i\text{-toosa} & \quad \text{‘I am angry’} \\
i-ni\text{-toosa} & \quad \text{‘you are angry’} \\
i-wi\text{-toosa} & \quad \text{‘he is angry’} \\
i-mi\text{-toosa} & \quad \text{‘she is angry’}
\end{align*}
\]

However, other paradigms are defective, evidencing an erosion of the actor prefix (Kern 1892: 119).

\[
\begin{align*}
\text{Galela} \\
i-wi\text{-pereki} & \quad \text{‘he is old’} \\
mi\text{-pereki} & \quad \text{‘she is old’}
\end{align*}
\]

Modern Tobelo shows evidence of a similar change in progress. While citation forms of stative verbs appear with initial $i$-, the prefix is frequently elided in fast speech. The $i$- prefix is never lost when it functions as an actor argument marker referencing non-human actors; aphaeresis is possible only when the $i$- is functioning as a pleonastic form with no argument status.

The use of constructions with ‘experienter’ objects to code stative verbs is actually quite common in Eastern Indonesia and New Guinea. For example, Reesink (1998) identifies a set of ‘experiential verbs’ in Sough (East Bird’s Head) which inflect via the undergoer (‘object’) paradigm. Actor pronouns are prefixes, while undergoer pronouns are suffixes. However, the 3rd person singular actor prefix is $\emptyset$-, and hence constructions such as the following are reminiscent of
Semantic alignment in North Halmahera, Indonesia

10. Donohue

Tobelo stative verbs.

(65) Sougb (East Bird’s Head)

(Ø-)areb-et

(3sg.A.)-sick-1sg.U

‘I am sick.’

The verb areb is actually a complex form consisting of ara ‘something’ and eb ‘do’; thus, (65) is literally ‘it did something to me’. It seems likely that these constructions have arisen from some sort of inanimate causer, as noted by Boelaars (1950) for Marind (Trans New Guinea).

If stative constructions in North Halmaheran languages arise from experiencer verb constructions, then languages such as Sahu must have undergone a reanalysis by which all intransitive verbs, including statives, were indexed via the actor paradigm, presumably by analogy with active intransitives. Yet in spite of this reanalysis, sensitivity to lexical aspect can still be detected in Sahu. It has been reported that certain Sahu stative verbs may occur without actor prefixes in certain contexts.

(66) Sahu

wala (i-)lamo’o

house (3a)-large

‘Large house’/’The house is large.’

(67) Sahu

bele (i-)kiau

banana (3a)-young

‘Young banana’/’The banana is young.’

Visser and Voorhoeve ascribe the difference between the presence or absence of the actor prefix in examples such as (66) and (67) to a difference between predication and attribution, respectively. In fact, either construction can be used either attributively or predicatively. Visser and Voorhoeve go on to acknowledge as much in noting that constructions involving the adverbial particle la ‘very’ may yield either meaning. Thus, ngo’om la i-tiara may mean ‘the very long road’ or ‘the road is very long’.

The lack of a morphological distinction between attributive and predicative adjectival verb constructions has been reported also for Tobelo (Holton 1999). In Sahu the crucial point is that the bare-stem constructions without the actor prefix co-exist with pure nominal forms containing the article ma-, as in (68).

(68) Sahu

wala ma-lamo’o

house art-large

‘The large house.’
Unlike the bare-stem forms above, these nominal forms do not admit a predicative interpretation. It seems plausible to interpret the bare-stem forms in (66) and (67) as verbs which lack pronominal prefixes. Thus, a class of Sahu stative intransitive verbs can be morphologically delineated by their ability to occur without an actor pronominal prefix. And this class consists precisely of those which in Tobelo and Galela occur with undergoer pronominal prefix.

Syntactically aligned North Halmaheran languages such as Tidore reflect a somewhat different path of evolution through which undergoer prefixes have been lost entirely. Other North Halmaheran languages appear to be presently evolving in that direction. For example, in Sahu the use of undergoer prefixes is currently giving way to the use of independent pronouns (Visser and Voorhoeve 1987: 30). Older speakers maintain the undergoer prefixes, as in (69):

(69) Sahu
\[ \text{ngoi ti-ni-elingi} \]
\[ \text{1PRO 1SG.A-2SG.U-think.of} \]
'I think of you.' (formal speech)

while younger speakers tend to replace undergoer prefixes with independent pronouns or full nominals, as in (70):

(70) Sahu
\[ \text{ngoi ti-elingi ngana} \]
\[ \text{1PRO 1SG.A-think.of 2PRO} \]
'I think of you.' (informal speech)

The loss of undergoer prefixes in Tidore, West Makian, and (eventually) Sahu is ostensibly due to contact with Austronesian languages, particularly the long-standing presence of Malay (Voorhoeve 1988). However, even those languages which have lost undergoer prefixes may show some evidence for a stative category. Although actor prefixes are optional in Tidore, van Staden (2001: 79) notes that actor prefixes occur less frequently on 'adjectives' (stative verbs in the other North Halmaheran languages) than on verbs. This trace of a stative category thus survives in Tidore in spite of the formal realignment of the pronominal system.

10.6 The rise and fall of semantic alignment

Given knowledge of the current synchronic variation in the formal alignment patterns, we can outline a possible scenario by which semantic alignment may have arisen in North Halmaheran languages. The internal comparative evidence suggests that proto-North Halmaheran probably reflected a syntactically aligned system much as in modern Tobelo. Experiencer object constructions gave way
to the use of a pleonastic subject prefix with stative verbs. The erosion of the now pleonastic subject prefix in such constructions led to a distinction between subjective and objective intransitive verbs, that is, a distinction in the marking of $S_A$ and $S_P$. This change is complete in Galela; nearly complete in Pagu and Modole; somewhat complete in Tabaru; and perhaps beginning in modern colloquial Tobelo. In contrast, in Tidore and West Makian, loss of the undergoer prefix has led to further realignment by which all intransitive verbs inflect via the single remaining (actor) paradigm. Sahu appears to be headed along a similar path. Sahu drops the undergoer prefix with stative verbs and is in the process of reanalysing them as active. These various stages in the evolution of alignment can be summarized along a cline from syntactic to semantic alignment. Erosion of the pleonastic actor prefix yields semantically aligned systems (right arrow in Figure 10.3), while further erosion of undergoer prefixes yields syntactically aligned systems (left arrow).

The return path of historical evolution from semantic alignment to syntactic alignment is somewhat speculative (hence the dashed arrow). Additional data may help to clarify whether languages like Tidore evolved through a previously semantically aligned stage or, alternately, directly from a syntactically aligned system through loss of undergoer prefixes. Nevertheless, the evidence for the cline from syntactic toward semantic alignment is substantial. The development of a pleonastic subject prefix and the subsequent reanalysis of object markers as undergoer markers may constitute an important mechanism in the rise of semantic alignment.

Even more striking is the degree to which the North Halmaheran languages share important semantic features in spite of differences in formal alignment patterns. As we have seen, both Tobelo and Galela exhibit nearly identical alternations between active intransitive and stative intransitive verbs, yet the requirement of a pleonastic subject prefix in Tobelo but not Galela renders the Tobelo system syntactically aligned and the Galela system semantically aligned. The modern multiplicity of formal alignment systems belies an underlying semantic unity. The formal analysis of alignment patterns in North Halmaheran languages is inherently unstable: even a slight change in a phonological or morphological
feature can lead to complete reanalysis of the formal system for person marking. And yet the underlying semantics of the systems remain consistent. While the North Halmaheran languages may differ formally in the expression of alignment via verbal prefixes, all of the North Halmaheran languages can be said to exhibit semantic properties consistent with semantically aligned systems. In short, the North Halmaheran languages are semantically aligned in spirit, if not always in form.
11

Verb classification in Amis

NAOMI TSUKIDA

11.1 Introduction
Amis is an Austronesian language spoken in Taiwan. The verbs of this language can be classified into four groups by morphological criteria. This classification at first seems quite different from ordinary semantic alignment systems, but it turns out that the notions that have appeared repeatedly in the literature on semantic alignment systems, i.e. dynamicity, affectedness, and control (see e.g. Mithun 1991), play an important role in explaining the classification. Moreover, if we investigate it more closely, a phenomenon analogous to A and P marking in ordinary semantic alignment system emerges (see 11.4.2).

Section 11.2 is a brief introduction to the language. In section 11.3 I discuss the verb classification in Amis. Section 11.4 is a comparison and discussion. Section 11.5 is the summary.

11.2 About the language
Amis is spoken in the Hualien and Taitung district of Taiwan. Amis data are based on the Fata’an dialect. It is one of the Austronesian languages of Taiwan, usually called Formosan. There is much controversy about the internal classification of Formosan languages and about their position in the Austronesian languages (Adelaar 2005: 11; Ferrel (1969), for example, classified Amis as part of the Paiwanic group, and Blust (1999) classified it as East Formosan. Most scholars agree that Taiwan is the homeland of Austronesian languages, and Blust (1999) even classifies the nine branches of Formosan languages directly under Proto-Austronesian.

The ethnic population of Amis is about 140,000. Not all these people can speak the language, especially members of the younger generation. Amis is an endangered language, in spite of the rather large population.

Sentential word order is predicate-initial, but the position of other constituents is less rigid. Amis has Nominative, Genitive, Oblique, and Locative cases distinguished mainly by propositions and suffixes; unique inflected forms exist
Naomi Tsukida

Table 11.1. The Amis verb paradigms

<table>
<thead>
<tr>
<th></th>
<th>1. mi</th>
<th>2. om</th>
<th>3. ma</th>
<th>4. ϕ</th>
</tr>
</thead>
<tbody>
<tr>
<td>AV Connegative</td>
<td>pi-CX</td>
<td>ka-(C-o)m-X</td>
<td>ka-CX</td>
<td>ka-CX</td>
</tr>
<tr>
<td>AV Neutral</td>
<td>mi-CX</td>
<td>(C-o)m-X</td>
<td>ma-CX</td>
<td>CX</td>
</tr>
<tr>
<td>AV Neutral relative</td>
<td>mi-CX-ay</td>
<td>(C-o)m-X-ay</td>
<td>ma-CX-ay</td>
<td>CX-ay</td>
</tr>
</tbody>
</table>

*C stands for the first consonant of the stem and X for the rest of the stem.

for pronouns. In noun phrases, the order of Head and Adjective or Genitive is free.

The Formosan languages are close to the better-known Philippine languages typologically; most of them have Philippine-type 'Focus system', and the Amis language is no exception. In a Focus system one of the nominals is selected as the subject of the clause and its semantic role is indicated by the verb form. Not only Agent and Patient but also Location, Beneficiary, and Instrument can be chosen as the subject. A Focus system is thus a kind of voice phenomenon, but all the voice forms are equally marked, unlike the Active/Passive or Ergative/Antipassive alternations, in which one of the two voice forms is more marked than the other.

Amis has two voices: Agent Voice (AV) and Goal Voice (GV). Agent Voice signals that the subject is either the A of transitive verbs or the S of intransitive verbs. Goal Voice signals that the subject is either Patient, Goal, Location, or Recipient. Other languages with a Philippine-type Focus system may have dedicated Instrumental/Beneficiary Voice, or Locative Voices; Amis has cognate morphemes, but they function rather to derive deverbal nouns than as part of the Focus system.

The Agent Voice in Amis has four types of conjugation. Verbs can be classified according to which conjugation a particular verb takes; so Amis verbs fall into four different classes. We will examine the system in detail in the following section.

11.3 Verb classification in Amis

Amis has four conjugations, shown in Table 11.1: the AV connegative form, AV neutral form, and AV neutral relative form.1 The notation (C-o)m-X indicates that om is inserted into the stem CX as an infix, but that the first syllable of the resulting form may or must (depending on the verb) be omitted.

1 The connegative form appears with the negative predicate caay. The connegative is also the stem for several derivations. The neutral form is used as a predicate to express present, past and future events/situations. The neutral relative form is used as the head of a referential phrase, or as a modifier in a referential phrase.
Verb classification in Amis

<table>
<thead>
<tr>
<th></th>
<th>−state</th>
<th>+state</th>
</tr>
</thead>
<tbody>
<tr>
<td>−affected</td>
<td>mi, ϕ-A (+control)</td>
<td>ϕ-B (−control)</td>
</tr>
<tr>
<td>+affected</td>
<td>om (+control), ma ma</td>
<td>ma ma</td>
</tr>
</tbody>
</table>

Below are some examples of verbs that take each conjugation (ordered alphabetically by the Amis spelling). 2

1. **mi** conjugation: **mi-clem** 'dive', **mi-cofay** 'go back on the way', **mi-dagoy** 'swim', **mi-holol** 'play'.
2. **om** conjugation: **m-aro** 'sit down', **m-isi** 'urinate', **m-ota** 'vomit', **(r-o)m-akat** 'walk'.
3. **ma** conjugation: **ma-fer** 'fly', **ma-fkac** 'run', **ma-lipahak** 'be glad', **ma-tlaaf** 'belch', **ma-tlook** 'hiccup'.
4. **ϕ** conjugation: **cigagan** 'be famous', **taini** 'come', **tanektek** 'be strong, be stout', **tatiih** 'be bad'.

Three conjugations include transitive members as well. Examples:

1. **mi**: **mi-lpel** 'catch', **mi-nanom** 'to drink', **mi-palo** 'hit', **mi-pres** 'suppress', **mi-saga** 'make'.
2. **om**: **(k-o)m-aen** 'eat', **(r-o)m-adiw** 'sing', **(s-o)m-owal** 'speak'.
3. **ma**: **ma-dakaw** 'take a ride on', **ma-fana** 'know', **ma-pawan** 'forget'.

The transitive **om** verbs are exhaustively listed above. Transitive **ma** verbs also include verbs of emotion. The **mi** conjugation includes more transitive members, and only a few examples have been given above.

11.3.1 Semantic description

Verbs in each conjugation are semantically characterized as follows. **ϕ** conjugation verbs are classified into **ϕ**-A and **ϕ**-B by their morphology and semantics (see 11.3.1.1 and 11.3.1.4 for detail).

- **mi, ϕ-A**: −state, −affected, +control
- **om**: −state, +affected, +control
- **ma**: ±state, +affected, ±control
- **ϕ-B**: +state, −affected, −control

The above is summarized in Table 11.2.

---

2 The Amis orthography used here is adapted from that of the Bible Society, which has published an Amis Bible and an Amis dictionary (Fey 1986). Symbols used in a non-standard way are: 'epiglottal plosive (cf. Edmondson et al. 2005), c voiceless alveolar affricate, d voiced or voiceless alveolar lateral fricative, e schwa, g velar nasal, l alveolar flap, r alveolar trill.'
The feature ±state indicates whether the situation denoted by the verb is state
or not. The feature −state indicates activity, or dynamic event. The notion of
static/active has been an important term to explain semantically aligned verb
classification systems, with many of them referred to as static/active languages.

The feature +affected indicates either that the verb expresses a situation where
the subject is the most affected entity or that the situation is caused through being
affected by another entity. The feature −affected indicates that the subject is not
the most affected entity or that the situation is not caused by another entity but is
inherent or permanent.

The notion of affectedness has also been employed to explain intransitive splits
in other languages. For example Mithun (1991) used this notion in explaining the
split in Eastern Pomo:

It is affectedness. . . . being dry, dangerous, empty, hanging, etc., like being satiated or afraid
are considered resultant, possibly temporary conditions: something has happened to their
participants. (Mithun 1991: 532)

This notion is also used to explain the use of middle voice. Klaiman (1991:
105) writes: ‘the middle implicates the logical subject’s affectedness.’ She explains
‘affectedness’ as follows:

Middle encodes a range of specific meanings,. . . , such that the referent of the nominal
which a verb assigns as logical subject coincides with the locus of the principal effect of the
verbally denoted action. (pp. 105–6).

I accept both of these definitions, and will use ‘affectedness’ in a broad sense.

The feature +control means that the Actor controls the situation denoted by
the verb. It may be instigated voluntarily or involuntarily, intentionally or unin-
ventionally. Control is irrelevant to the mi conjugation.

11.3.1.1 mi, ϕ-A: −state, −affected, +control Mi verbs and some of the ϕ verbs
express −state, −affected, +control situations, i.e. actions or processes in which
the subject is not the locus of the principal effect. These verbs are usually +control,
i.e. the Actor controls the action.

(1) −state, −affected, +control: mi-’atef ’go playing with friends at night’,
mi-’icep ’chew betelnuts’ (cf. ‘icep ’betelnut’), mi-clem ’dive’, mi-cofay ’go
back on the way’, mi-dagoy ’swim’, mi-hapinag ’make clear’, mi-holol ’play’,
mi-lafin ’stay at night’, mi-lcad ’imitate’, mi-loyoh ’cut down sth.; ’mi-odawis
’gather betelnuts’, mi-stol ’hit’, mi-sakalafi ’prepare evening meal’, mi-to’as
’bring up someone’.

Two exceptional verbs appear to be +affected, −control.

(2) −state, +affected, −control: mi-safo’og ’doze’, mi-saker ’snore’.

Both of these verbs include the prefix sa-. This prefix attaches to noun stems and
derives a verb that inflects in the mi conjugation, and has the meaning ‘make
Verb classification in Amis

(noun): mi-sa-loma ‘make a house’ < loma ‘house’. If we regard fo’og and ker as a kind of noun stem, then we can explain the above exceptional cases by saying that this sa- requires these verbs to inflect in the mi conjugation.

ϕ verbs are classified into two. One, ϕ-A, consists of those which express action controlled by the Actor.

(3) pa-digwa ‘make a telephone call’ (cf. digwa ‘telephone’), pa-fli ‘give’, pa-ka-fana ‘teach someone something’, ta-ini ‘come’ (cf. it-ini ‘here’), ta-ira ‘go’ (cf. it-ira ‘there’), tala-lotok ‘go to mountain’ (cf. lotok ‘mountain’), tala-picodadan ‘go to school’ (cf. picodadan ‘school’).

(Morphological causative verbs, which use the prefix pa-, all belong to this category.)

These verbs all include specific prefixes, namely pa-, ta-, and tala-. These prefixes block the addition of the conjugation prefix mi-. In fact, verbs with the prefix pa-can be used both with and without the conjugation prefix mi-. ‘Give’, for example, may be realized as either pa-fli or mi-pa-fli.

11.3.1.2 om: −state, +affected, +control There are very few om verbs. The following lists in (4) and (5) are exhaustive. There are verbs of motion, posture, bodily evacuation, speaking/singing/crying, and some others.

(4) −state, +affected, +control: mo-’sos3 ‘fart (inaudibly)’, mo-’tot ‘fart (audibly)’, m-aro ‘sit down’, m-ayi ‘defecate’, m-isi ‘urinate’, (k-o)m-aen ‘eat’, (l-o)m-oad ‘stand up’, mo-nanaw ‘wash one’s face’, m-ota ‘vomit’, (r-o)m-adiw ‘sing’, (r-o)m-akat ‘walk’, (s-o)m-owal ‘speak, talk’, (t-o)m-agic ‘cry’.

(5) +state, +affected, +control: m-oko ‘live (at a place)’, (t-o)m-ireg ‘be standing’

Most om verbs are characterized as −state, +affected, +control. The principal locus of the effect is the Actor, who controls the situation. Vomiting is sometimes uncontrollable, but one can control it to some extent, so I regard it as basically +control. As for dynamicity, most of them express action, but two stative verbs occur as well.

We can see from Table 11.2 that the semantics of om verbs overlaps that of ma verbs. There are very few om verbs in the language, compared to other related languages (see 11.4.1). This seems to imply that the use of ma has expanded through time. There would have been more om verbs in an earlier stage of Amis, and the om verbs of present Amis have somehow survived the ‘invasion’ of ma, described in the following section.4

3 Mo-’sos ‘fart (inaudibly)’, mo-’tot ‘fart (audibly)’, and mo-nanaw ‘wash one’s face’ appear to include prefix mo-. This probably resulted from metathesis, possibly caused by epiglottal plosive or nasal n at the stem initial position.

4 In Sakizaya, another dialect of Amis, there are mo verbs, which correspond to om verbs in other dialects. In Sakizaya, there are many more verb stems which take the mo prefix than there are verb stems
There are many verbs that inflect in the ma conjugation, and consequently the ma conjugation covers a wide range of semantic fields. They include both +state verbs and −state verbs, and both +control verbs and −control verbs. A shared feature of these verbs, however, is that they all express +affected situations, i.e. a situation in which the subject is the most affected entity, or the situation is caused through being affected by another entity. I will present the examples in four sections, according to stativity and control.

Non-stative controllable verbs

Ma-fkac 'to run', ma-fer 'to fly', and the like are classified in this subcategory. The Actor controls these actions, and he himself is affected the most in the situation. Reciprocal verbs, except those derived from emotional verbs, are also classified in this subcategory. In reciprocal situations Actors control the action and Actors try to affect each other.

(6) −state, +affected, +control: ma-fer 'fly', ma-cakat 'ascend', ma-dakaw 'ride', ma-flec 'return', ma-fkac 'run', ma-galef 'advance', ma-had 'grow', ma-lanam 'have breakfast', ma-lahok 'have lunch', ma-lafi 'have supper', ma-nafoy 'walk on all four, ma-sadak 'go out', ma-sakro 'dance', ma-tawa 'laugh'.

Active reciprocal verbs (−state, +affected, +control): ma-da-dama 'help one another' (cf. mi-dama 'help'), ma-da-do'do 'follow one another' (cf. mi-do'do 'follow'), ma-ka-(kota-)kotay 'take turn one another' (cf. mi-kotay 'take turn'), ma-la-(lica-)licay 'greet one another' (cf. mi-licay 'ask'), ma-la-(loo-)lood 'fight one another' (cf. mi-lood 'go to war'), ma-la-((mo-)l)moh 'meet one another' (cf. mi-limoh 'meet one someone').

Non-stative non-controllable verbs

There are verbs that express non-stative process which the Actor cannot control. Examples are:

(8) −state, +affected, −control: ma-fa'sig 'sneeze', ma-loyoh 'fall down', ma-sowaf 'yawn', ma-tlaaf 'belch', ma-tlook 'hiccup', ma-tolo' 'stumble'.

Weather verbs also are non-stative and refer to non-controllable situations.

(9) −state, +affected, −control: ma-dodem 'be cloudy, dim', ma-fali 'be windy', ma-fetili' 'roll (for thunders)', ma-lonen 'quake (for earth)', ma-'orad 'rain'.

Stative controllable verbs

Some verbs express a stative situation that is controlled by the Actor. The main members are those words that express cognition, emotion, or behaviour patterns. The following verbs are examples.

(10) +state, +affected, +control: ma-'amitir 'be a tomboy', ma-'amlas 'be violent', ma-'iget 'be angry', ma-'inap 'be jealous', ma-apac 'be late', ma-fahka 'be surprised', ma-fana' 'know', ma-faraw 'be in a hurry', ma-fokil which take the om prefix in other dialects. Many verb stems are used with either mi or mo, without any difference in meaning, though there are a few exceptions which show semantic differences.
'be ignorant, not know', ma-foti' 'sleep', ma-froh 'be furious', ma-galay 'want', ma-godo 'be ashamed', ma-hmek 'be happy, admire', ma-kalah 'be in a hurry', ma-lalok 'be earnest', ma-lipahak 'be glad, happy', ma-olah 'like, love', ma-osi 'hate', ma-raraw 'be worried', ma-rarom 'be sad, lament', ma-ri'ag 'be cruel', ma-sawawa 'be childish' (cf. wawa 'child'), ma-talaw 'fear'.

Reciprocal verbs that are derived from emotional verbs are included in this category also. Emotional reciprocal verbs express stative controllable situations.

Examples:

(11) Emotional reciprocal verbs (+state, +affected, +control): ma-sa-osi 'hate each other', ma-la-olah 'like each other'.

Stative non-controllable verbs Verbs of physical states are stative and non-controllable. The states are caused through being affected by another entity. Examples:

(12) +state, +affected, −control: ma-hapinag 'be clear', ma-lcad 'be same', ma-lolay 'get tired', ma-pawan 'forget', ma-raay 'be far', ma-stol 'get a bruise', ma-to'as 'be old'.

11.3.1.4 θ-B: +state, −affected, −control Verbs that belong to θ-B express states that are considered to be inherent or permanent.

(13) +state, −affected, −control: aqtoł 'smell bad', adihay 'be many (things)', aloman 'be many (people)', calfacef 'be foggy', calkah 'rough', dadahal 'be wide', fa'cal 'be good', fa'det 'be hot, be sultry', fa'sis 'smell good', fo'kal 'be white', galay 'be good', ka'so 'be tasty', kalomkam 'be fast, be quick', kohtig 'be black', li'tec 'be cold', malasmas 'drizzle', mamag 'be little', mimig 'be small, be little', po'ner 'be low', rona 'be different', sa'pi 'be cool', ta'gal 'be bright, be clear', tahhay 'be fast, be quick', takaraw 'be high', tanektek 'be strong, be stout', tata'ag 'be big, be large', tatih 'be bad', to'man 'be dark'.

There are several verbs that seem to express temporal or affected state.

(14) +state, +affected, −control: alem'em 'have a muscular ache', cago'ot 'be dissatisfied', rahker 'be satisfied, rihaday 'be happy'.

There are verbs that involve specific prefixes.

(15) ci-gagan 'be famous' (cf. gagan 'name'), ci-tayal 'have work' (cf. tayal 'work'), ci-wawa 'have child' (cf. wawa 'child'), paka-fkac 'be able to run' (cf. ma-fkac 'run'), paka-dagoy 'be able to swim' (cf. ma-dagoy 'swim'), kara-tagic 'cry very often' (cf. t-om-agic 'cry'), kara-tot 'fart very often' (cf. mo-tot 'fart').

Some of the θ verbs contain an -al- infix which is already fossilized: kalomkam 'to be fast, quick', calkah 'to be rough', calfacef 'to be foggy', malasmas 'to drizzle', and alem'em 'to have a muscular ache'.

5 Some of the θ verbs contain an -al- infix which is already fossilized: kalomkam 'to be fast, quick', calkah 'to be rough', calfacef 'to be foggy', malasmas 'to drizzle', and alem'em 'to have a muscular ache'.
11.3.2 Semantic contrast

Many Amis verb stems can inflect in different conjugations, with contrastive meanings. I will refer to such a situation by the term ‘X vs. Y alternation’. There are mi vs. ma alternations, mi vs. om alternations, om vs. ma alternations, and ma vs. ϕ alternations. The semantic contrasts between the different conjugations are predictable from their meaning, which we described in 11.3.1. Let us look at them in turn.

11.3.2.1 Mi vs. ma There are many verb stems that can conjugate in mi and ma conjugations, with contrastive meanings. The contrast may be summarized as follows: mi verbs express an action which is controlled by an Actor and affects another entity, while ma verbs express a state caused through being affected by another entity or an uncontrolled, automatic action. Such contrasts conform to what we saw in 11.3.1:

mi − state, − affected, + control
ma ± state, + affected, ± control

We can observe several types of semantic contrast, which are illustrated in the following examples.

In the following pair, we observe a mi verb mi-patay meaning ‘to kill’—an action that is controlled by an Actor to affect another entity, with its corresponding ma verb ma-patay meaning ‘to be dead’—a state that was caused by another entity.

16) a. mi-patay cira tina tamdaw.
   MI-kill 3SG.NOM this.OBL person
   ‘S/He killed this person.’

b. ma-patay kina tamdaw.
   MA-dead this.NOM person
   ‘This person is dead.’

An Agent or Cause can appear overtly with ma verbs. The Agent or Cause appears in Genitive form.

17) ma-patay nira kina tamdaw.
   MA-dead 3SG.GEN this.NOM person
   ‘This person was killed by him/her.’
   ‘This person is dead because of him/her.’
   ‘S/He killed this person.’

Similar pairs are:

18) mi-adah ‘cure’ vs. ma-adah ‘be cured, recovered’; mi-da’ak ‘make suffocate’ vs. ma-da’ak ‘suffocate’; mi-dodem ‘make dim’ vs. ma-dodem ‘be cloudy’.

It is not always the case that a notional Cause can appear overtly as Genitive NP with ma verbs. It cannot, for example, appear with ma-fana ‘to know’ (example (19c)) or with ma-sadak ‘to go out’.
Verb classification in Amis

(19) a. *mi-fana’ cira tîna ratoh takaanan.*
   MI-notify 3SG.NOM this.OBL news 1SG.OBL
   ‘S/He notified me of this news.’

   b. *ma-fana’ kako tîna ratoh.*
   MA-know 1SG.NOM this.OBL news
   ‘I know this news.’

   c. *‘ma-fana’ nîra kako tîna ratoh.*
   MA-notify 3SG.GEN 1SG.NOM this.OBL news
   ‘S/He has notified me of this news.’

Similar pairs are:

(20) *mi-fana’ ‘notify’ vs. ma-fana’ ‘know’; mi-lcad ‘imitate’ vs. ma-lcad ‘be same’; mi-lolay ‘be tiring’ vs. ma-lolay ‘be tired’; mi-loyoh ‘fall dawn (tr.)’ vs. ma-loyoh ‘get dizzy’; mi-sadak ‘take out (tr.)’ vs. ma-sadak ‘go out’; mi-sinanot ‘prepare’ vs. ma-sinanot ‘be ready’; mi-stol ‘hit (tr.)’ vs. ma-stol ‘get a bruise’; mi-to’as ‘bring up (tr.)’ vs. ma-to’as ‘be old’

In some pairs *mi* verbs express deliberate, controlled action and *ma* verbs automatic, uncontrolled action:

(21) a. *mi-hanhan cira.*
   MI-breathe 3SG.NOM
   ‘S/He breathes consciously (as was told by a doctor, for example).’

   b. *ma-hanhan cira.*
   MA-breathe 3SG.NOM
   ‘S/He breathes.’

In the pair above the most affected entity seems to be the same with the *mi* verb and the *ma* verb, i.e. the one who breathes, because both refer to the process of breathing. Stativity seems to be the same also.

The same can be said about the following pairs. In (22a), (23a), and (24a) with *mi* verbs, the Agent tries to control and affect his body, his word, and the like. The sentences (22b), (23b), and (24b) express automatic or uncontrolled actions or already realized states.

(22) a. *mi-fer cira.*
   MI-fly 3SG.NOM
   ‘He tries to fly. He jumps.’

   b. *ma-fer ko ’ayam.*
   MA-fly NOM bird
   ‘A/The bird flies.’

(23) a. *mi-kori cira.*
   MI-be on a diet 3SG.NOM
   ‘S/He is on a diet.’

   b. *ma-kori cira.*
   MA-skinny 3SG.NOM
   ‘S/He is skinny.’

(24) a. *mi-gilay cira to sowal nîra.*
   MI-speak with an accent 3SG.NOM OBL word 3SG.GEN
   ‘S/He speaks with a dialectal accent on purpose.’
11.3.2 Mi vs. om Some verb stems that otherwise takes om conjugation may also take mi conjugation. Om verbs express actions which mainly affect the Agent, whereas mi verbs express actions that affect others.

(25) a. Mi-load cira to kilag.
   MI-make something stand up 3SG.NOM OBL wooden stick
   ‘S/He made a/the wooden stick straight.’

b. l-om-oad cira.
   OM-stand up 3SG.NOM
   ‘S/He stood up.’

Mi-load has corresponding ma-load as well.

(26) ma-load nira kina kilag.
   MA-make something stand up 3SG.GEN this.NOM wooden stick
   ‘S/He made this wooden stick straight.’ or
   ‘This wooden stick was made straight by him/her.’

Normal crying is expressed by (t-o)m-agic, but ‘to cry for something’ is expressed by its mi counterpart. We can say that this mi verb expresses an action where the Agent tries to influence that something to come into his property. Mi-tagic has a corresponding ma verb as well.

(27) (t-o)m-agic ‘cry’
   mi-tagic ‘cry for something’
   ma-tagic ‘be cried over’

11.3.2.3 om vs. ma Transitive om verbs, k-om-aen ‘to eat’, s-om-owal ‘to speak’, and r-om-adiv ‘to sing’ have their ma counterparts. In such pairs om verbs express actions or processes such that the locus of its principal effect is the subject, which is usually the Agent, and ma counterparts express states of the subject that are realized through the subject’s being affected by something other than the subject. Both express +affected situations and they show a contrast in stativity and control.

Examples:

(28) a. k-om-aen cira to fotig. −state, +affected, +control
   OM-eat 3SG.NOM OBL fish
   ‘S/He eats a fish.’

b. ma-kaen nira ko fotig. +state, +affected, −control
   MA-eat 3SG.GEN NOM fish
   ‘S/He ate a/the fish. A/The fish was eaten by him/her.’

This accords with what we saw in 11.3.1, namely that om verbs express −state, +affected, +control and ma verbs ±state, +affected, ±control.
11.3.2.4 Ma vs. ϕ Some stems inflect in both the ma and the ϕ conjugations. In the ma conjugation these verbs express temporary states implying that the state was caused through the effect of some other entity, and in the ϕ conjugation they express inherent or permanent states, without such implications as are found with the ma inflection. Example (29a) describes a room being dark without implying any cause or background relating to its being dark, while example (29b) indicates that the room is dark because someone put out the light.

(29)

a. to’man kina loma’.
dark this. NOM room
(lit.) This room is dark.
‘It is dark in this room.’

b. ma-to’man kina loma’.
MA-dark this. NOM room
‘(lit.) This room was made dark.’
‘It is dark in this room (because someone put out the light).’

Other examples include ma-fa’det ‘to be hot, warm (because someone heated it)’ vs. fa’det ‘to be hot, warm’.

11.3.3 Ma verb that takes the case frame (GEN C) NOM T

Amis has a Philippine-type focus system (see section 11.2). Sentence (16a) is in AV and has its corresponding GV sentence in (30).

(30) patay-en nira kina tamdaw.
kill-GV 3SG.GEN this. NOM person
‘S/He will kill this person.’
‘S/He definitely killed this person.’

The case frame of (17) is exactly the same as that of the GV sentence (30). I have analysed earlier sentences like (17) as GV. There are some good reasons for this analysis. One is that ma forms with the case frame (GEN C) NOM T does not have corresponding GV form. Several ma verbs have the same case frame as (16a), and they have corresponding GV: e.g. ka-fana’-en. I compare their case frames in Table 11.3. The fact that ma forms with the case frame (GEN C) NOM T do not have a corresponding GV may mean that they themselves are GV forms, possibly of the corresponding mi verbs.

This approach is not the only solution, however. Not all the verbs have corresponding GV, and one can say that ma verbs that have the case frame (GEN C) NOM T are just such stems. It may happen that even a mi verb does not have a corresponding GV form. Mi-clem ‘to dive’, mi-dagoy ‘to swim’, or taira ‘to go’ (the first two are mi verbs and the last one is ϕ verb), for example, do not have a corresponding GV form. They are explained by semantics. We cannot expect any patientive NP to occur with these verbs.
Table 11.3. Case frames of 

<table>
<thead>
<tr>
<th>mi with the case frame nom A (obl P)</th>
<th>ma with the case frame (gen C) nom T</th>
<th>ma with the case frame nom A (obl P)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AV</strong> mi-patay nom A (obl P). A kills P.</td>
<td>ma-patay (gen C) nom T. T is dead (because of C). C killed T.</td>
<td>ma-fana’ nom A (obl P). A knows P.</td>
</tr>
<tr>
<td><strong>GV</strong> patay-en gen A nom P. A will surely kill P.</td>
<td>*ka-patay-en</td>
<td>ka-fana’-en gen A nom P. A surely knows P</td>
</tr>
</tbody>
</table>

If we regard ma verbs as GV, a problem arises. We come to have two GV forms for one mi verb, while others have only one. This imbalance is shown in Table 11.4.

How do ma forms with the case frame (gen C) nom T differ from GV en forms? The presence of a gen C phrase with ma verbs is optional (53 per cent in a text I collected), whereas Agent is rarely absent in GV clauses (86 per cent of GV declarative sentences had Agent phrases). In Cebuano also, Agent is rarely omitted in GF sentences (Shibatani 1988a). Although the proportion of the sentences with ma verbs accompanied by a Cause phrase is still higher than that with passive predicates accompanied by an Agent phrase in English or Japanese, it is significantly lower than the proportion of the GV sentences with an Agent phrase.

Morphologically, ma verbs with the case frame (gen C) nom T and ma verbs with the case frame nom A (obl P) have the same inflectional possibility, whereas en form does not show any, as summarized in Table 11.5.

Ma verbs that take the case frame (gen C) nom T are apparently related to corresponding mi verbs, regardless of whether they are a GV form or passive form. What is the relationship between them? I assume it to be derivational, not inflectional, whereas I assume that the relationship between mi verb and its corresponding en form is inflectional. This new approach explains the fact that such verbs have full inflection of other ma forms, as shown in Table 11.5.

Table 11.4. Imbalance caused by regarding ma form as GV

<table>
<thead>
<tr>
<th>mi with the case frame nom A (obl P)</th>
<th>ma with the case frame nom A (obl P)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AV</strong> mi-patay nom A (obl P). A kills P.</td>
<td>ma-fana’ nom A (obl P). A knows P.</td>
</tr>
<tr>
<td><strong>GV</strong> patay-en gen A nom P. A will surely kill P.</td>
<td>ma-patay (gen C) nom T. T is dead (of C). ka-fana’-en gen A nom P. A surely knows P</td>
</tr>
<tr>
<td>A definitely killed P.</td>
<td>C killed T.</td>
</tr>
</tbody>
</table>
Verb classification in Amis

Table 11.5. Case frames of \(mi\) verbs and \(ma\) verbs

<table>
<thead>
<tr>
<th>Example</th>
<th>(mi) verb</th>
<th>(ma) verb that takes the case frame (Gen C)</th>
<th>(ma) verb that takes the case frame (Nom A)</th>
<th>(\phi) verb</th>
<th>(en) form</th>
</tr>
</thead>
<tbody>
<tr>
<td>AV neutral relativized form</td>
<td>(mi)-patay</td>
<td>(ma)-patay-ay</td>
<td>(ma)-fana'-ay</td>
<td>taira-a-y</td>
<td>patay-en-ay</td>
</tr>
<tr>
<td>AV connegative</td>
<td>(pi)-patay</td>
<td>(ka)-patay</td>
<td>(ka)-fana’</td>
<td>ka-taira</td>
<td>*ka-patay-en</td>
</tr>
<tr>
<td>Instrument</td>
<td>(sa)-(pi)-patay</td>
<td>(sa)-(ka)-patay</td>
<td>(sa)-(ka)-fana’</td>
<td>(sa)-(ka)-taira</td>
<td>*sa-patay-en</td>
</tr>
<tr>
<td>Location</td>
<td>(pi)-patay-an</td>
<td>(ka)-patay-an</td>
<td>(ka)-fana’-an</td>
<td>ka-taira-an</td>
<td>*patay-en-an</td>
</tr>
</tbody>
</table>

This approach also explains the semantic irregularity between \(mi\) forms and \(ma\) forms (see 11.3.2.1). The case frame correspondence is also irregular. Some can take a Cause in Genitive case (example (17)) and some cannot (example (19c)), for example. The semantic relationship and case correspondence between \(mi\) forms and \(en\) forms, on the other hand, is quite regular.

11.4 Comparison

I first compare Amis to other Austronesian languages. Then in 11.4.2 I will compare it to ordinary semantic alignment systems.

11.4.1 Comparison with other Austronesian languages

AV forms are said to have existed already in Proto-Austronesian (henceforth PAn). Adelaar (2005: 6) summarizes:

There are four formal verb classes: (a) verbs which took the actor infix \(<um>\); (b) verbs without \(<um>\) affix in actor voice; (c) verbs consisting of a derived stem beginning with (possibly causative) \(\ast pa\)-; and (d) verbs consisting of a derived stem beginning with (possibly stative) \(\ast ka\) -. Both \(\ast pa\) - and \(\ast ka\) - appeared as \(\ast ma\) - after infixation of \(<um>\).

Amis \(om\) verbs are apparently the descendants of (a) type verbs. In other languages, there are many verbs with cognate morpheme of \(\ast <um>\), but in Amis there are very few such verbs. For example, Seediq, another Formosan language which belongs to the Atayalic group, has a group of verbs that involve infixation of \(em\), the cognate of \(<um>\) in that language (Tsukida 2005). This group has quite a large number of members, and roughly corresponds to Amis \(mi\) and \(om\) verbs. Seediq \(em\) verbs cover both \(-\) affected and \(+\) affected verbs.

Amis \(\phi\) verbs are possibly descendants of (b) verbs. In Amis, they involve specific prefixes.
As for *mi* verbs, we cannot know their origin. It is possible that the *pi* prefix which is present in connegative form appears as *mi* after infixation of *om*, but we do not know where this *pi* comes from either. Amis *mi* verbs semantically correspond to Tagalog *mag* verbs, as shown in the following examples.

(31) Tagalog *mag-tayo* 'erect (sth.)’ vs. *t-um-ayo* ‘stand up’
    Amis *mi-load* ‘erect (sth.)’ vs. *l-om-oad* ‘stand up’.

(32) Tagalog *mag-init* ‘make hot, heat’ vs. *um-init* ‘become/get hot’
    Amis *mi-talem* ‘make sharp’ vs. *ma-talem* ‘be/become sharp’

Tagalog *mag-* is claimed to imply ‘greater frequency’ or ‘intensity of the action’ (Himmelmann 2005b: 365), but Amis *mi* has no such connotation.

*Ma* verbs possibly derive from (c) and (d) verbs, but there are two problems with this hypothesis. One is that *pa-* is said to appear as *ma-* after infixation of *<um>*, but this does not hold for Amis. In present-day Amis, *pa-* the causative prefix, appears without infixation of *<om>*-, but optionally with *mi-*. It is possible, however, that there were two homophonous *pa-* prefixes, one which still productively expresses causative meaning and appears optionally with *mi-*, and one which became *ma-* as the result of infixation of *om* (diachronically) and whose causative meaning is obscured. Another problem is that *ϕ-*B verbs exist in Amis, which certainly involve *ka* in connegative form (see Table 11.1), but without *ma* in neutral form. This means that these verbs did not undergo infixation of *<um>*.

Moreover, the affixation of *ka-* in connegative form is not limited to stative verbs in Amis. It is used with dynamic verbs as well (see 11.3.1.3). Zeitoun and Huang (2000) studied the morpheme *ka-* and its cognates in several Formosan languages. The authors claim that *ka-* is a stative marker, but their data does not include Amis, and their claim does not hold for Amis. In Amis, *ka-* is just a grammatical marker to mark connegative.

Foley (2005) is an investigation of the unaccusative-unergative split across a range of Austronesian languages. Presuming that Tagalog *um* is an unergative marker and that *ma* is an unaccusative marker, Foley makes the following claim:

The system of the Philippine languages probably represents something close to the original type, with two distinct, but crosscutting bases for the unergative-unaccusative split, one by an aspectual type contrast between things that happen (states and changes of state) and things that are done (actions), and the other by the volition, ability, etc or lack of thereof of the instigating participant, the actor (Foley 2005: 425).

Foley says that *ma* appears if the predicate expresses a non-volitional, inabilitative state, and *um* otherwise, i.e. when the predicate expresses either volitional, abilitative verbs or action verbs. Amis *om* verbs share such *um* meanings, but so do many *ma* verbs. As for Amis *ma* verbs, some of them share such *ma* meanings, but there also are many *ma* verbs that express volitional actions.
Foley (2005) explains the Austronesian unergative-unaccusative split, morphologically realized as the -um-/ma- split, in terms of two distinct, cross-cutting bases. The Amis split in verbal morphology is also explained in terms of two distinct, cross-cutting bases (see Table 11.2). For both, one of the bases is stativity, but the other is different. For Amis it is affectedness, and for Foley’s Austronesian unergative-unaccusative split it is volitionality.

The notion of affectedness here is somewhat similar to Foley’s (2005) term ‘undergoes a change in state’, but not exactly the same. ‘Affectedness’ as used in this chapter has a broader sense than Foley’s ‘undergoes a change in state’. In the sense of the term as employed here (see 11.3.1), it is possible that a volitional actor is affected at the same time, but it seems impossible in Foley’s scheme. In his macro-role hierarchy (Foley 2005: 391), volitionality is at the top of the actor pole and ‘undergoes a change in state’ is placed at the extreme of the opposite undergoer pole. It is difficult to adopt Foley’s hierarchy in the explanation of Amis data.

11.4.2 Comparison with ordinary semantic alignment systems
At first sight, the Amis system of verb classification seems to have little in common with ‘ordinary’ semantic alignment systems (e.g. Mithun 1991). The Amis system does not involve either case markers or marking on the verb for subject and object. If we look at the system closely, however, interesting analogies emerge.

Many transitive stems of Amis appear in both the ma and mi conjugations, with contrastive meanings and case frames (see 11.3.2). The ma conjugation is used when the notional Patient is the subject, and the mi conjugation is used when the notional Agent is the subject. Most of the intransitive stems take either mi conjugation or ma conjugation in Amis. This is strikingly analogous to P-marking and A-marking in ordinary semantic alignment systems. Transitive verb stems in Amis can take either of these markings depending on which argument the subject is, but intransitive verbs can only take one. In ordinary semantic alignment systems, intransitive verbs can be classified according to which marking they take. In Amis, intransitive verbs are classified according to which conjugation they take.

While most semantic alignment systems show a twofold division of intransitive verbs, i.e. A-marked and P-marked, the Amis system is fourfold. It is not limited to the number of case markers. I have to admit that the analogy of A and P marking can be applied only partially, but the fact that it can be applied to a part of an unrelated system is nevertheless interesting.

There are other languages with three or more classes of intransitive verbs. Muskogean languages such as Chickasaw have three classes of intransitive verbs: A-marked verbs, P-marked verbs, and Dative-marked verbs, which mark the subject as a sentient experiencer (Foley 2005: 427). In the Papuan language Saweru verbs are classified into three classes based on inflection (Donohue 2001, 2004b).
### Table 11.6. Traditional split intransitivity and verb classification in Amis

<table>
<thead>
<tr>
<th></th>
<th>Intransitive</th>
<th>Transitive</th>
<th>Intransitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amis</td>
<td>$mi$-verb</td>
<td>$mi$-verb</td>
<td>$ma$-verb</td>
</tr>
<tr>
<td></td>
<td>NOM S</td>
<td>NOM A (obl P)</td>
<td>NOM A (gen A) NOM P</td>
</tr>
<tr>
<td>Ordinary semantic alignment systems</td>
<td>A-verb$^a$</td>
<td>A-verb-P</td>
<td>NOM S</td>
</tr>
<tr>
<td>Semantics</td>
<td>−state,</td>
<td></td>
<td>+state,</td>
</tr>
<tr>
<td></td>
<td>−affected,</td>
<td></td>
<td>+affected,</td>
</tr>
<tr>
<td></td>
<td>+control</td>
<td></td>
<td>−control</td>
</tr>
</tbody>
</table>

$^a$ This table involves only Amis $mi$ and $ma$ verb forms of Amis. Amis $ma$ verbs with non-A obl P case frame are also omitted.

$^b$ My use of this notation reflects the parallel that I see with ordinary semantic alignment systems; I do not intend to limit the use of the term to such a language where A and P are dependently marked on verbs, A by prefix and P by suffix.

Donohue (2001) points out the commonality between the Saweru and Chickasaw systems. In Icelandic there is a four-way split in the marking of Ss. Donohue (this volume) has pointed out the similarity of the Icelandic system of non-canonical marking for subjects and objects (Andrews 2001). So we can see that the Amis situation, while extreme, is not so strange in a typology of semantic alignment.

What is really interesting is that the same semantic features that have been used to describe ordinary semantic alignment systems, that is, ±state, ±affectedness and ±control, are valid in describing the classification in Amis. Amis $mi$ verbs show features similar to that of ‘A-marked verbs’ in ordinary semantic alignment systems, covering ‘−state, −affected, +control pole’, and Amis $ma$ verbs show features similar to that of ‘P-marked verbs’, covering ‘+state, +affected, −control pole’. The above discussion is summarized in Table 11.6.

### 11.5 Summary

This chapter introduced verb classification in Amis and explained it in terms of semantic features. I employed stativity and affectedness in order to explain the classification. $Mi$ verbs and $\psi$-A verbs are characterized as −state, −affected, $om$ verbs −state, +affected, $ma$ verbs ±state, +affected, and $\psi$-B verbs +state, −affected (11.3.1). Many Amis verb stems can inflect in different conjugations, with contrastive meanings. We have seen that the semantic contrasts observed are consistent with the above characterization (11.3.2).

The Amis situation was subsequently compared with other Austronesian languages and with ordinary semantic alignment systems. I pointed out several interesting points of difference between my findings in the Amis situation and previous studies on Austronesian languages (11.4.1). I also pointed out intriguing points,
some different from and some analogous to the ordinary semantic alignment systems (11.4.2).

There is more than one conjugation type for AV while there is only one for each of the other voices. After having analysed the semantics of each AV conjugation, it is clear that how the Actor is involved in the situation, whether it is affected or not, is an important matter in Amis.
Part IV
The Americas
12

The emergence of agentive systems in core argument marking

MARIANNE MITHUN

It has been proposed that core argument categories (sometimes referred to as ‘alignment’) are highly stable over time and strongly resistant to borrowing (Nichols 1992: 181). The proposal certainly seems reasonable. Categories such as subjects and objects, ergatives and absolutes, or agents and patients are often deeply embedded in the grammar, encoded morphologically by case markers or pronominal shape. They are typically distinguished in every clause. Evidence pointing to stability can be found in a number of language families. In North America, for example, all languages of the Iroquoian, Caddoan, and Siouan families, each with a probable time depth of thousands of years, show clear agentive patterns. The agentive patterns may be older than the families themselves: it has been hypothesized that these three families might be remotely related to each other, with a common parent older than Proto-Indo-European (Latham 1860, Chafe 1976: 43–53).

Parallels in core argument categories have even been taken as diagnostic of deep genetic relations. During the first half of the 20th century, when Edward Sapir was attempting to group the recognized language families of North America into larger superstocks, he made just such a suggestion to Alfred E. Kroeber.

(1) Sapir 1920 letter to Kroeber (Golla 1984: 349; emphasis added)

The interest [of polysynthesis] is psychological rather than historical. To me, it is worth less than such an obscure feature as prevalence of stems with initial vowel…. or classification of pronouns into transitive and intransitive vs. active and static (D [Penutian] is characterized by the former, F [Hokan-Siouan] by the latter).

F. Hokan-Siouan
   I. Hokan proper
      (Shasta down to Seri, Chontal)
   II. Yuki
   III. Coahuiltecan
   IV. Keres

V. Siouan-Yuchi
   1. Siouan
   2. Muskogi-Natchez
   3. Yuchi
   4. Tunica-Chitimacha-Atakapa
A number of the languages grouped together by Sapir in his Hokan-Siouan super-stock show agentive or active/stative argument patterns, though certainly not all. It should be noted that the hypothesis of a Hokan-Siouan superstock has since been abandoned, and the existence of a Hokan group remains controversial (Goddard 1996: 308–16, Campbell 1997: 290–305, Mithun 1999: 303–4).

Here we will examine the stability of core argument categories by considering the distribution of agentive systems in North America. Such systems are relatively rare cross-linguistically. Nichols (1992: 187) found that nominative/accusative or neutral systems occurred in 65 per cent of the 172 languages in her genetically and areally balanced sample, ergative systems in 19 per cent, and agentive or stative/active systems in just 14 per cent. Within North America, however, she found agentive systems in nearly a third (31 per cent) of the languages sampled. North America is characterized by great genetic diversity, with around 300 languages grouped into 58 distinct language families in the traditional sense—i.e. the largest genetic units considered established on the basis of the comparative method. It also contains some strong linguistic areas, particularly in California, on the Northwest Coast, and in the Southeast. It thus provides a good testing ground for the hypothesis of genetic stability and resistance to areal effects.

12.1 Northern California I

California alone shows substantial diversity: 22 of the 58 language families north of Mexico are represented within its borders. Agentive systems occur in two areas in Northern California.

12.1.1 Yuki

Yuki consists of three closely related languages or dialects: Coast Yuki, Yuki Proper, and Huchnom (Mithun 1999: 574–6). All show the same core argument categories. Material cited here comes from Kroeber (1911), Sawyer and Schlichter (1984), and Schlichter (1985). Forms from Kroeber are identified by the initial K. Forms from Sawyer and Schlichter are identified by the initials of the speakers: Arthur Anderson (A) and Minnie Fulwider (F). Numbers following the initials refer to the pages on which they are cited.

In Yuki, the grammatical roles of participants are expressed primarily in independent pronouns.

(2) Yuki pronouns

\[ \text{miT} \quad \text{ma:} \quad \text{e:} \quad \text{we} \quad \text{F} \]  
\[ \text{2SG} \times \text{SG} \quad \text{fish} \quad \text{give-PAST} \]

‘You gave me a fish.’

There are three sets of pronouns.
(3) Yuki pronoun forms

<table>
<thead>
<tr>
<th>Set I</th>
<th>Set II</th>
<th>Set III</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>?ap</td>
<td>?i</td>
</tr>
<tr>
<td>2SG</td>
<td>mel/mi</td>
<td>mis</td>
</tr>
<tr>
<td>1EXCL.PL</td>
<td>?us</td>
<td>?usq</td>
</tr>
<tr>
<td>1INCL.PL</td>
<td>mi/me</td>
<td>miy/mey</td>
</tr>
<tr>
<td>2PL</td>
<td>mo?os/mos</td>
<td>mo?osiyq</td>
</tr>
</tbody>
</table>

Examples of the uses of each set are in (4), (5), and (6).

(4) Yuki Set I pronouns

?ap  ?ap lis k’an la?aktek  ‘I talked fast.’ A111
1SG  kawaye ?ap munmek  ‘I’m going to steal the horse.’ A111
mel?  ?impis me? ko’omha?  ‘Where do you come from?’ A244
2SG  mi?  ?i: hqw cænewe  ‘You gave me a fish.’ F244

(5) Yuki Set II pronouns

?i:  kima-se ?i: hqw cænek  ‘They gave me a fish.’ F133
1SG  kima-se ?emotmilmek  ‘He’s going to gamble with me.’ F133
mis  kima-se mis hqw cæmek  ‘They’ll give you a fish.’ F244
2SG  ?iye miyuctha  ‘What’s the matter with you?’ F244

(6) Yuki Set III pronouns: obliques and possessives

?i:t  ?i:tu?i ki?la?aktek  ‘He left before me.’ A133
metaq?  metaq?woyol’ ?oct hæc.  ‘Smoking is bad for you.’ A244

As in many languages, 3rd person pronouns are relatively infrequent in speech: continuing 3rd person topics are not re-identified pronominally in every clause.

(7) Unmentioned 3rd persons

a  hucmil  pat-ek.  A24
    frybread bake-decl
    '(She)’s making frybread.'

b  me?  t’uk tal-tel  A106
    you hit neg-imperative
    ‘Don’t hit (him)’!

The 3rd person pronouns are used for emphasis or contrast. They are derived from the demonstrative ki? ‘that, there’.

(8) Yuki basic 3rd person forms

<table>
<thead>
<tr>
<th>Set I</th>
<th>Set II</th>
<th>Set III</th>
</tr>
</thead>
<tbody>
<tr>
<td>3SG</td>
<td>ki?</td>
<td>ki?q</td>
</tr>
<tr>
<td>3PL</td>
<td>kimas:se kimas:q kimas:et/kimas:qt</td>
<td></td>
</tr>
</tbody>
</table>
The Set I pronouns might appear at first to represent subjects and the Set II pronouns objects. But some participants that English speakers would encode as subjects are referred to with Yuki Set II pronouns, with both transitive and intransitive predicates.

(9) Yuki Set II with intransitives

\[
\begin{align*}
\text{hilyu} & : 'I'm sick' & F187 \\
\text{mis} & : 'you're sick' \\
\text{ki?} & : 'he's sick' \\
\text{usa} & : 'we're sick'
\end{align*}
\]

\[
\begin{align*}
\text{hamlo?ote} & : 'I'm hungry' & A109 \\
\text{hapway k'i?ip?te} & : 'I'm too full of food' & A88 \\
\text{t'olo?ote} & : 'I feel ashamed' & A21 \\
\text{mika?la?mit?e} & : 'I am unlucky, have bad luck' & F24 \\
\text{k'an'k t'i?i} & : 'I have a sore knee' & A17 \\
\text{şaq şašek} & : 'I've got a toothache' & A17 \\
\text{lo?ca?i?la?ote} & : 'I am phlegmatic' & A159 \\
\text{mat ha?kiml?e} & : 'I can't shoot' & A133 \\
\text{hu?šle} & : 'I'm happy' & F102 \\
\text{'pinhacmek} & : 'I'm sleepy' & A190 \\
\text{šul} & : 'I have a fever' & F37
\end{align*}
\]

(10) Yuki Set II with transitives.

\[
\begin{align*}
\text{mis} & : 'I hear you.' & A104 \\
\text{mis} & : 'I love my father.' & F79 \\
\text{nak} & : 'I'm afraid of the dark.' & A59
\end{align*}
\]

Sets I and II do not represent subjects and objects after all. From the examples seen so far, we might hypothesize that this is an active/stative system, in which the choice between the two pronominal sets is triggered by aspect or Aktionsart. Set I pronouns appear with activities and other events ('talk', 'steal', 'come', 'give', 'gamble', 'leave', 'hit'), while Set II pronouns appear with states ('be sick', 'be hungry', 'be full of food', 'be ashamed', 'be unlucky', 'have a sore knee', 'have a toothache', 'be phlegmatic', 'be unable to shoot', 'be happy', 'be sleepy', 'have a fever', 'love', 'be afraid of'). Further examples show, however, that aspect is not criterial. Set I pronouns also appear with states.
Agentive systems in core argument marking

(11) Yuki Set I pronouns with states: ʔap 'I'

mul huyap ʔap mehek  I'm in the middle of the creek.  F26
ʔap ʔtamap noʔohék  I'm living on the mountain.  A111
ká ʔap/onk'e sat'akah ʔap  I'm sitting here on the cold ground.  F188

Set II pronouns also appear with actions and other events.

(12) Yuki Set II pronouns with events: ʔ: 'I'

a. ʔ: i: a:nį:ssuyitwìk  I fell down  F78
ʔ: k'aptek  I choked  F44
ʔ: poʔok  I got burned  F37
ʔ: yotom hitlek  I burped  A38
ʔ: yičmik  I tremble  K
ʔ: k'awtek  I yawned  A243
ʔ: ča:htek  I sweated  A133
ʔ: nahumtek  I forgot  A87
ʔ: hilyu:tek  I got sick  F187
ʔ: hwilumek  I am getting tired  A219

b. ʔ: kwe: naʔamt-ek  A28
1SG.II knife lose-decl
'I lost my knife.'

c. ko:pe meʔa  ʔ: k:i:pt-ek  A71
coffee too.much 1SG.II overconsume-decl
'I drink too much coffee.'

The categories are based on the semantic roles of core participants, forming what can be termed an agent/patient or agentive pattern. The Set I pronouns represent grammatical agents, participants who instigate or control events like 'talk' or 'steal', or states like 'be sitting' or 'be living (somewhere)'. The Set II pronouns represent grammatical patients, those who are not in control but are significantly affected. They may perform actions ('belch', 'yawn'), but they do not control them. A basic pronoun set is lexicalized with each verb. Connected speech shows little evidence of on-line judgments about degrees of control or affectedness.

Speakers can, however, choose not to portray affectedness. The grammatical agent pronouns (Set I) are the unmarked, default choice. Examination of the textual and dictionary material that exists indicates that control or affectedness is more often specified for 1st persons (or for 2nd persons in questions) than for third. This prevalence is not entirely surprising. Speakers might be expected to have a keener sense of what they themselves are feeling, of their own control and affectedness, and more right to express it.
302  Marianne Mithun

Yuki choices about portraying affectedness

<table>
<thead>
<tr>
<th>Yuki</th>
<th>English</th>
<th>Transcription</th>
</tr>
</thead>
<tbody>
<tr>
<td>?i: lohwitšek</td>
<td>I (patient) coughed’</td>
<td>A53</td>
</tr>
<tr>
<td>(kiʔ) lohtek</td>
<td>he (agent) coughed’</td>
<td>A53</td>
</tr>
<tr>
<td>wok ?i: nαnklαmek</td>
<td>I (patient) learned to dance’</td>
<td>A124</td>
</tr>
<tr>
<td>(kiʔ) nαnkal bɔ́</td>
<td>he (agent) learns slowly’</td>
<td>A124</td>
</tr>
</tbody>
</table>

Case is also marked on some lexical nouns in Yuki, all of them referring to human beings or personified animals, such as ‘woman’, ‘my father’, ‘Wailaki’ (a neighbouring group), ‘Jackrabbit’, and ‘Coyote’, the trickster protagonist of numerous tales (Kroeber 1911: 355, 375–6, Schlichter 1985: 57). The case marking on nouns also follows an agentive pattern, with grammatical agents unmarked and grammatical patients carrying the suffix -q. The alternation between unmarked agent nouns and suffixed patient nouns can be seen in the extracts below from a text by comparing the nouns referring to Coyote. When Coyote is in control, allowing himself to be found or watching, the noun is unmarked. When he is not in control, moving involuntarily to snap or hearing something, the patient suffix -q appears.

(14) Yuki text extract (Kroeber 1911: 375–8)

a. sαéi haye ki hulk’oí eit haye t’q-sa-tl-mil.
   and now that Coyote now find-caus-refl-past
   ‘And now Coyote made himself be found. . . .’

b. sαéi hulk’oí náu-mil.
   and Coyote watch-past
   ‘And Coyote watched. . . .’

c. hulk’o-a muk-law-et-mil
   →Coyote-pat snap-move involuntarily-past
   ‘And Coyote involuntarily moved to snap it.’

d. sαéi aŋa hulk’o-a hqwlt-mil.
   →and again Coyote-pat hear-past
   ‘And again Coyote heard it.’

Transitivity is not necessarily marked on the verb. Many of the same verb forms are used for both intransitives and transitives.

(15) Yuki transitivity: nqwh ‘look, see, watch’

a. nαtpit-qm ʔap nqwh-ek.  F182
   window-loc 1sg.agt see-decl
   ‘I’m looking through the window.’

b. ʔap nqwh-ek.  F182
   house 1sg.agt see-decl
   ‘I see the house.’
Agentive systems in core argument marking

(16) Yuki transitivity: haʔah ‘be, get on, pick up, carry, take, pack’
   a. sahol’ ʔol-ap haʔah-ek.
      eagle  tree-on be-DECL
      ‘The eagle is on the tree.’
   b. haʔah-ek
      carry-DECL
      ‘He’s carrying it in his arms.’

12.1.2. Wider genetic relations
Yuki has been grouped genetically with one other language, Wappo (Radin 1929: 7, Elmendorf 1968, 1981). (Sawyer 1980 notes that at least some resemblances between the two may be due to contact.) The languages do show striking similarities in pronominal shape.

(17) Similarities in some pronoun shapes (Radin 1929: 134, Elmendorf 1981)

<table>
<thead>
<tr>
<th></th>
<th>Yuki Set I</th>
<th>Yuki Set II</th>
<th>Wappo Set I</th>
<th>Wappo Set II</th>
</tr>
</thead>
<tbody>
<tr>
<td>1sg</td>
<td>?ap</td>
<td>?i</td>
<td>?ah</td>
<td>?i</td>
</tr>
<tr>
<td>2sg</td>
<td>miʔ</td>
<td>mis</td>
<td>miʔ</td>
<td>mi</td>
</tr>
<tr>
<td>1excl.pl</td>
<td>?us</td>
<td>?usq</td>
<td>isi</td>
<td>isa</td>
</tr>
<tr>
<td>1incl.pl</td>
<td>mi/me</td>
<td>miyʔ/mayʔ</td>
<td>m?si</td>
<td>m?si</td>
</tr>
<tr>
<td>2pl</td>
<td>moʔos/mos</td>
<td>moʔosiyʔ/m?si/m?si</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In Wappo as in Yuki, 3rd person pronouns have been formed from demonstratives, a process that apparently took place independently in the two languages, since the demonstrative bases are different. (The 3r pronouns in both languages are used for 3rd person arguments that are coreferential with the subject of their clause or a higher clause.)

(18) 3rd person pronouns from different demonstratives

<table>
<thead>
<tr>
<th></th>
<th>Yuki Set I</th>
<th>Yuki Set II</th>
<th>Wappo Set I</th>
<th>Wappo Set II</th>
</tr>
</thead>
<tbody>
<tr>
<td>3sg</td>
<td>kiʔ</td>
<td>kiʔʔ</td>
<td>tsepdfʔte</td>
<td></td>
</tr>
<tr>
<td>3pl</td>
<td>kima:se</td>
<td>kima:sqʔ</td>
<td>tsēkoti, őni</td>
<td>tsēkoto, on-</td>
</tr>
<tr>
<td>3r</td>
<td>kip</td>
<td>kipʔ</td>
<td>me</td>
<td></td>
</tr>
</tbody>
</table>

But the Wappo pronouns represent subjects and objects, not agents and patients.

(19) Wappo subject pronoun ałę ‘I’ (Radin 1924, 1929)

- a’-uluʔat:a ‘I divided it’ (1929: 103)
- a’-uowátita ‘I hit it’ (1929: 103)
- a’-teléuše ‘I come and get it’ (1929: 32)
- a’-našd’ise ‘I jump around’ (1929: 37)
The Yuki and Wappo pronouns are thus similar to each other in shape but not in function. This fact suggests that the pronouns themselves are a common inheritance, but that the core argument pattern has shifted since the two languages diverged.

### 12.1.3 The Pomoan family

A look at the geographical location of Yuki provides further evidence that core argument categories may not be immune to change. Immediately to the south of Yuki are the Pomoan peoples. Figure 12.1 shows the languages of Northern California. San Francisco Bay appears at the bottom of the map.

The Pomoan family consists of seven mutually unintelligible languages. Each contains three sets of independent pronouns. The inventory of forms in Central Pomo in (21). Material cited here comes from speakers Frances Jack, Eileen Oropeza, and Florence Paoli (p.c.).

### 21 Central Pomo pronouns

<table>
<thead>
<tr>
<th></th>
<th>Set I</th>
<th>Set II</th>
<th>Set III</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>?a:</td>
<td>ţo:</td>
<td>k'ë</td>
</tr>
<tr>
<td>2SG</td>
<td>ma</td>
<td>mto</td>
<td>mk'ë</td>
</tr>
<tr>
<td>3SG</td>
<td>mu:l</td>
<td>mú:tú</td>
<td>mú:k'ë</td>
</tr>
<tr>
<td>3SG,R</td>
<td>ţí:</td>
<td>ţí:to</td>
<td>ţí:k'ë</td>
</tr>
<tr>
<td>1PL</td>
<td>ya</td>
<td>yal</td>
<td>yá:k'ë</td>
</tr>
<tr>
<td>2PL</td>
<td>máya</td>
<td>máyal</td>
<td>máya:k'ë</td>
</tr>
<tr>
<td>3PL</td>
<td>mú:túya</td>
<td>mú:túyal</td>
<td>mú:túya:k'ë</td>
</tr>
<tr>
<td>3PL,R</td>
<td>tiya</td>
<td>tiyal</td>
<td>tiya:k'ë</td>
</tr>
</tbody>
</table>

As in Yuki, the Set III pronouns represent possessors and obliques. The basic 3rd person forms are recent developments from the demonstrative mu:l 'that.' Other Pomoan languages show different 3rd person pronominal forms. Continuing 3rd person topics are not usually re-identified in every clause with a pronoun.
Figure 12.1. Northern California (from Heizer 1978: ix)

(22) Central Pomo continuing 3rd person topics (Frances Jack, speaker p.c.)
Me:n-da bal q'á 'el bé-č'.
such-at this water the hold-INCEPTIVE.PFV
'During this time, (she) picked up the water.
Mi: bé-m.
there hold-ESSIVE
(She) held (it) there.'
Marianne Mithun

At first glance, Set I pronouns might appear to be subjects and Set II objects.

(23) Central Pomo independent pronouns
   a. ʔa: čá-č-
       run-semelfactive-pfv
       'I ran away.'
   b. ʔa: múč u ʔéy=čadi-w.
       him away=chase-pfv
       'I chased him away.'
   c. Muːl tóː ʔéy=čadi-w.
       he me away=chase-pfv
       'He chased me away.'

Yet as in Yuki, both Set I and Set II pronouns appear where English speakers would use subject pronouns.

(24) Central Pomo intransitives
   a. Set I ʔaː: 'I'  b. Set II ʔaː: 'I'
      ʔaː wiːq’ʔleː 'I would go'
      ʔaː pó’diclaw 'I dove in'
      ʔaː sbiː 'I got up'
   c. Set I ʔaː: 'I'  d. Set II ʔaː: 'I'
      ʔaː má:báhʔc’ in 'I’m conceited'
      ʔaː le bane:ʔt’aw 'I’m lazy'
      ʔaː qá:ʔc’ 'I fell'
      ʔaː qá:ʔc’ 'I fell'
      ʔaː qá:ʔc’ 'I fell'

Set I pronouns are used with actions like ‘go’, ‘dive’, and ‘get up’ above, while Set II pronouns are used with states like ‘be sick’, ‘be tired’, and ‘be ticklish’. The distribution could suggest an active/stative pattern. Set I pronouns also appear with states, however, and Set II pronouns also appear with events.

(25) Central Pomo Set I with states
(26) Central Pomo Set II with events
   ʔaː ꞌnáːc’ ‘I’m hiding’
   ʔaː ꞌnáːc’ ‘I’m hiding’
   ʔaː má:báhʔc’ in ‘I’m conceited’
   ʔaː má:báhʔc’ in ‘I’m conceited’
   ʔaː ꞌt’aw ‘I’m lazy’
   ʔaː ꞌt’aw ‘I’m lazy’
   ʔaː ꞌtaw ‘I’m lazy’
   ʔaː ꞌtaw ‘I’m lazy’
   ʔaː ꞌt’aw ‘I’m lazy’
   ʔaː ꞌt’aw ‘I’m lazy’
   ʔaː ꞌt’aw ‘I’m lazy’
   ʔaː ꞌt’aw ‘I’m lazy’

Both sets appear with intransitives as above, and with transitives, as below, so the system is not ergative. Transitivity is not overtly marked.

(27) Central Pomo transitives
    a. ʔa: qó 约谈 el bé-č-
       ISG.AGT water the hold-inceptive-pfv
       'I picked up the water.'
    b. ʔaː ꞌmú:tu ꞌʔyáq-an.
       ISG.PAT COP=3.PAT mentally-know-ipfv
       'I remember her.'

The system distinguishes grammatical agents (Set I) and patients (Set II). Set I pronouns typically refer to participants who instigate, perform, and control events.
Agentive systems in core argument marking

and states. Usually these three features coincide. When they do not, the feature of control is criterial. A sneezer performs the sneeze but is not in control. Central Pomo speakers use Set II pronouns for sneezers.

(28) Central Pomo performance without control: Set II

\[
\begin{align*}
to: & \quad \text{Pesya} \quad \text{I sneezed} \\
\text{to:} & \quad \text{Téq'ya} \quad \text{I choked} \\
\text{t:} & \quad \text{Téq'} \quad \text{I burped}
\end{align*}
\]

A few verbs can appear with either set of pronouns. These represent separate lexicalizations. For most verbs, speakers have no choice.

(29) Central Pomo lexical doublets

\[
\begin{align*}
\text{A:} & \quad \text{Smá mt.í:ˇc'ya} \quad \text{I went to bed} \\
\text{A:} & \quad \text{Smá mt.í:ˇcka} \quad \text{I must've fallen asleep} \\
\text{Al:} & \quad \text{C:ım ném} \quad \text{I rammed it, ran into it (intentionally)} \\
\text{Al:} & \quad \text{C:ım ném} \quad \text{I bumped into it} \\
\text{A:} & \quad \text{K'lú:k'luw} \quad \text{I coughed} \\
\text{A:} & \quad \text{K'lú:k'luw} \quad \text{I coughed (involuntary)}
\end{align*}
\]

Participants are cast as grammatical patients only when they are significantly affected by the situation and the speaker chooses to empathize with the affectedness. Inanimate objects are never referred to with patient forms, nor are most animals. The agent forms are a default choice.

(30) Central Pomo empathy

\[
\begin{align*}
a. \quad \text{Mu:t} & \quad \text{A:} \quad \text{Hk'úm.} & \quad \text{3SG.AGT} \quad \text{1SG.AGT} \quad \text{Kill.PFV} \\
\text{b.} & \quad \text{Mu:t} & \quad \text{A:} \quad \text{Hk'úm.} & \quad \text{3SG.PAT} \quad \text{1SG.AGT} \quad \text{Kill.PFV} \\
& \quad \text{'I killed it.' (a fly)} & \quad \text{'I killed him.' (a person)}
\end{align*}
\]

Speakers may choose not to express the affectedness of other persons, or even, on occasion, of themselves, when it is not considered relevant.

Some nouns referring to certain people, particularly some kinsmen, have patient case forms. This is reminiscent of the situation in Yuki, where only some animates, generally humans and personified animals, have patient case forms. The shape of the Central Pomo patient case ending is entirely different from that in Yuki, however: -I. Example (31b) below shows that the noun 'woman' does not have a patient case form, but the noun 'wife' does.

(31) Central Pomo patient case on nouns (Frances Jack, speaker p.c.)

\[
\begin{align*}
a. \quad \text{K'he bák:iyiya-I} & \quad \text{A:} \quad \text{Hčel-an.} & \quad \text{My father.in.law-PAT 1SG.AGT Seek.PFV} \\
& \quad \text{'I'm looking for my father-in-law (PATIENT).'}
\end{align*}
\]

b. \quad \text{Mú:khet' mézi'ta} \quad \text{J'aldá:w, miyá:} \quad \text{dáqate-I} \\
\quad \text{3.POSS woman die 3.KIN.POSS wife-PAT} \\
\quad \text{'His woman died, his wife (PATIENT).'}
12.1.4 The Yuki and Pomoan systems

No genetic relationship has ever been suggested between the Yuki and Pomoan groups, nor is one advanced here. Northern California is a well-known linguistic area, however. Communities were small, and exogamy and multilingualism were the norm. Similarities in shape suggest that Pomoan 1st and 2nd person singular pronouns may have been borrowed into Yuki and/or Wappo. Suggestive similarities are underlined below.


<table>
<thead>
<tr>
<th></th>
<th>1.SG.AGT</th>
<th>1.SG.PAT</th>
<th>2.SG.AGT</th>
<th>2.SG.PAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Northern</td>
<td><code>ha</code></td>
<td><code>to</code></td>
<td><code>ma</code></td>
<td><code>mi-to</code></td>
</tr>
<tr>
<td>Central</td>
<td><code>ha</code></td>
<td><code>to</code></td>
<td><code>ma</code></td>
<td><code>mo</code></td>
</tr>
<tr>
<td>Kashaya</td>
<td><code>ha</code></td>
<td><code>to</code></td>
<td><code>ma</code></td>
<td><code>to</code></td>
</tr>
<tr>
<td>Southern</td>
<td><code>ha:to</code></td>
<td><code>ha:to</code></td>
<td><code>ha:ma</code></td>
<td><code>mi:-to</code></td>
</tr>
<tr>
<td>Eastern</td>
<td><code>ha</code></td>
<td><code>wi</code></td>
<td><code>ma</code></td>
<td><code>mi</code></td>
</tr>
<tr>
<td>Northeastern</td>
<td><code>ha</code></td>
<td><code>ha:-to</code></td>
<td><code>ha:-ma</code></td>
<td><code>mi:-to</code></td>
</tr>
<tr>
<td>Southeastern</td>
<td><code>ha</code></td>
<td><code>wi:ti</code></td>
<td><code>ma</code></td>
<td><code>ti</code></td>
</tr>
<tr>
<td>Yuki</td>
<td><code>ap</code></td>
<td><code>ti</code></td>
<td><code>mi</code></td>
<td><code>mi-s</code></td>
</tr>
<tr>
<td>Wappo</td>
<td><code>ap</code></td>
<td><code>ti</code></td>
<td><code>mi</code></td>
<td><code>mi</code></td>
</tr>
</tbody>
</table>

Sawyer and Schlchter (1984: 280) identify the Yuki element -p as human singular and -s as patient case. Other pronouns show quite different shapes.

The Yuki and Pomoan systems are similar in several ways. Both consist of independent pronouns that distinguish grammatical agents, patients, and obliques. Only a few nouns, all referring to persons, carry patient case suffixes. Grammatical agents and grammatical patients appear in both intransitive and transitive clauses, though transitivity is not necessarily specified overtly, and in both events and states. The criterial features for case choice are control and significant affectedness. Speakers may opt not to portray affectedness for humans, and it is never specified for non-humans. Both systems are agentive, or more specifically patientive, since the patient case is semantically marked. Third person pronouns, recently formed from demonstratives, are usually omitted for continuing topics.

The similarities in the finest details are too close, particularly given their relative rarity cross-linguistically, to be due to chance. It appears that the Pomoan system, which is fully developed in all seven members of the family, was borrowed into Yuki. But how could such a fundamental and abstract system be borrowed, particularly without the full paradigms of pronouns that carry it? Because earlier stages of the languages are unattested, we cannot know for certain, but a likely path of development is easy to imagine.

In a language in which (i) intransitive and transitive verbs are not distinguished formally, (ii) topical 3rd persons are usually not mentioned, (iii) few nouns are marked for case, and (iv) word order is predicate-final, it would be a simple matter to reanalyse a nominative/accusative system as an agentive one or vice versa. Reanalysis could take place along the lines below. Transitive clauses with
omitted 3rd person subjects could be reanalysed as intransitive, and objects could be reanalysed as grammatical patients.

(33)  (subject) object  transitive verb  ‘It scared me (object)’ >

      me  scared

      patient  intransitive verb  ‘I (patient) was scared’

Both Pomoan and Yuki exhibit all of the features that would facilitate such a reanalysis. It is not impossible that the development of the pattern took place independently in Pomoan and Yuki, but the strong bilingualism in the area and the cross-linguistic rarity of such systems suggest that the parallelism in their systems is due to reanalysis stimulated by contact.

12.2  Northern California II

The Yuki, Wappo, and Pomoan systems involve free pronouns. As independent words, the pronouns would be more accessible to speaker consciousness and analysis than affixes. Agentive patterns are much more commonly found in bound pronominal affixes, however. Agent and patient pronominal affixes appear in another part of Northern California, also known as an especially strong linguistic area characterized by long-standing patterns of intermarriage, cultural borrowing, and multilingualism.

12.2.1  Karuk

At the northern edge of California is the Karuk language, also visible in Figure 12.1 above. It is considered an isolate, though proposals have been made linking it remotely with a variety of other languages and families in a larger group termed Hokan. An early description of the language is in Angulo and Freeland (1931). A fuller grammar with texts is in Bright (1957). Additional texts are in Harrington (1932a, 1932b). Discussion of the pronominal system is in Macaulay (1992, 2000).

Core arguments are specified in Karuk by pronominal prefixes on verbs.

(34)  Karuk pronominal prefixes (Bright 1957: 33, 62, 61)

      ni-mniš  ‘I cook’
      Ñu-mniš  ‘he/she cooks’
      kun-imnis  ‘they cook’

First person singular pronominal prefixes distinguish grammatical role.

(35)  Karuk grammatical role (Bright 1957: 62, 61)

      ni-mmah  ‘I see him’
      Ña-mmah  ‘he sees me’

The system is more complex than first meets the eye, however. Only one core argument is identified in any verb. The choice of which argument to specify is
based on a person hierarchy. Second person plurals (used elsewhere in California to show respect) are given priority over 1st persons, and 1st persons are given priority over all others: 2pl > 1 > 2sg > 3. If a 1st person acts on a 3rd (‘I see (him)’), or a 3rd person acts on a 1st (‘(he) sees me’), only a 1st person prefix appears.

There is evidence of an incipient agentive system. The prefix ni- ‘I’ appears with both intransitive and transitive verbs and in both events and states.

(36) Karuk 1st person pronominal (Bright 1957: 63, 48, 125, 125, 62)

ni-ppackxet ‘I won’

ni-čpmēš ‘I will go back’

ni-łaxviš ‘I will take it’

ni-yuwpēn ‘I opened my eyes’

ni-xvi:pha ‘I get angry’

ni-¿ix ‘I was afraid of him’

The prefix na- appears in transitives like ‘he sees me’ above. It also appears in some intransitives, as below.

(37) Karuk 1st person pronominal (Bright 1957: 62, 250)

ná-¡axhi ‘I bleed’

ná-kkûha ‘I am sick’

ná-sáyxrhva ‘I am lonesome’

Bright (1957: 59) notes that certain Karuk verb stems can appear with either 1st person singular prefix ni- or na-, such as ‘be hungry’: ni-xxürihi or na-xxürihi ‘I am hungry’. Other stems with this option include ¡axhi ‘bleed’, kûhi ‘suffer pain, be sick’, sayríchwa ‘be lonesome’, ihitci:xva ‘be jealous’, iñčева ‘be afraid’, iñxaska ‘be thin, lose weight’, fbríhi ‘be bald’, furnká ‘be nervous, cranky, fretful’, hðähi ‘be late, be offended’, iñnaha ‘defecate’, iñxax ‘be cross-eyed’, ikvís:i ‘fall asleep’, ikvís:rih ‘be tired’, imčak ‘burn oneself, get burnt’, imčax ‘be hot’, imči:txko: ‘have one’s bones protrude through one’s skin’, imfráhi ‘feel pain’, ixtih ‘thirst for’, ixtup ‘have an erection’, pakyav ‘have good luck with’, ířih ‘be unwilling, lazy, tired’, and ykvhx ‘have enough’. The prefix ni- thus represents 1st person singular agents, and the prefix na- 1st person singular patients. The appearance of na- with event verbs like ‘fall asleep’ and ‘burn oneself’ shows that the system is agentive rather than active/stative. The agentive system is only incipient, however. Grammatical role is distinguished only in 1st person singular (where speakers have direct knowledge about control and affectedness and the right to specify it), and Bright points out that use of the na- prefix with verbs like ‘be hungry’ is optional.
12.2.2 Chimariko

Interestingly, another language isolate in the same area also shows an agentive system, also with a hierarchical overlay. In Chimariko, another isolate visible on the map in Figure 12.1, core arguments are also represented in verbs by pronominal affixes (Mithun, to appear). Only one argument is represented in any verb. The choice of argument is similarly based on a person hierarchy: 1, 2 > 3; that is, 1st and 2nd persons are given priority over 3rd (‘I found (her)’). If both core arguments are speech act participants (1st or 2nd person), the agent is given priority over the patient (‘I found (you); ‘You found (me)’). Material cited here comes from the fieldnotes of J. P. Harrington, organized and analysed by George Grekoff.

(38) Chimariko pronominal prefixes (Grekoff, n.d.)

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pi-mumni</td>
<td>'I was running'</td>
</tr>
<tr>
<td>mi-mumni</td>
<td>'you were running'</td>
</tr>
<tr>
<td>hi-mumni</td>
<td>'he or she was running'</td>
</tr>
<tr>
<td>Pi-č’ut</td>
<td>'I hit him/her'</td>
</tr>
<tr>
<td>mi-č’ut</td>
<td>'you hit him/her'</td>
</tr>
<tr>
<td>hi-č’ut</td>
<td>'he or she hit him/her'</td>
</tr>
</tbody>
</table>

The 1st person singular and plural pronominal affixes have two forms: one for grammatical agents and the other for grammatical patients. Both appear in intransitive and transitive verbs.

(39) Chimariko pronominals

a. Set I

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pi-mumni</td>
<td>'I was running'</td>
</tr>
<tr>
<td>Pi-siyakutni</td>
<td>'I looked back'</td>
</tr>
<tr>
<td>Pi-k’ot</td>
<td>'I am talking'</td>
</tr>
<tr>
<td>Pi-c’i’ta</td>
<td>'I caught (the crawfish)'</td>
</tr>
<tr>
<td>Pi-xa’łyta</td>
<td>'I made (the water hot)'</td>
</tr>
<tr>
<td>Pi-winquutta</td>
<td>'I dumped (the crawfish)'</td>
</tr>
</tbody>
</table>

b. Set II

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ēchu-saxnit</td>
<td>'I am coughing'</td>
</tr>
<tr>
<td>ēk-awin</td>
<td>'I am afraid'</td>
</tr>
<tr>
<td>ēchu-t’i’niman</td>
<td>'I am glad'</td>
</tr>
<tr>
<td>ēchu-seyin</td>
<td>'I am named, called'</td>
</tr>
<tr>
<td>ēchu-ćxemum</td>
<td>'I jerk, I am shaking, twitching'</td>
</tr>
<tr>
<td>ēchu-akimxanan</td>
<td>'I am going to wash away'</td>
</tr>
<tr>
<td>ēchu-sato’muxanan</td>
<td>'I am going to starve to death, give out'</td>
</tr>
</tbody>
</table>

Set I pronominals typically identify semantic agents in control. Other verbs that appear with Set I prefixes mean ‘jump around,’ ‘eat,’ ‘sing,’ ‘(man) to get married,’
‘act with the foot’, ‘walk draggingly or drag along’, ‘breathe’, ‘hide’, ‘stop or turn around’, ‘swim’, ‘fly’, ‘raise hand’, ‘cross (water)’, and many more. Set II pronominals generally identify a significantly affected participant who is not in control. Other verbs that appear with Set II prefixes mean ‘have a rash or irritation’, ‘be afraid or fear’, ‘hurt or ache’, ‘be dried or parched’, ‘cough’, ‘be glad’, ‘fall’, ‘give out or be short of breath’, ‘sneeze’, ‘sweat’, ‘yawn’, ‘twitch (body part)’, and ‘get well or recover’.

Third persons are not represented in either Karuk or Chimariko when 1st persons are present, and transitivity is not necessarily marked overtly. Under such circumstances, it is easy to see how reanalysis of a nominative/accusative system as an agentive system (or the reverse) could be triggered by contact. We know that there has long been heavy multilingualism in the area, including among the Chimariko and the Karuk. On the model of the Chimariko system it would be easy for Chimariko-Karuk bilinguals to reinterpret an originally transitive Karuk verb like ‘it scares me (object)’ as an intransitive ‘I (patient) am afraid’ and the original object prefix na- as a patient prefix.

12.2.3 Wiyot

There are several other languages in the immediate area that are not related to either Karuk or Chimariko but that show the effects of long-standing intense contact, with multilingualism and intermarriage. Two of them, Wiyot and Yurok, are remotely related to the Algonquian languages in a family called Algic. In all of the Algic languages, core arguments are represented by pronominal affixes on verbs. These languages are known for their hierarchical systems but not for agentive patterns.

The Wiyot pronominal suffixes fall into two sets. The first set are used with verbs such as kámar ‘steal food’, hutw- ‘bring it’, lág ‘go’, dar- ‘act’, haskap ‘bite a piece’, and thahlab- ‘talk to me’. These pronominal endings can be seen in the verbs kámar ‘I steal food’, kámar-at ‘you steal food’, kámar-il ‘he/she steals food’, and kámar-ił ‘one steals food’ (Teeter 1964: 70–71). The same agent suffixes are used with intransitive and transitive verbs: lág-ił ‘he/she goes’, tathánab-il ‘he punched him’, hutw-us-il ‘he/she brings it for you’ (Teeter 1964: 73–4).

The second set of pronominal suffixes are used with verbs such as tíškóhw ‘be liked’, láhg- ‘be heavy’, dóatak ‘be a large roundish object’, tokwan- ‘be slapped’, kohw ‘be tall’, kwápt ‘be covered up’, lalát ‘be burned’, and tuli ‘be wet’. These pronominal endings can be seen in tíšków-al ‘I am liked’, tíšków-alat ‘you are liked’, tíškóhw ‘he/she is liked’ (Teeter 1964: 77). Other verbs cited by Reichard in her 1925 grammar include those with such meanings as ‘be willing’, ‘be bruised’, ‘be cold’, ‘be hungry’, ‘be lazy, dull’, ‘run splinter into body’, ‘lie’, ‘sleep alone’, ‘get power’, ‘be a person of large proportions’, ‘be awkward, clumsy; and be comical’ (Reichard 1925: 76). As can be seen, derived passives occur with this second set of pronominal suffixes.
From the descriptions it is not entirely clear whether the system is basically an agent/patient one, an active/stative one, or a mixture. In most situations, the two patterns would yield the same result: one who is burned is both a semantic patient (not in control) and in a state. Glosses such as Reichard’s ‘run splinter into body’ and ‘get power’ suggest an agent/patient basis. Wiyot does not appear to exhibit a fully formed agentive system, however. Many verbs that would appear with grammatical patients in other languages appear with the first set of pronominals mentioned above rather than the second, such as *pačol-il* ‘he/she/it falls’ and *lol-il* ‘it falls out, drips’ (Teeter and Nichols 1993: ii. 230). Furthermore, the forms of the transitive patient pronominals do not generally match those of the intransitive patients.

### 12.3 The Northwest Coast

Perhaps the best-known linguistic area in North America is the Northwest Coast. It contains languages from a number of language families, many with members outside of the area. Among the families represented are Tlingit-Eyak-Athabaskan, Haida, Tsimshianic, Salishan, Wakashan, and Chimakuan.

#### 12.3.1 Tlingit: Tlingit-Eyak-Athabaskan family

The Tlingit language is a member of the Tlingit-Eyak-Athabaskan family, genetically unrelated to any of the languages discussed so far. It is described in Boas (1917), Swanton (1911a), Story and Naish (1973), and Leer (1991). In all languages of the family, core arguments are identified by pronominal prefixes on verbs. Two 1st person singular pronominal prefixes can be seen below, *xa*- and *xad*. (The different orthographies used in the various sources are preserved in material cited here. Some authors describe the pronominals in the *xad* series as prefixes, others as loosely attached clitics.)

(40) Tlingit pronominal prefixes *xa*- and *xad*- (Leer 1991)

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>ya'n</em>-xa-gid</td>
<td>I am walking along</td>
</tr>
<tr>
<td>xa-layčx</td>
<td>I am making it</td>
</tr>
<tr>
<td>ka-xa-či'xw</td>
<td>I am kneading it</td>
</tr>
<tr>
<td><em>xa</em>-tis'gʷ</td>
<td>I am washing it</td>
</tr>
<tr>
<td><em>xa</em>-ždačx</td>
<td>I kill it</td>
</tr>
<tr>
<td>xad-wusiti'n</td>
<td>she saw me</td>
</tr>
<tr>
<td>xad-yandusqé:ž...</td>
<td>they say to me</td>
</tr>
<tr>
<td>xad-gasgídž</td>
<td>I always fall down</td>
</tr>
<tr>
<td>xad-ganitzq'wž</td>
<td>I get sick again</td>
</tr>
<tr>
<td>xad-kawdusasáy</td>
<td>I am hot</td>
</tr>
<tr>
<td>xad-yač'e:</td>
<td>I am fine</td>
</tr>
</tbody>
</table>
The pronominals follow an agentive pattern. Both appear where English would show subjects. Both occur in intransitive and transitive verbs. Both occur with events and states. The agent \textit{xw-} (\textit{xat-}) can be seen with events above (‘I am walking’) and with states below (‘I am sitting comfortably’).

(41) Tlingit agent pronominals with events (Story and Naish 1973)
\begin{itemize}
  \item \textit{yan sh ka-xw-jix’akw} ‘I’m sitting very comfortably, just the way I want to’
  \item \textit{nisdaat ka-xw-didáťl} ‘I was doubled up last night (it was so cold)’
  \item \textit{ee jik-xw-aawk’ts} ‘I’m alert (wondering what you’re going to do to me)’
  \item \textit{tléilkoo-x-shatéeneen} ‘I was blind (did not have sight)’
  \item \textit{kinax o-o-iz’ik’áa} ‘I’m really lazy’
  \item \textit{āyéet sa-xw-sit’án} ‘I’m expecting my son’
  \item \textit{sa-xw-dháa} ‘I was ready’ (to be killed as a clan representative, to atone for death of an equal on the opposite side)
\end{itemize}

The patient \textit{sat} (\textit{xat-}) appears with states as above (‘I’m tired out’) and events.

(42) Tlingit patient pronominals with events (Story and Naish 1973)
\begin{itemize}
  \item \textit{xat k’iwaawáash} ‘I yawned’
  \item \textit{xat oowák’ich} ‘I sobbed’
  \item \textit{xat yadát’kw} ‘I frequently have hiccups’
  \item \textit{xat woodziġít} ‘I woke up’
  \item \textit{xat kawdik’ee} ‘I failed completely’
  \item \textit{dei xat googanáa} ‘I’m going to die’
  \item \textit{yaa xat nalt’ix’} ‘I am beginning to freeze’
  \item \textit{kei xat naldál} ‘I’m getting heavier’
  \item \textit{yei xat naltéít} ‘I’m gaining weight’
  \item \textit{yei xat nalkún} ‘I’m losing weight’
  \item \textit{kindayígin xat woodziģét} ‘I fell flat on my back’
  \item \textit{yan xat kkey-shakawdligás} ‘I fell on my knees and skidded along’
  \item \textit{yaa xat k’ayanadáax} ‘I kept running’ (even after reaching the finish line my legs continued to keep moving forward fast)
\end{itemize}

Pronominal set is generally lexicalized with each verb. The agent prefix \textit{xw-} (\textit{xat-}) is used in both sentences below, for example, even though the breaking and burning on those occasions was beyond the control or intention of the speaker.

(43) Tlingit non-fluid pronominals (Story and Naish 1973)
\begin{itemize}
  \item \textit{ax taa’al’ee ka-xw-aal’ée} ‘I broke my needle’
  \item \textit{sakwénin xw-ali’túch} ‘I’ve burned the toast/reduced it to charcoal’
\end{itemize}

Tlingit is the only language in its family to show agentive patterning in the pronominal prefixes, however. Of the 37 languages in the family, 36 show subject/object patterning. The pronominals themselves are cognate: the cognates of
the Tlingit agent prefixes represent subjects in all other languages. They appear in the same position in the verb, immediately before a valency prefix (called the 'classifier') and the stem. In all of the languages, additional prefixes precede the agent/subject markers, most of which are also cognate across the family. Examples of the pronominal prefixes cognate with the Tlingit agent $x$-

The examples show, these are subject pronominals, used to represent both semantic agents and patients, in both events and states, and in both intransitives and transitives.

(44) Eyak subject $x$-

ich' $\bar{x}$-tah
ich' qanuh $\bar{x}$-tah
$\bar{x}$-e : 
$\bar{y}$u: yax$\bar{x}$-dah
$\bar{x}$-dæk
$\bar{q}$e $\bar{x}$-daqc:
$\bar{x}$-e:k
$\bar{x}$-t'uh
gala$\bar{x}$-tah
$\bar{x}$-sinh

(45) Navajo subject sh-

yi-sh-háád
nahjá' 'ii-sh-lé
hashi-sh-ch'ág
'li-sh-kaad
'adah 'ii-sh-dáht
'li-sh-zhééh
náa-sh-bál
náa-sh-ch'ág
ni-sh-chon
shi-sh-ch'ín
'adááii-sh-nih

Though the vast majority of the languages in the family show subject and object pronominals, the numbers alone do not prove that the nominative/accusative pattern was characteristic of their common parent, because of the internal relationships within the family, sketched in Figure 12.2. There could have been one shift from an original subject-object system between the breakup of Proto-Tlingit-Eyak-Athabaskan and Tlingit, or one shift from an original agentive system to a subject-object system between the breakup of the common parent and that of Eyak-Athabaskan.

Evidence for one scenario comes from a neighbour. The current territory of the Tlingit can be seen in Figure 12.3. As reported by De Laguna, the Tlingit expanded into this area relatively recently, moving from the south northward and inland
around 1700. The southernmost part of their original territory is now occupied by Haida speakers, though Tlingit town names remain there. De Laguna describes intense contact and intermarriage between the two groups (1990: 213).

The Tlingit are known to have absorbed increments of Haidas and Tsimshians, together with ceremonial prerogatives (like the Henya Eagle clan), just as the Tsimshian and Haida have received some Tlingit groups. The Tlingit have also taken in various Eyak and Athabaskan bands, usually treating these foreign groups as clans, and ascribing them new crests. Such crests identify Tlingit clans, subclans, and lineages and serve as guides in aligning Haida or Tsimshian clans with those of the Tlingit for marriage or potlatching.

The intense contact between Tlingit and Haida suggests an explanation for the difference in core argument categories between Tlingit and all other members of its family.

12.3.2 Haida

Haida is generally considered a language isolate, though possible connections to Tlingit-Eyak-Athabaskan have long been the subject of inquiry (Swanton 1905, 1908, 1911b, Lawrence 1977, Levine 1977, Leer 1991, Enrico 2003). Haida pronouns are free. 1st person singular pronouns hl and dii can be seen below.

\[
\begin{align*}
    & hl \text{ sral-gan} & \text{’I fixed it’} \\
    & \text{Joe } hl \text{ qing-gan} & \text{’I saw Joe’} \\
    & \text{’laa } hl \text{ ist’ida-gan} & \text{’I warned him’} \\
    & hl \text{ nj-angqasaa-ang} & \text{’I am going to go’} \\
    & \text{dii } la \text{ gu’laa-gang} & \text{’he likes me’} \\
    & \text{dii } \text{ginga’aan } la \text{ qeenggaa} & \text{’he looks like me’} \\
    & \text{dii } hl \text{raaaga-ang} & \text{’I am afraid’} \\
    & \text{dii } rahtgal-gang & \text{’I am tired of it’} \\
    & \text{dii } gudang-gang & \text{’I want to’} \\
    & \text{dii } q’ud-ang-gan & \text{’I wasn’t hungry’} \\
    & \text{’laa } \text{ginga’andaqeeinggaa} & \text{’I look like him’}
\end{align*}
\]

The Haida pronouns show an agentive pattern. The form hl is used for semantic agents in both intransitives (’go’) and transitives (’fix’, ’see’, ’warn’), while the form
Agentive systems in core argument marking

Figure 12.3. Northern Northwest Coast of North America, early 19th century (from Suttles 1990: ix)
**Marianne Mithun**

dii is used for those affected in both intransitives ('be afraid', 'be hungry') and transitives ('like', 'look like', 'be tired of something', 'want something').

Like the Tlingit pronominals, these do not show an active/stative pattern. The agent pronoun hl is used both with actions like 'fix', 'see', 'warn', and 'go' above, and with states like 'be out' below. (The transcription systems used in the original sources are retained here.)

(47) **Haida agent hl with state:** Levine (1977: 183)

\[ gway-ay\ gù P\ uà P\ ŋj-\mbox{inn}-i. \]

island-the on focus 1SG.AGT exist-PAST-old

'I (AGENT) was out on the islands.'

The patient pronoun dii is used with stative verbs like 'be afraid', 'be tired', 'be hungry', 'want' and 'resemble' above, as well as with event verbs like 'fall', 'bump into', 'disappear out of sight', 'escape destruction', 'hiccup', 'get mad', and more.

(48) **Haida patient dii with event** (Lawrence 1977: 180)

\[ K’yuw-\mbox{aa-st} díi dlawíi-gan. \]

trail-the-from 1SG.PAT fall-PAST

'I (PATIENT) fell off the walk'

Because pronoun choice is semantically rather than syntactically based, transitive clauses may contain two grammatical patient arguments.

(49) **Haida double-patient transitives** (Lawrence 1977: 77, 147)

a. \[ dàn gú díi kuyáadaang. \]

2SG.PAT 1SG.PAT love 2SG.PAT 1SG.PAT like

'I (PATIENT) love you (PATIENT).'

b. \[ dàn gú díi guláagang. \]

2SG.PAT 1SG.PAT like

'I (PATIENT) like you (PATIENT).'

Argument categories are generally learned with each verb, rather than selected as speakers speak. Sometimes the pronominal category that appears with a particular verb seems at first surprising, until the original or literal meaning of the lexical item is uncovered. The verb 'believe' occurs with a grammatical agent, for example, and the verb 'go somewhere else' occurs with a grammatical patient.

(50) **Haida lexicalization**

\[ yahda\ ‘believe’ \]

AGENT literally ‘make true for oneself’ (Levine 1977: 135)

\[ gadas\ ‘go to live somewhere else’ \]

PATIENT originally ‘come off, ricochet off’ (Enrico 2003: 95)

A few verbs have been lexicalized with both sets of pronouns. The verb gwaawa/gwaawu, for example, means ‘refuse’ with an agent pronoun but ‘not want’ with a patient pronoun. Enrico notes that the verb gudrad is usually used with a patient pronoun to mean ‘remember’, but it can be used with an agent pronoun to indicate effort, as in ‘OK, I’ll try to remember’. A few verbs are used
Agentive systems in core argument marking

by some speakers with agent pronouns and by others with patient pronouns, with little change in meaning, such as 'sneeze', 'vomit', and 'vomit up', verbs that involve performance but not control (Enrico 2003). There is also some systematic dialect variation in pronominal category.

(51) Haida dialect variation Enrico (2003: 93, 102)

<table>
<thead>
<tr>
<th></th>
<th>Masset</th>
<th>Skidegate</th>
</tr>
</thead>
<tbody>
<tr>
<td>q’usab’la</td>
<td>agent</td>
<td>agent of patient</td>
</tr>
<tr>
<td>q’anda</td>
<td>agent of patient</td>
<td>agent</td>
</tr>
<tr>
<td>sk’al.aa’w</td>
<td>patient</td>
<td>agent of patient</td>
</tr>
</tbody>
</table>

As in some other languages, the agent/patient distinction does not run through the entire pronominal paradigm. Enrico (2003) notes that all dialects show it in 1st person pronominals: again, speakers are best able to specify their own control and affectedness. For 2nd person singulars, the Skidegate dialect to the south shows the distinction in all pronominals, but the Masset dialect to the north shows it only in non-clitics. For 3rd person singulars, Skidegate shows the distinction in topical clitics. The distinction does not appear in the remaining pronominals.

12.3.3 A mechanism for the Tlingit shift

Since the only language of the Tlingit-Eyak-Athabaskan family to show agentive patterning (Tlingit) is spoken in an area immediately adjacent to another language with agentive patterning (Haida), and since there is documentation of extensive contact between the two, it appears that Tlingit reanalysed its original nominative/accusative pattern as an agentive one under the influence of Haida. As in Yuki and Pomoan, basic 3rd persons were apparently not specified pronominally in Proto-Tlingit-Eyak-Athabaskan (Krauss 1965). The 3rd person gap in the Tlingit paradigm below is matched in the other languages in the family.

(52) Tlingit pronominal prefixes (Story and Naish 1973: 358)

<table>
<thead>
<tr>
<th></th>
<th>Masset</th>
<th>Skidegate</th>
</tr>
</thead>
<tbody>
<tr>
<td>ya-goot go</td>
<td></td>
<td></td>
</tr>
<tr>
<td>woo-xa-ya-goot</td>
<td>xwaag’oot</td>
<td>'I went'</td>
</tr>
<tr>
<td>woo-ee-ya-goot</td>
<td>yeeg’oot</td>
<td>'you went'</td>
</tr>
<tr>
<td>woo-ya-goot</td>
<td>woog’oot</td>
<td>'he/she went'</td>
</tr>
</tbody>
</table>

(More recently, 3rd person object markers have developed in certain contexts in the various languages, from different sources.) The 3rd person gap alone could provide the conditions for a reanalysis of transitive clauses as intransitives, and objects as grammatical patients ('It makes me sick' > 'I get sick'), providing transitivity was not overtly marked. But unlike Yuki and Pomoan, Proto-Tlingit-Eyak-Athabaskan did contain overt markers of argument structure in every verb, markers which persist in all daughters.
In all languages in the Tlingit-Eyak-Athabaskan family, lexical entries for verbs (called verb themes) consist of a stem, preceded by what is termed a ‘classifier’, possibly preceded by one or more thematic prefixes. The meanings of verb themes may or may not be predictable from their parts. Some samples are below.

(53) Tlingit verb themes (Story and Naish 1973)

a. \(ya-jaak\)w ‘beat up, assault, attack violently’
   \(\text{classifier - stem}\)

b. \(ka-shi-dook\) ‘be slushy’
   \(\text{theme - classifier - stem}\)

c. \(sha-ka-dli-gaan\) ‘lean (against something)’
   \(\text{theme - theme - classifier - stem}\)

\(sha\) ‘head’
\(ka\) ‘surface’

Traces can be seen in some Tlingit verb themes of a classifying function for the classifier prefixes.

(54) Classificatory function in Tlingit (Story and Naish 1973)

\(ya-hoo\) ‘swim on surface of water (human or animal), wade’
\(si-hoo\) ‘swim on surface of water (bird)’
\(ya-xaach\) ‘tow, usually by boat’
\(li-xaach\) ‘tow, usually by boat’ (especially a large object)

The classifier prefixes more often serve a more grammatical function in the modern languages. Verb themes can be grouped the basis of their argument structure: whether they require an agent, a patient, both, or neither.

(55) Tlingit verb theme categories (Story and Naish 1973)

a. Agent/patient transitives \(ya-jaakw\) ‘beat up, assault’
   \(shu-ka-ya-jaa\) ‘advise’
   \(\text{classifier - stem}\)

b. Agent intransitives \(ya-dli-gwaash\) ‘hop’
   \(sha-ka-dli-gaan\) ‘lean (against something)’

   \(\text{classifier - stem}\)

c. Patient intransitives \(ya-naa\) ‘die’
   \(li-teesh\) ‘be lonesome’

   \(\text{classifier - stem}\)

d. Impersonals \(kee-ya.aa\) ‘to dawn’
   \(ka-shi-dook\) ‘be slushy’

Differences in argument structure tend to correlate with the classifier prefix. The first two verbs below are based on the stem \(aat\). The intransitive \(ya-aat\) ‘go in’, with the classifier \(ya\), requires only one core argument, an agent. The transitive \(si-aat\) ‘cause to go out’, with the classifier \(si\), requires two core arguments, an agent and a patient. (In the examples below, the verb stem with classifier is listed first, followed by a full illustrative clause.)
Tlingit argument structure alternations (Story and Naish 1973: 377)

(56)  

<table>
<thead>
<tr>
<th>Verb Form</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ya-.aat neil wootooa.áat</td>
<td>'we went indoors'</td>
</tr>
<tr>
<td>si-.aat neil wootooosei.át</td>
<td>'we made them [the children] go in'</td>
</tr>
<tr>
<td>ya-.ee k’ins’ oowa.ée</td>
<td>'the potatoes are cooked'</td>
</tr>
<tr>
<td>si-.ee k’ins’ awsi.ée</td>
<td>'he cooked the potatoes'</td>
</tr>
<tr>
<td>ya-leet hoon-t oowalit</td>
<td>'it slid into the water'</td>
</tr>
<tr>
<td>li-leet hoon-t awlilít</td>
<td>'he slid it [the tree] into the water'</td>
</tr>
<tr>
<td>di-keil’ haa wdtikéil’</td>
<td>'we ran away'</td>
</tr>
<tr>
<td>li-keil’ haa wlikéil’</td>
<td>'he chased us'</td>
</tr>
</tbody>
</table>

But the system is far from regular. The verb stems below are all based on the root _haa_ with the prefix _ka-_ ‘surface, on’, but the semantics are not transparent.

Tlingit argument structure alternations (Story and Naish 1973: 273)

(57)  

<table>
<thead>
<tr>
<th>Verb Form</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ka-ya-haa</td>
<td>'move (esp indefinite motion), be invisible, move invisibly'</td>
</tr>
<tr>
<td>ka-di-haa</td>
<td>'disappear, especially turn into vapor, cease to exist'</td>
</tr>
<tr>
<td>ka-si-haa</td>
<td>'bury, excavate'</td>
</tr>
<tr>
<td>ka-li-haa</td>
<td>'cause to disappear mysteriously'</td>
</tr>
<tr>
<td>ka-di-haa</td>
<td>'sneak away, cause oneself to disappear by moving gradually out of sight'</td>
</tr>
</tbody>
</table>

The same classifier prefix appears in verbs with different argument structures.

Tlingit same classifier, different argument structures (Story and Naish 1973)

(58)  

<table>
<thead>
<tr>
<th>Verb Form</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ya-.eex’</td>
<td>'growl (bears), cry out, shout'</td>
</tr>
<tr>
<td>ya-.eex’</td>
<td>'call out to, shout to, holler at, invite'</td>
</tr>
<tr>
<td>ya-naa</td>
<td>'die'</td>
</tr>
<tr>
<td>ya-naa</td>
<td>'inherit'</td>
</tr>
<tr>
<td>si-taan</td>
<td>'fall' (of natural precipitation)</td>
</tr>
<tr>
<td>si-taan</td>
<td>'carry, take (usually long, complex object), lay'</td>
</tr>
<tr>
<td>li-gaas</td>
<td>'be forbidden, taboo, not allowed by custom'</td>
</tr>
<tr>
<td>li-gaas</td>
<td>'abstain from, refrain from'</td>
</tr>
<tr>
<td>ka-shi-x’ee’</td>
<td>'be slippery' (oil, ice, wet rocks, etc.)</td>
</tr>
<tr>
<td>ka-shi-x’ee’</td>
<td>'iron clothes'</td>
</tr>
<tr>
<td>di-t’aach</td>
<td>'swim' (human using strokes to slap water)</td>
</tr>
<tr>
<td>di-t’aach</td>
<td>'clap' (hands)</td>
</tr>
<tr>
<td>dzi-gaax</td>
<td>'keep crying'</td>
</tr>
<tr>
<td>dzi-gaax</td>
<td>'cry for, ask for'</td>
</tr>
</tbody>
</table>

Different classifiers appear in verbs with the same argument structure.
The classifiers were clearly part of the common parent language, since reflexes of them appear in all members of the family, but their functions have apparently become less transparent over time, as they have become part of lexicalized verbs that have continued to undergo semantic changes of their own.

Story and Naish (1973) and Leer (1991) point out that the classifiers can be arranged in four series, the ya-, si-, li-, and shi- series. All series appear with both intransitives and transitives, though verbs with ya-series classifiers are more often intransitive, and those with si- and li-classifiers are more often transitive.

The four prefix forms in each series are distinguished by two cross-cutting features. Those identified as +identifiable are used with identifiable times and participants. Those identified as +D show traces of an old element d, which does not always show up as such but which has left phonological traces such as voicing, called the D-effect. The +D classifiers appear to have functioned as detransitivizers. They appear across the family in middles.

But modern +D verbs can vary in their argument structures. Some require agents, some patients, some both, and some neither.
Agentive systems in core argument marking

(62) Tlingit +D verbs with varying argument structures

\[
\begin{array}{ll}
\text{di-haan} & \text{'stand up'} \\
\text{di-x éég} & \text{be burned (of flesh, skin), become shriveled'} \\
\text{di-téé} & \text{'put on (shirt, dress, etc.)'} \\
\text{a-yu-di-téé} & \text{'storm, be stormy, rough (of weather)'} \\
\text{cí'a-dzi-dook} & \text{give final pronouncement of one's opinion and be quiet'} \\
\text{dzi-geet} & \text{fall'} \\
\text{dzi-ëi} & \text{'(medicine man) ask for more (payment) for service'} \\
\text{ka-dli-yaas'} & \text{'stretch out legs; step, place one's foot'} \\
\text{dli-neitl} & \text{be fat'} \\
\text{dli-daas'} & \text{snare'}
\end{array}
\]

The Tlingit-Eyak-Athabaskan languages have no passives, but they do contain generic/indefinite pronominal categories, sometimes called the 4th person, meaning 'one' or 'someone'. This category can be traced back to their common parent: the Tlingit indefinite agent doo- is cognate with the indefinite subjects in the other languages.

(63) Tlingit indefinite agent doo- (Story and Naish 1973)

\[
\begin{array}{ll}
\text{kúux woo-doo-wa-át} & \text{'they (indefinite) went back'} \\
\text{a'aa teen ka-doo-tééy} & \text{'they carve with an adze'} \\
\text{kúux woo-doo-dzi-.eé} & \text{'they (indefinite) (doo-) cooked rice'} \\
\text{x'úx' kaw-too-shi-xít} & \text{'we (too-) wrote a letter'} \\
\text{x'úx' kaw-doo-ji-xít} & \text{'they (indefinite) (doo-) wrote a letter'}
\end{array}
\]

Interestingly, all Tlingit 4th person verbs now require +D (detransitive) classifiers in place of their usual si, li-, or shi- classifiers.

(64) Tlingit indefinite agent with +D classifiers (Story and Naish 1973)

\[
\begin{array}{ll}
\text{kóox woo-too-si-.ée} & \text{'we (too-) cooked rice'} \\
\text{kóox woo-doo-dzi-.ée} & \text{'they (indefinite) (doo-) cooked rice'} \\
\text{x'úx' kaw-too-shi-xít} & \text{'we (too-) wrote a letter'} \\
\text{x'úx' kaw-doo-ji-xít} & \text{'they (indefinite) (doo-) wrote a letter'}
\end{array}
\]

These doo- constructions have apparently been interpreted as functional equivalents of passives, and coincidentally as formal intransitives.

(65) Tlingit passive equivalents (Story and Naish 1973)

\[
\begin{array}{ll}
\text{a. } & \text{sh déin xát kawdoojik'án } \text{sh déi shi-k'aan 'hate'} \\
\text{b. } & \text{yát-g xát woodoodliyéx } \text{yát-g+ li-yeix 'adopt'} \\
\text{c. } & \text{yát-g xát woodoodliyéx } \text{yát-g+ li-yeix 'adopt'} \\
\end{array}
\]

These constructions have apparently been interpreted as functional equivalents of passives, and coincidentally as formal intransitives.
The reinterpretation of the indefinite subject verbs as intransitives without specified agents could set the stage for the reanalysis of objects as patients.

12.4 The Southeast

The other North American languages with agentive systems are not scattered randomly over the continent. Many were located in the Southeast at contact, the area visible in the darkened area in Figure 12.4. In fact all attested languages in the Southeast north of Timucua in south Florida show agentive systems. These include all the languages of the Muskogean family and the isolates Chitimacha, Tunica, Natchez, and Atakapa. Their locations can be seen in Figure 12.5. No speakers remain of any of the isolates.

12.4.1 The Muskogean family

The Muskogean languages are Creek, Mikasuki-Hitchiti, Apalachee, Koasati, Alabama, and Choctaw-Chickasaw. In all of them, core arguments are identified by pronominal affixes in verbs.
Figure 12.5. The Southeast of North America, general situation in the 16th–18th centuries (from Fogelson 2004:x)
The pronominal affixes show an agentive pattern. Both grammatical agents and grammatical patients appear in both intransitive and transitive verbs, but transitivity is not marked overtly. As in the languages seen earlier, the system is based on semantic role rather than aspect or Aktionsart. Agent pronominals appear with both events and states, as do patient pronominals.

As in the languages seen earlier, both sets of pronominals appear in both intransitives, as above, and transitives, as above and below.
Creek transitives with patient pronominals (Martin 2004: 140)

\[ \text{ó:wa-n ca-yá:ci:s} \quad \text{‘I want water'} \]
\[ \text{ó:wa-n ca-hos-t-tó:-s} \quad \text{‘I forget water'} \]

For the most part, the choice of pronominal paradigm is lexicalized with each verb. Some verb stems have been lexicalized with both sets.

Some Creek alternatives (Martin 2004: 139)

a. \( \text{ca-hákti:s-í:s} \quad \text{‘I’m sneezing’ PAT} \) (more usual)
\( \text{hákti:s-éy-s} \quad \text{‘I’m sneezing’ AGT} \)
b. \( \text{hosí:l-éy-s} \quad \text{‘I’m urinating’ AGT} \) (more usual)
\( \text{ca-hósi:l-í:s} \quad \text{‘I’m urinating’ PAT} \) (unable to control it)
c. \( \text{li:tk-éy-s} \quad \text{‘I’m running’ AGT} \) (more usual)
\( \text{ca-li:tk-ís} \quad \text{‘I’m running’ PAT (out of control downhill) (requested)} \)

Third persons are not overtly marked.

Creek zero 3rd persons (Martin 2004: 135, 136)

\( \text{nafk-} \quad \text{‘hit’} \)
\( \text{na:fk-éy-s} \quad \text{‘I am hitting’} \)
\( \text{na:fk-íck-ís} \quad \text{‘you are hitting’} \)
\( \text{na:- -fk-ís} \quad \text{‘(S/he) is hitting’ (no mark in third person)} \)
\( \text{ca-na:fk-ís} \quad \text{‘(s/he) is hitting me’} \)
\( \text{ci-na:fk-ís} \quad \text{‘(s/he) is hitting you’} \)
\( \text{na:fk-ís} \quad \text{‘(s/he) is hitting (him/her)’} \)
\( \text{inokk-} \quad \text{‘be, get sick’} \)
\( \text{ca-nókk-ís} \quad \text{‘I am sick’} \)
\( \text{ci-nókk-í:-s} \quad \text{‘you are sick’} \)
\( \text{(i)nókk-í:-s} \quad \text{‘(he/she) is sick’} \)

All of these features can be reconstructed for Proto-Muskogean (Booker 1980). All contain pronominal affixes that distinguish grammatical agents and patients, all show zero 3rd person forms, and none marks transitivity. Nouns are marked for case, but the case markers follow a nominative/non-nominative pattern.

12.4.2 Atakapa isolate

Atakapa, a language isolate spoken in the same area, also shows an agentive pattern in its pronominal affixes on verbs, though the actual forms of the affixes are different from those in the Muskogean languages. The 1st person singular agent -ó appears in both intransitives (‘travel’, ‘dance’, ‘play’) and transitives (‘wash’, ‘mock’, ‘see’), and in both events (‘enter’) and states (‘be chief’).
(73) Atakapa agent suffix -ô 'I' (Swanton 1929: 125–6)

\begin{itemize}
  \item tsak'c-ô 'I chew'
  \item want-ô 'I travel'
  \item wan puxkint-ô 'I go dancing, will go dancing'
  \item cakhu-ô 'I see them'
  \item wi cakits-ô 'I wake somebody up'
  \item na iwéuckint-ô 'I am mocking you'
  \item òkòrkouc patsem-ô 'I wash a shirt'
  \item wic ankamc-ô 'I play'
  \item wi hattsickic-ô 'I shall be chief'
  \item ñot inik-ô 'I enter the house'
\end{itemize}

The first person patient prefix hi- also appears in both intransitives ('be tired', 'be strong') and transitives ('bother', 'beat', 'pinch'; 'be afraid of', 'hate'), and in both events ('kill', 'fall', 'get pushed') and states ('be sick', 'be lying').

(74) Atakapa patient prefix (h)i – 'I, me' (Swanton 1929)

\begin{itemize}
  \item hi-kuckict haxc 'Don't bother me!' \\
  \item al atna hi-mic 'Give me a little meat!' \\
  \item kò-hi-pamulet 'They seized and beat me' \\
  \item na hi-tsündta 'You'll pinch me' \\
  \item hi-nimahaxc 'Don't kill me!' \\
  \item hi-lacak 'I am tired or lazy' \\
  \item hi-walcat 'I have dreamed' \\
  \item hi-côkec 'I am sick' \\
  \item wihi-laktikit 'I shall be strong' \\
  \item hi-tixt 'I am bent, I am laid or lying' \\
  \item hi-ťańc 'I get pushed' \\
  \item wi hi-makawet 'I fell' \\
  \item icak hi-yamanian 'I was afraid of him' \\
  \item icak hi-imilc 'I hate this man'
\end{itemize}

Swanton points out that there is an indefinite 3rd person pronominal (h)i- that has the same form in all uses as the 1st person singular patient (Swanton's 'objective'): kòi hi-tèù 'one who likes to talk'. Second and 3rd person singular agents are zero. Both facts suggest ways by which the agentive system could have developed out of a nominative/accusative one. As in Yuki, original transitive clauses with zero 3rd person agents could have been interpreted as intransitives, and their original objects as grammatical patients. Alternatively, original transitive clauses with indefinite subjects could have been reinterpreted as intransitives, and the indefinite prefix hi- reinterpreted as a patient pronominal.

12.4.3 Chitimacha isolate

A second isolate originally spoken in the area, Chitimacha, also shows an agentive pattern. The 1st person singular agent is -k and the patient is -ki. Both appear
in intransitives (‘die’, ‘feel cold’) and transitives (‘beat’), with both events (‘make sick’) and states (‘feel cold’). Second and 3rd persons are unmarked on verbs.

(75) Chitimacha agentive patterning (Swadesh 1946: 326)

\[
\begin{align*}
\text{geti-} & \quad \text{‘I beat (him)’} \\
\text{geti} & \quad \text{‘(he) beat (him)’} \\
\text{nikpa-ki} & \quad \text{‘(he) made me sick’} \\
\text{natma-ki} & \quad \text{‘(he) told me’} \\
\text{dadiwa-ki} & \quad \text{‘I feel cold’} \\
\text{nu:p-ki-ču:š} & \quad \text{‘if I die’}
\end{align*}
\]

Other verbs that appear with patient suffixes include kap ašiše- ‘to become weared’, kap agihte- ‘to be greedy’, wokt- ‘to taste, feel’, teki- ‘to suffer pain’, ša- ‘to sleep or fall asleep’, and gæst- ‘to shiver’.

Examples of the suffixes in use can be seen in the text excerpt below, transcribed and translated by Swanton and analysed by Kimball (1992).

(76) Chitimacha (Speaker Benjamin Paul)

\[
\begin{align*}
\text{we ašant’a-nk hup hi čuy-k o:ksni-ču-k} \\
\text{the old.man-by to thither go.SG.AGT steal-FUTURE-SG.AGT}
\end{align*}
\]

‘I’ll go to the old man and I’ll steal it.

\[
\begin{align*}
\text{wetkš we ašant’i hečt šuš c’isn tep hamčmiš hi k’ap tk} \\
\text{then the old.man watch wood piece fire have thither grab}
\end{align*}
\]

Then (he) watched the old man, seized a piece of wood having fire,

\[
\begin{align*}
\text{kun-huk’u nuhé-i} \\
\text{HEARSAY-it.is TUN-AORIST}
\end{align*}
\]

and ran. . .

\[
\begin{align*}
\text{‘učč hiš k’an kâc kw i:’ si’i, kakwi k’ay-i-k} \\
\text{who 3 not know-PARTICIPIAL be know be.NOT-AORIST-SG.AGT}
\end{align*}
\]

Still no one knows. I do not know.

\[
\begin{align*}
\text{huyk’is ’am onak ni šik-ki k’an ši-ki} \\
\text{good something all PREVERB forget-1SG.PAT not be-1SG.PAT}
\end{align*}
\]

I am not forgetting all good things. I have not forgotten everything yet.’

A reanalysis of transitives with zero subjects as intransitives could have occurred.

12.4.4 Tunica isolate

A fourth isolate in the area also shows a three-way distinction in its pronominal affixes. One set of affixes is used for agents of both intransitives and transitives.

(77) Tunica agent prefix ?a- ‘I’ (Haas 1946: 366)

\[
\begin{align*}
\text{Pá-sani} & \quad \text{’I came’} \\
\text{Ruhpehk?á-hkimí} & \quad \text{’I hit him’}
\end{align*}
\]
A second set, also used for inalienable possessors, is used for intransitive patients.

(78) Tunica intransitive patient 'I 't (Haas 1946: 366)

'i-yaši 'I am angry'

'2-šépá 'I am glad'

A third set, also used for alienable possessors, is used for transitive patients.

(79) Tunica transitive patients 'ihk- 'me' (Haas 1946: 366)

'ih-pek'uhki 'he has hit me'

'ihk'-êh'uhki 'he kicked me'

Masculine and impersonal genders are distinguished in the 3rd person. The impersonal 'refers to a non-realist or nameless entity which can never be substantially expressed within the sentence' (Haas 1941: 56–7).

(80) Impersonal (Haas 1941: 57)

šihtuna ya-ti-hê dark become-IMPERSONAL.AGT-SUBORDINATE 'when it got dark'

The impersonal is also used for feminine persons. In addition, as in many languages, impersonal constructions are used much like passives in other languages to background an agent.

(81) Impersonal with passive effect (Haas 1941: 59)

'uwênitištuk'Oh' ani 'uwk-wêni-ti-štuk'ahó-âni 3M.SG.PAT-find-IMPERSONAL.AGT-cannot-be 'one cannot find him' = 'he cannot be found'

Haas notes furthermore that transitive clauses with impersonal agents were used as alternate forms for involuntary actions (1941: 59):

In addition to the types of transimpersonals discussed above there is evidence that stems denoting involuntary action (e.g., 'to breathe'; 'to cough') were formerly used as transimpersonals. The more usual procedure now is to treat such stems as intransitives. Note the following pairs of words given by the informant as equivalents:

<table>
<thead>
<tr>
<th>Transimpersonal</th>
<th>Intransitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>'ih$h'hakatí</td>
<td>or 'h$h'akaní</td>
</tr>
<tr>
<td>'ihk'$h'aktí</td>
<td>or 'ih$h'akani</td>
</tr>
<tr>
<td>'ihk'$ôwikatí</td>
<td>or 'ôwikani</td>
</tr>
<tr>
<td>'ül$h'hukatí</td>
<td>or 'l$h$huká</td>
</tr>
</tbody>
</table>

In addition, there is one stem denoting an involuntary action which is always transimpersonal, viz. hiyu 'to wake up', as in 'ih$t'h'oki 'I woke up'.

The 'transimpersonal' forms contain the impersonal agent suffix -ti and the 1st person transitive patient prefix 'ih(k)-'.

12.4.5 Natchez isolate

The fourth isolate in the Gulf region, Natchez, also shows an agentive pattern in its pronominal affixes. Both the 1st person singular agent prefix *t-* and the 1st person patient prefix *ni-* appear in intransitives and transitives, and in events and states.

(82) Natchez agent prefix *t-* ‘I’ (Haas p.c. 1982)

\[ \text{toy-t-awa'n} \quad \text{‘I win’} \]
\[ \text{kip-t-ewa} \quad \text{‘I’m going around it’} \]


\[ \text{cakna-ni-lan} \quad \text{‘he is sticking me’} \]
\[ \text{macop-e-ni-w} \ \text{‘I will lose’} \]
\[ \text{ca:ya} \quad \text{_{M}Mnu:’a} \]
\[ \text{ca:ya} \quad \text{_{M}M-ni-w-a.} \]
\[ \text{deer-Article be.tired.of-PAST-1SG.PAT-AUX-MOD} \]
\[ \text{‘I am tired of deer meat.’} \]

Additional verbs with patient prefixes have such meanings as ‘be ashamed’, ‘be afraid’, ‘be happy’, ‘be excited’, ‘be intoxicated’, ‘be cold’, ‘be thirsty’, ‘be tired’, ‘be old’. Verbs with grammatical patients corresponding to English subjects are conjugated with auxiliaries inflected for tense in the zero form used for 3rd person agents.

12.4.6 The Southeast

The languages indigenous to the Southeast are not demonstrably related to each other. The Muskogean family forms an uncontroversial genetic unit, but the status of the isolates Atakapa, Chitimacha, Tunica, and Natchez is much less clear. The four isolates are sometimes referred to as ‘the Gulf languages’, but primarily on the basis of their location along the Gulf of Louisiana. It has been hypothesized that Natchez might be related to the Muskogean family as a whole at a time depth comparable to that of Proto-Indo-European (Brinton 1873, Swanton 1907, Haas 1956). Haas commented (p.c. 1989) that if there are any relationships among the languages in the area, the closest to the Muskogean family would be Natchez, followed by Tunica, then Atakapa, and finally, though questionably, Chitimacha, but she made it clear that any such relationships were too remote to be established with certainty.

Though the shapes of the pronominal affixes vary from language to language, it is striking that all show agentive patterns. All of the languages also show just the features that could foster the realanalysis of nominative/accusative systems as agentive systems and vice versa. All contain zero 3rd person agent affixes, indefinite pronominal agent affixes, or both. None shows a formal distinction between intransitive and transitive verbs. Under these conditions it would be easy to reanalyse original transitive clauses with omitted or impersonal 3rd person agents as intransitives, and original object pronouns as grammatical
patients: ‘(X) scares me (object)’ or ‘Something scares me (object)’ > ‘I (patient) am scared’.

The Southeast is well known as a strong linguistic area, one that shows the effects of long-standing, intense contact across language boundaries and pervasive multilingualism. This situation, along with the fact that all of the languages contain the structures that would facilitate the reanalysis of accusative patterns as agentive patterns, suggests that such reanalysis was triggered by contact. The languages do vary in the proportions of verbs that appear with patient pronominals, suggesting that the pattern is less deeply established in some of the isolates, but it is not clear where the agentive pattern may have originated. The pattern may even have originated outside this group.

Several other large language families are also represented in the Southeast, in particular the Iroquoian family (Cherokee, Tuscarora, Nottoway, and Meherrin), the Caddoan family (Caddo), and the Siouan–Catawba family (Ofo, Biloxi, and Tutelo). As noted at the outset, all three of these families show well-established agentive patterning in their pronominal affixes (Mithun 1991). All three families covered a wide area at contact, extending throughout the Eastern Woodlands, west over Great Plains, and north of the modern Canadian border. The isolate, Euchee (Yuchi), generally thought to be related to Siouan–Catawba, was spoken in Georgia until 1836–40, when speakers were forced to move to Oklahoma. It shows an active-stative pattern (Linn 2001). It is not unlikely that this system developed from an earlier agentive system, since the vast majority of actions involve agents, and the vast majority of states involve patients.

12.5 Conclusions

Parallelism in core argument categories may not be as good a diagnostic of deep genetic relationship as once thought. We are learning more all the time about the processes by which systems of one type can develop into those of another. Here we have seen that the distribution of agentive patterns shows an extremely strong areal distribution in North America, occurring in three of the strongest linguistic areas, California, the Northwest Coast, and the Southeast. It is easy to see the mechanisms by which agentive patterning could be transferred through contact, even when the markers that carry it, usually pronominals, are not transferred themselves. In situations where pronominal reference to topical 3rd persons is usually omitted, and there is no overt marker of transitivity, it would be easy to reinterpret transitive clauses like ‘(It) is hurting me’ as intransitives like ‘I am in pain’ and object markers as grammatical patient markers. Nearly all of the languages seen here show just these features.

A second mechanism could also result in such a reanalysis. In languages with an indefinite or generic pronominal category (‘one’, ‘someone’, ‘something’), transitive clauses with indefinite subjects are often used in ways similar to passives in other languages, as a means of backgrounding agents. When such usage becomes
well established and the indefinite subjects are fully backgrounded, the clauses may be reinterpreted as intransitives, as in Tlingit and Atakapa. Both of these processes could occur spontaneously in languages under the right conditions, but the areal distribution of the agentive systems in North America strongly suggests that bilingualism in another language which already had such a system could provide a strong stimulus for the reanalysis.

The more we learn about the genetic and areal distributions of core argument patterns cross-linguistically, and about the precise patterns to be found in individual languages, the better we will be able to uncover the internally and externally triggered mechanisms by which they can develop and change. Work by Harris (1985) on Kartvelian, Aldai (this volume) on Basque, Gutiérrez and Zavala (2005) on Chol, Holton (this volume) on North Halmaheran languages, and Malchukov (this volume) on a variety of languages in North America and New Guinea are already contributing to our understanding of such mechanisms.
Argument dereferentialization in Lakota

REGINA PUSTET AND DAVID S. ROOD

13.1 Semantic alignment and pragmatic perspective

Typological research has revealed that many languages are equipped with some means of suppressing core arguments of transitive clauses (Keenan 1985). The structural devices used for this purpose include impersonalization strategies, such as the agent-suppressing they impersonal in English examples like they told me, where there is no specific referent for the agent (they) in the universe of discourse. The sole function of impersonals is that of backgrounding the transitive agent with respect to discourse saliency. Lakota has a rich repertory of such backgrounding constructions, which are used with both transitive and intransitive base verbs. Some of these have not yet been described in great detail.

However, argument suppression does not involve only backgrounding: many languages have structural devices for moving arguments out of a less discourse-salient into a more discourse-salient position, a process which is referred to as foregrounding. The most discourse-salient position in a clause is often associated with notions such as subject, topic, or focus, depending on the language studied and the choice of terminology. Argument suppression may also occur as a consequence of shifting a non-foregrounded argument into foregrounded position, since the original foregrounded argument loses this position.

Languages with semantic alignment such as Lakota (Siouan language family, Central North America) have been said to lack morphosyntactic devices which serve the purpose of foregrounding (Foley and Van Valin 1984: 155–9, Hopper and Thompson 1980: 280, Keenan 1985: 243–45). Put differently, foregrounding does not exist in such languages, at least not to the extent that the expression of grounding relations is mandatory in each clause (e.g. Foley and Van Valin 1984, Van Valin 2001, 1980). Constructions which fulfil this function are,

We are indebted to the Lakota native speakers Della Bad Wound, †Neva Standing Bear, and Dorothy Rose Wilson for providing the language data that made this study possible.
Argument dereferentialization in Lakota

in particular, passives and antipassives. These constructions reverse the status of arguments regarding grounding as compared to the situation holding in the transitive base clause. If cases serve as markers for foregrounding in a nominative-accusative or ergative language, the cases which fulfill this function are the nominative and the absolutive respectively (e.g. Van Valin 1980: 322). These cases are pragmatically meaningful in that they convey the notion of foregrounded status. Argument marking in semantic alignment languages has been interpreted as being geared towards the expression of semantic roles only; pragmatic roles are thought to be irrelevant in such systems (e.g. Foley and Van Valin 1984). The very fact that semantic alignment languages do not have unified coding for the argument of intransitive clauses is one of the strongest arguments in favour of the assumption that case marking in these languages is exclusively rolesensitive, at least in those semantic alignment languages in which the varying coding formats for intransitive subjects can be semantically aligned with case role marking in transitive clauses. Lakota is an example of a language which shows such alignment in terms of a semantic parameter that is best, though not entirely uncontroversially, described as agency (Mithun 1991: 514–18, Pustet 2002). Alignment splits can, however, also occur along the lines of semantic and/or grammatical categories such as tense, aspect, and subordination vs. non-subordination.

In languages in which the notion of foregrounding is not part of the conceptual ‘load’ conveyed by basic transitive clauses, the alternation between the two expression formats for a single event exemplified by the English active vs. passive examples (1) and (2) cannot be expected to exist because there is no need to shift pragmatic perspective (similarly, cf. Van Valin 1980: 324–5).

(1) The raccoon chased the dog. 

(2) The dog was chased by the raccoon.

In the extant documentations of Lakota, either no mention is made of a passive (Boas and Deloria 1941, Buechel 1939), or else Lakota is explicitly said to lack a passive or similar construction by means of which examples like (2) can be translated directly from English (Van Valin 1985: 368). However, recent fieldwork has revealed that Lakota is in fact equipped with a passive-like construction which is functionally equivalent to the English passive, although the construction has some characteristics which set it apart from the ‘classical’ agent-demoting and simultaneously patient-promoting passives discussed in the typological literature (e.g. Keenan 1985, Shibatani 1985, 1988b, Siewierska 1984). This construction might be a new development which has been triggered by prolonged contact with English. At least, it has not been documented in the existing descriptions of Lakota.

---

1 Other grammatical devices which express foregrounding include inverse constructions in obviations systems and Philippines-style focus systems.
13.2 The data

Core case relations are not morphologically marked on Lakota NPs. Obliques are coded by case-marking suffixes or postpositions. Lakota exhibits a complex system of person-marking affixes on the predicate: intransitive ‘subject’, transitive agent and patient, and, at least in specific semantic and syntactic constellations, benefactives, possessors, and other oblique roles are cross-referenced by person markers on the predicate. With few exceptions, such as the monomorphemic 3rd person plural patient affix *wicha-*‘s’, these markers are composed of two parts: a number-insensitive person marker and a number marker. To a large extent, the person-marking paradigms for transitive agent and patient are morphologically distinct. Singular number is not marked overtly, while plurality is expressed by means of the suffix *-pi.* This number-marking system is used with both transitive agents and transitive patients. Intransitive arguments are coded by means of these sets of markers as well (see Table 13.1). The clause in (3) illustrates the usage of Lakota person/number markers in transitive predicates; in this case, the person markers appear as infixes to the root *nah.* ‘to hear’:

\[(3) \text{ Na-má-ya-hú’u-pi.} \]
\[
\text{STEM-1SG.PAT-2AGT-hear-PL.AGT} \]
\[
\text{‘You guys hear me.’} \]

13.2.1 Agent suppression

Lakota has an impersonalizing construction which serves to suppress reference to the agent of transitive events. This construction is homonymous with the coding format for 3rd person plural agent, which involves a zero agent marker and the plural suffix *-pi*:

\[(4) \text{ sú’ka } na-má-O-hú’u-pi.} \]
\[
\text{dog DEF STEM-1SG.PAT-3AGT-hear-PL} \]
\[
\text{‘The dogs hear me.’} \]

The impersonalizing *-pi*-construction is glossed by means of the acronym AGIPS (‘agent impersonalizer’):

\[(5) \text{ Hená ū wástįkala kága-pi.} \]
\[
\text{those with dried corn make-AGIPS} \]
\[
\text{‘With these (things) dried corn was made.’} \]

\[(6) \text{ Mary Light e-má-ciya-pi.} \]
\[
\text{Mary Light STEM-1SG.PAT-say to-AGIPS} \]
\[
\text{‘I am called/my name is Mary Light.’} \]

2 Number marking is implicitly present in certain portmanteau morphemes such as *ma-*‘1st person singular patient’ (cf. (3)).
Argument dereferentialization in Lakota 337

(7) ‘Tûku e-ní-ciya-pi sò?’ eyá yûkhá
what STEM-2PAT-say to-AGIPS QS say then
‘Užžeštka e-má-ciya-pi k’yu’, eyá kéye.
rose STEM-1SG.PAT-say to-AGIPS ASS say QUOT
‘“What is your name?” he said. It (the bush) replied: “My name is rose.” ’

This construction is analogous to the impersonalizing they construction in English, as in (8),

(8) They sent him home.

which is semantically equivalent to the agentless passive in he was sent home.

The following examples clearly indicate that the element -pi, when used in an agent-suppressing function, does not convey the notion of plurality. The first group of examples features the verb thû ‘to give birth to’, which, ontologically speaking, admits only one agent per instantiation in reality. The agent in question is the mother giving birth. Nevertheless, in examples (9)–(11), the verb thû ‘to give birth to’ carries the element -pi, which, in these cases, must be regarded as an agent impersonalizer:

(9) Wichás.a wâ chčá wâ kíci-thû-pi.
man IDEF.SG child IDEF.SG 3SG.POSS-give birth to-AGIPS
‘A man’s son was born.’ (Deloria 1932: 106)

(10) Lakhóta ki lé pte-sâ wâ thû-pi cha
Indian DEF this buffalo-white IDEF.SG give birth to-AGIPS QL
w-Ø-lýntsí-pi.
Nsp.PAT-3AGT-happy-pl.AGT
‘The Indians celebrate the birth of a white buffalo/that a white buffalo was born.’

(11) Waníkiya Thû-pi Ápétu
Saviour (Jesus) give birth to-AGIPS day
‘The Saviour’s Birthday’, i.e. Christmas.

Examples (12) to (14) also validate the hypothesis that the impersonalizing -pi lacks number marking function. In these cases, the discourse context implies a singular referent for the semantic agent of the -pi-construction.

Context: the injured person had been hit by a rock thrown by a dwarf-like being.

(12) Ihúhíni yûkhá hêchel hû el aphpá-pi hé hêchel yazá kéye.
in the morning then so leg at hit-AGIPS that so hurt QUOT
‘In the morning the spot where he had been hit hurt.’
<table>
<thead>
<tr>
<th>Person</th>
<th>SG</th>
<th>DU</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ma-Xa</td>
<td>μ(k)-X</td>
<td>μ(k)-X-pi</td>
</tr>
<tr>
<td>2</td>
<td>ni-X</td>
<td>–</td>
<td>ni-X-pi</td>
</tr>
<tr>
<td>3</td>
<td>Ø-X</td>
<td>–</td>
<td>Ø-X-pi</td>
</tr>
</tbody>
</table>

(b) Stative intransitive person affixes

<table>
<thead>
<tr>
<th>Person</th>
<th>SG</th>
<th>DU</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ma-X</td>
<td>μ(k)-X</td>
<td>μ(k)-X-pi</td>
</tr>
<tr>
<td>2</td>
<td>ni-X</td>
<td>–</td>
<td>ni-X-pi</td>
</tr>
<tr>
<td>3</td>
<td>Ø-X</td>
<td>–</td>
<td>Ø-X-pi</td>
</tr>
</tbody>
</table>

(c) Transitive person affixes

<table>
<thead>
<tr>
<th>1SG.PAT</th>
<th>2SG.PAT</th>
<th>3SG.PAT</th>
<th>DU.PAT</th>
<th>1PL.PAT</th>
<th>2PL.PAT</th>
<th>3PL.PAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>ma-ya-X</td>
<td>[coded by separate reflexive paradigm]</td>
<td>ma-X</td>
<td>–</td>
<td>–</td>
<td>ma-ya-X-pi</td>
<td>ma-X-pi</td>
</tr>
<tr>
<td>ni-X</td>
<td>[coded by separate reflexive paradigm]</td>
<td>μ-ni-X</td>
<td>μ-ni-X-pi</td>
<td>–</td>
<td>ni-X-pi</td>
<td></td>
</tr>
<tr>
<td>Case</td>
<td>3sg.pat</td>
<td>1pl.pat</td>
<td>2pl.pat</td>
<td>3pl.pat</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3SG.PAT</td>
<td>wa-X</td>
<td>ya-X</td>
<td>[coded by separate reflexive paradigm]</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1PL.PAT</td>
<td>–</td>
<td>μ-ya-X-π</td>
<td>μ(k)-X-π</td>
<td>[coded by separate reflexive paradigm]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2PL.PAT</td>
<td>chi-X-π</td>
<td>–</td>
<td>ni-X-π</td>
<td>[coded by separate reflexive paradigm]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3PL.PAT</td>
<td>wich-a-wa-X</td>
<td>wich-a-ya-X</td>
<td>wich-a-X</td>
<td>[coded by separate reflexive paradigm]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a X symbolizes the position of the verb stem when the person affixes function as prefixes (rather than as infixes, which is also possible).
Context: the boy’s mother is the only person pulling out prairie turnips here.

(13) Yuŋká hokšíla ki léchel intišaŋí lé thípsila yuží-pí then boy DEF here nevertheless this prairie turnip pull-up-AGIPS cha ohlóka cha ektá-šna čýokas’-hq 3khé. QL hole QL at-HAB look at-PROG QUOT

‘The boy, nevertheless, kept peeping through the holes left by the pulled-out prairie turnips.’

Context: the person asking for a story was a single person, namely Regina Pustet.

(14) Héchel apétu ki lé él wó’oyaka waží oyág-ma-ši-pí. so day DEF this on story IDE SG tell-INSG.PAT-ASK-AGIPS

‘Today I was asked to tell a story.’

Another group of examples which prove that -pi has the potential of suppressing agents by dereferentializing them is characterized by the fact that in the surrounding discourse context, an argument which qualifies as a filler for the semantic agent slot in the pi-construction cannot be identified. Where -pi-constructions occur at the beginning of narratives, particularly good examples for such configurations appear. Examples (15)–(17) each constitute the opening sentence in the narratives they are taken from.

(15) Tókhí lé waníyetu tópa sece cha héháni lé about this year four maybe QL at that time this kiyékhíyapi o’inaží hécha cha hé líla tháka kága-pí. airport such QL that INTENS big make-AGIPS

‘About four years ago an airport was built, a very big one.’

(16) Ho hehál tókhéškhe pápa kága-pí ki hé oblákjí-kte. well then how dried meat make-AGIPS DEF that tell.INSG.AGT-FUT

‘Next I will tell about how jerky was made.’

(17) Lehál wagmíza skuyá ki hé ü tókhéškhe íh’á-pí ki now corn sweet DEF that with how boil.vt-AGIPS LK hé oblákjí-kte. that tell.INSG.AGT-FUT

‘Now I will tell about how people cooked with sweet corn.’

Further, -pi must be interpreted as dereferentializing in contexts in which the implied agent is semantically generalized to the effect that no particular extralinguistic entities are referred to:
Argument dereferentialization in Lakota

(18) Tukté’el makhóche wā₂ el wigli yaké ḥáṭṭahq hé táku kihq wayyĩka-pi ṣkhé.

‘Wherever there is mineral oil in the ground something like that can be observed.’

An additional question that arises in the attempt to determine the exact functional scope of impersonalizing-pi concerns the syntactic argument structure imposed by this construction. As stated above, impersonalizing -pi is homonymous with the marking format for 3rd person plural agent, in which the 3rd person is expressed by a zero affix. Would it be appropriate to analyse an impersonalizing -pi-construction as more ‘abstract’ in function than the homonymous plural agent form of a verb, i.e. as a separate valence-reducing construction that eliminates the transitive agent phrase? This would imply that a predicate containing impersonalizing -pi is intransitive: only the patient would then be present at the structural level. Unfortunately, an unequivocal answer to this question cannot be provided at this point because, due to the fact that in Lakota transitive agents are always zero-marked in the 3rd person, the clause given in examples (19) and (20) can be analysed in two ways, i.e. either as lacking an agent or as containing an agent:

(19) Thaló ki hé Ô-kablá-pi.

meat DEF that 3SG.PAT-slice-AGIPS

‘The meat was sliced.’

(20) Thaló ki hé Ô-Ô-kablá-pi.

meat DEF that 3AGT-3SG.PAT-slice-AGIPS

‘People sliced the meat.’

In some languages, passives (and, presumably, antipassives as well) can be used with intransitive base verbs. For instance, some German intransitives, such as lachen ‘to laugh’, can be passivized. The result is complete suppression of arguments at the semantic level by default, since in a structurally intransitive verb there is only one argument that lends itself to suppression. In syntactic terms, this construction is still intransitive since the subject slot is filled with the dummy argument es ‘it’:

(21) Es wird ge-lach-t.

it 3SG.PRES.PASS.AUX PPP-laugh-PST

‘There is laughing going on.’

The Lakota -pi-impersonal is used with intransitive base verbs as well; however, dummy insertion as in German or other modifications of the valence frame of the respective verbs do not take place.
Éyaš lehàn héchu s’e wachi-pi.
but now that way dance-agips
‘But today people dance like this.’

Hená è cha ū chethi-pi.
those IP QL with build a fire-agips
‘With these things people started a fire.’

Lé ū, wahítke na itázipa ki lená ū, wayé yá-pi.
this with, arrow and bow DEF these with, hunt go-agips
‘With this, with these arrows and bows, people went hunting.’

Lená léchu-pi chá-s.na hená ū akisn-pi.
these do this-agips then-hab those because of recover-agips
‘Whenever people do that they get well from it.’

13.2.2 Patient suppression

In Lakota, transitive patients can be ‘blotted out’ in ways analogous to transitive agent suppression by means of the -pi-construction. Several predicate affixes fulfil this function: wa- ‘inanimate patient dereferentializer’, taku- ‘inanimate patient dereferentializer’, and wicha- ‘animate patient dereferentializer’.

13.2.2.1 wa- ‘inanimate patient dereferentializer’  Using the most neutral formulation possible, the basic function of wa- consists in blocking the patient slot in transitive verbs for transitive patient markers. Such constructions can be translated into English by simply omitting the patient if the English verb in question allows this, as in the case of the transitive base verb yúťA ‘to eat (something)’:

(26) Wa-yáta-pi-kte.
PATIPS-cal.2AGT-PL-FUT
‘You guys will eat.’

As with impersonal -pi, the wa-construction might be interpreted in two ways: wu- either eliminates an argument slot, this time the transitive patient slot, or fills it. In the former case, detransitivization takes place, and the construction could be analysed as an antipassive; in the latter case, a translation like ‘non-specific patient’, ‘things’, ‘stuff’ would be appropriate. In the absence of additional structural clues supporting one of these interpretations, it is hard to decide which one is more adequate.

Unlike -pi, wa- may appear more than once in a given predicate:

(27) Wa-w-ó-O-kiya-pi.
PATIPS-PATIPS-STEM-3AGT-help-PL
‘They helped (various people with various things).’
Argument dereferentialization in Lakota

Usually, *wa-* serves to suppress reference to inanimate entities. However, there are occasional examples in which the argument in question must be interpreted as animate. Thus, in (27), there are two patient slots: one indicates the beneficiary of the act of helping, the other indicates the argument expressed by means of the preposition *with* in the English translation ‘to help with.’ From a semantic point of view, the beneficiary in (27) must be animate.

Additional examples of *wa-* referring to animate arguments include:

(28)  \[\text{Wa-Ø-kté-pi.}\]
\[\text{PATIPS-3AGT-kill-PL}\]

‘They killed.’

The patient of an act of killing is necessarily animate.

(29)  \[\text{John Jack wa-’iwa-Ø-Ø-ye.}\]
\[\text{John Jack PATIPS-remind-3SG.AGT-3SG.PAT-STEM}\]

‘John reminds people of Jack.’

The native speaker who provided this example remarked that it implies people in general as cognitive recipients of the act of reminding.

*Wa-* may occur with intransitive base verbs, but such combinations are not very productive; presumably, they are lexicalized remnants of a historical stage in which *wa-* was used more widely with intransitives. Only stative intransitives (lexemes denoting property concepts and material entities) have been found in combination with *wa-*. Examples include:

(30)  \[\text{wa-thó}\]
\[\text{‘grass, green grass, leaves, garden’}\]
\[\text{wa-green/blue}\]

(31)  \[\text{wa-xə}\]
\[\text{‘faded things’}\]
\[\text{wa-faded}\]

(32)  \[\text{wa-cik’ala}\]
\[\text{‘small things’}\]
\[\text{wa-small}\]

(33)  \[\text{wa-ká}\]
\[\text{‘elders’}\]
\[\text{wa-old}\]

(34)  \[\text{wa-há}\]
\[\text{‘hides (pl noun)’}\]
\[\text{wa-hide}\]

(35)  \[\text{wa-çi̋k’ala}\]
\[\text{‘small things’}\]
\[\text{wa-faded}\]

(36)  \[\text{wa-sí̋j’fat (n)\text{ (noun)}\text{ )}\]
\[\text{wa-fat (noun)}\]

(37)  \[\text{wa-mní-tu}\]
\[\text{‘whale, shark, octopus, sea lion, seal, any large ocean animal’}\]
\[\text{wa-water-LOC}\]

(38)  \[\text{wa-hú-topa}\]
\[\text{‘quadruped’}\]
\[\text{wa-leg-four}\]

13.2.2.2 *takú-* ‘inanimate patient dereferentializer’  Like *wa-*, *takú-* suppresses specification of transitive patients. The argument in question is, without exception, inanimate. It seems safe to assume that *takú-* ultimately derives from the indefinite-interrogative pronoun *taku* ‘things, something, what’. Elicitation

\[\text{3. Elicitation}\]

* - *taki* is probably a shortened form of *takúku* ‘(all kinds of) things’. This element in turn derives from a reduplicated form of *taku* ‘things, something’. The stress shift involved is idiosyncratic; at least it cannot be explained in terms of rules of Lakota phonology.
Regina Pustet & David S. Rood

shows that takú- can be substituted for patient-suppressing wa- in almost any case, although takú- is by far less frequent than wa- in discourse.

(39) Takú-blůtšá. PATIPS-1SG.AGT.finish 'I finished things.'

(40) Takú-Ø-blůtšá-pi. PATIPS-3AGT-INSTR-broken-PL 'They broke (many) things.'

(41) Takú-wa-kayęğa. PATIPS-1SG.AGT-SW 'I sew.'

(42) Takú-Ø-nah'ũ-pi. PATIPS-3AGT-hear-PL.AGT 'They heard things.'

(43) Lël takú-cho-wa-k'į̄. here PATIPS-STEM-1SG.AGT-roast 'I'm roasting things here.'

13.2.2.3 wichá- ‘animate patient dereferentializer’ Evidence from Lakota discourse suggests that the regular 3rd person plural animate patient marker wichá- can be used non-referentially as well, and thus may serve to suppress reference to animate patients in transitive clauses, just as wa- suppresses reference to inanimate patients. For instance, wichá- may appear at the beginning of stories at a point at which no argument that qualifies as coreferential with the semantic patient of a wichá-construction can be identified, as in (44):

(44) YuΔhá iháni Iktómi kākhena yā-hał kęye. YuΔhá then long ago Iktomi about go-prog quot then γúndhelakha tuktěl lowá-pi na wachi-pi cha suddenly somewhere sing-agips and dance-agips so na-wicha-h'ũ̄ kęye cha. STEM-PATIPS-hear quot qL 'A long time ago Iktomi was travelling around. Suddenly he heard people singing and dancing somewhere.'

(Note that lowá-pi and wachi-pi, glossed as ‘singing’ and ‘dancing’, are additional examples of the intransitive use of impersonal -pi.)

Similarly, in (45), wichá- in wichá-khuwa-pi ‘they chased people’ does not refer to specific people, but rather, to people in general. The same applies to wichá- in e-wicha-kiya-pi ‘they called them’.

(45) Eháni hél wanāgũ̄ eyá wiwila wichāša e-wicha-kiya-pi long ago there spirit liK STEM-PATIPS-call-agips cha héchacha hená eyáša Ŭ-wicha-khuwa-pi. QL that kind those always 3AGT-PATIPS-chase-PL.AGT 'Long ago spirits called “spring men” sometimes chased people.'

In the excerpt from a recipe for a toothache remedy given in (46), wichá- in ñ-wicha-kiye ‘it helps’ denotes people in general. The preceding context does
not contain any referentially specific 3rd person plural arguments which can be interpreted as coreferential with wicha- in this case.

(46) Hé ugnás hi yazá-pi hátqha hé
    that maybe tooth ache-agips when that
    i-y-ó-gnaka-pi-kte, pus-yá-kel, hátqha hé
    mouth-El-LOC-place-agips-fut dry-ADV-kind of when that
    ó-wicha-kiye.
    STEM-PATIPS-help.vt
    ‘When people maybe have a toothache, they put it in their mouth, dried as it is. It helps (people).’

13.2.3 More details on the usage of argument-suppressing -pi and wa-

By and large, argument-suppressing -pi, wa-, takú- and wichá- seem to target core arguments only, i.e. transitive agents and patients, respectively, and intransitive ‘subjects’. There are, however, occasional exceptions to this rule.

Agent-suppressing -pi may cooccur with patient-suppressing wa- or wichá- in a single verb form:

(47) Wa-yátka-pi ki lila s’ag-yáha jykče.
    PATIPS-drink-agips def intens strong-ADV run
    ‘Drinking is going on very strongly.’

(48) Cha wa-glúzaz.a-pi na’js nówá-pi ki,
    QL PATIPS-poss.wash-agips and bath-agips def
    hená nakú lila wówasukiye óta.
    those also intens rule many
    ‘For doing laundry and taking baths there also were lots of rules.’

(49) Apétu iyohila owáchekiye ki lená wa-ká-hila-pi na
    day each church def these nsp.pat-instr-ring-agips and
    hená wichá-ha-pi.
    those PATIPS-bury-agips
    ‘Every day the church bells rang, and people were buried/there were funerals.’

In many cases, such combinations result in a concept that is translated by an English noun:

(50) wa-khályá-pi (pronounced: wa-khála-pi)
    PATIPS-heat.v-agips
    ‘coffee’

(51) wa-kága-pi
    PATIPS-make-agips
    ‘statue’

(52) wa-l’áyethu-pi
    PATIPS-decorate-agips
    ‘decorations’
Table 13.2. Summary of argument-suppressing constructions in Lakota

<table>
<thead>
<tr>
<th>Coding format for AG</th>
<th>Coding format for PAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>-pi- ‘agent dereferentializer’</td>
<td>–</td>
</tr>
<tr>
<td>wa- ‘inanimate patient dereferentializer’</td>
<td>Same as in active transitive clause</td>
</tr>
<tr>
<td>takú- ‘inanimate patient dereferentializer’</td>
<td>Same as in active transitive clause</td>
</tr>
<tr>
<td>wichá- ‘animate patient dereferentializer’</td>
<td>Same as in active transitive clause</td>
</tr>
</tbody>
</table>

13.2.4 Interim summary: lack of promotion of constituents in argument-suppressing constructions

None of the numerous argument-suppressing constructions dealt with so far has a characteristic that is often felt to be essential for the functional definition of passive or antipassive: the syntactic promotion of an argument. In the English passive, for instance, the basic transitive patient (the boy in (53) and (54)) is promoted to the status of subject, a position occupied by the agent in the corresponding active clause:

Active:

(53) The teacher sent the boy home.

Passive:

(54) The boy was sent home (by the teacher).

Pronominal arguments provide a clearer example for promotion of patients in passive clauses, since in this case the accusative case marking of the patient of the active clause changes to nominative in the passive:

(55) She sent him home.  (56) He was sent home by her.

Although, just like standard passives and antipassives, the above Lakota constructions serve to either remove or neutralize—depending on the interpretation—one of the basic transitive arguments, the syntactic status of the remaining argument remains unchanged. Table 13.2 summarizes the properties of the above Lakota constructions in this regard. The respective case-marking formats for the residual argument in argument-suppressing constructions—the agent in patient-suppressing constructions and the patient in agent-suppressing constructions—are in any case identical to the case marking that the latter arguments exhibit in the corresponding active transitive base clauses. Analogous structures are documented for passives in languages like Mojave (Langacker and Munro 1975: 810) and Welsh (Comrie 1977: 55).
Argument dereferentialization in Lakota

On these grounds, it seems safe to conclude that, as stated in works like Foley and Van Valin (1984) or Van Valin (2001), Lakota is a role-dominated language, in which case marking is exclusively dictated by semantic content rather than by the need to shift case markers to serve the function of coding pragmatic roles, such as that of a foregrounded participant. This assumption, however, which has for a long time defined the approach taken to Lakota in this respect (e.g. Foley and Van Valin 1984, Van Valin 1980, 1985, 2001, Van Valin and Foley 1980), is called into question by the data presented in section 13.2.5.

13.2.5 Agent suppression plus overt agent NP: a ‘foregrounding passive’?
Recent language data reveal that Lakota is also equipped with a construction which is functionally equivalent to an English passive in which both agent and patient are overtly expressed. Just like the agent-dereferentializing -pi-impartial dealt with in section 13.2.1, this construction is based on the element -pi:

(57) Wicháša ki mathó O-kté-pí.
    man   def bear   3SG.PAT-kill-PASS

‘The man was killed by the/a bear/bears.’

13.2.5.1 General structure of the pi-passive In order to bring out the peculiarities of this construction, it is necessary first to discuss some general facts about argument marking in Lakota. Full NPs in core argument function are never case-marked. Word order can be taken as an indicator of agent vs. patient role in transitive clauses in which both agent and patient figure as full NPs: usually, the AGT precedes the PAT, as in (58), although occasionally the reverse is true, as in (59). Left-dislocation of the patient phrase in this example codes contrastive focus.

(58) Súįka hé igmúthąka hé theb-O-O-yé.
    dog   that mountain lion that STEM-3SG.AGT-3SG.PAT-eat up

‘The dog ate the mountain lion.’

(59) Súįka hé igmúthąka hé theb-O-O-yé.
    dog   that mountain lion that STEM-3SG.AGT-3SG.PAT-eat up

‘The mountain lion ate the dog.’

In Lakota, there is an elaborate system of determiners which includes several types of article conveying the basic distinction of definite vs. indefinite, and three types of demonstrative. At least with animate full NPs, marking by means of one or more of these types of determiner is mandatory in virtually all contexts. There are three marking formats:

(a) Article only:

(60) wicháša ki/wə
    man   def/1DF.SG

‘the/a man’
(b) Demonstrative only—the demonstrative obligatorily follows the lexical head in this case:

\[(61)\] 
\[\text{wichása hé} \]
\[\text{man that} \]
\['that man'\]

(c) Article plus demonstrative—here the demonstrative may either precede or follow the lexical head plus article complex:

\[(62)\] 
\[\text{hé wichása ki} \]
\[\text{that man DEF} \]
\['that man'\]

\[(63)\] 
\[\text{wichása ki hé} \]
\[\text{man DEF that} \]
\['that man' \]

One of the few contexts in which an animate NP may lack determiners is when it is used in a non-specific, non-referential, sometimes generic sense, like *ithýkala* 'mice' in (64):

\[(64)\] 
\[\text{Igmú cat ki def ith kýkala mouse} \]
\[\text{wichá-yuta-pi.} \]
\[\text{3pl.pat-3agt.eat-pl.agt} \]

'Bears killed the man.'

Given these facts, the construction exemplified by (57), in which the animate agent NP *mathó* 'bear(s)' follows the patient NP and lacks determiners, can be dealt with without positing a special, possibly passive-like, construction type. On the assumption that *mathó* 'bear(s)' designates the species in general rather than an individual bear or individual bears, and that *-pi* is a 3rd person plural marker which is coreferential with the agent *mathó* 'bear(s)', and also with a 3rd person agent marker *Ø-* which is attached to the verb, example (57), repeated here for convenience, can be translated by 'Bears killed this man'.

\[(65)\] 
\[\text{Wichása ki mathó Ø-Ø-kté-pi.} \]
\[\text{man DEF bears 3agt-3sg.pat-kill-pl.agt} \]

'Bears killed the man.'

As a matter of fact, however, there are three additional translations for the agent NP in example (57)/(65): 'the bear', 'a bear', and 'the bears'. These possibilities defy the rules for determiner use in canonical transitive and intransitive clauses in Lakota; in all three cases, the animacy of the agent NP requires the presence of determiners. What is more, the singular translations 'the bear' and 'a bear' are not compatible with the element *-pi* if the latter is analysed as a regular plural marker. As a number marker, in any case, *-pi* requires a plural interpretation. So the hypothesis that (57)/(65) represents a regular active transitive clause in which *-pi* fulfils the function of plural agent marking, rather than that of coding a passive-like structure, is not tenable. The numerous examples given in section 13.2.1 also show that impersonalizing *-pi* does not mark number. Interpreting *-pi* as an impersonalizing, passive-like marker in example (57)/(65) solves
the problems which arise when singular translations are given for the agent NP \textit{mathó} 'bear(s)'.

Confronting a Lakota native speaker with English passive clauses containing an agent has so far produced only the construction exemplified by (57)/(65) in translation: the patient, rather than the agent, is clause-initial; the agent does not combine with determiners; and the agent phrase is not restricted in terms of number, i.e. can be given either a singular or a plural reading.

Similarly, in example (66), -pi must be analysed as a passive rather than as a plural agent marker.

(66) \textit{Igmúańka hé źíka theb-Ø-ýá-pi.}

\text{mountain lion that \textsc{dog} STEM-3SG.PAT-eat up-PASS} \\
\text{‘That mountain lion was eaten by \textsc{dogs/a dog/the dog}.’}

Again, testing the hypothesis that -pi is a plural agent marker, \textit{igmúańka hé ‘that mountain lion’ cannot be interpreted as coreferential with the potential plural agent marker -pi because the phrase is unequivocally marked as singular by means of the demonstrative hé ‘that’, whose plural form is \textit{hená} ‘those’. This leaves źíka ‘\textsc{dog(s)}’ as the default agent. And since źíka ‘\textsc{dog(s)}’ can be translated as a singular in example (66), it must be concluded that, at least with singular interpretations of this NP, -pi codes an impersonal or passive rather than a plural agent. The same is true for examples (67) and (68):

(67) \textit{Wíchás.a hé mathó Ø-khuwá-pi.}

\text{man that bear \textsc{bear} 3SG.PAT-chase-PASS} \\
\text{‘That man was chased by \textsc{bears/a bear/the bear(s)}.’}

(68) \textit{Hoksíla hé wichícla a-Ø-phá-pi.}

\text{boy that \textsc{girl} STEM-3SG.PAT-hit-PASS} \\
\text{‘That boy was hit by \textsc{girls/a girl/the girl(s)}.’}

Given the above facts, there is only one possible alternative to assigning passive-like status to the -pi-construction in examples like (57)/(65): the agent NP can, at least theoretically, be analysed as an incorporated noun. In such a scenario, the pi-clauses in question must be considered intransitive at the structural level. The agent has been absorbed by the verb, and has therefore been removed from the valence frame of the verb. However, the incorporation hypothesis has to be rejected. First, transitive agents are not usually incorporated into the verb in Lakota, although this rare construction is documented in examples such as (69):

(69) \textit{Iyăyí na wayáka ye, táku źug-Ø-wáphapha-pi.}

\text{go and see IMP something dog-3AGT-bark at-PL.AGT} \\
\text{‘Go and look, dogs are barking at something.’}

More importantly, however, the agent in -pi-constructions does not behave like an incorporated noun in that it does not exhibit the stress pattern associated with
noun incorporation, which is most clearly observable when monosyllabic nouns are involved. These lose stress completely when incorporated, like *pté ‘buffaloes’* in the following example. In configurations like (70), stress will invariably fall on the first syllable of the verb, regardless of whether this syllable is stressed in the base verb or not:

(70)  
\[ \text{Pte-Ø-kté-pi,} \]  
\[ \text{buffalo-3AGT-kill-PL.AGT} \]  
‘They killed buffaloes.’

In the *pi*-clause (71), *pté ‘buffalo(es)’* figures as a non-incorporated noun which carries independent stress:

(71)  
\[ \text{Wicháša ki hé pté paslóháq Ø-ilpéya-pi.} \]  
\[ \text{man} \]  
\[ \text{def that buffalo knock down 3SG.PAT-throw-PASS} \]  
‘That man was pushed/knocked down by buffaloes/a buffalo/the buffalo(es).’

The stress pattern in (71) cannot be altered to produce the stress pattern characteristic for noun incorporation—i.e. removing stress from *pté ‘buffalo(es)’* produces an ungrammatical example:

(72)  
\[ *\text{Wicháša ki hé pte-pásloháq Ø-ilpéya-pi.} \]  
\[ \text{man} \]  
\[ \text{def that buffalo-knock down 3SG.PAT-throw-PASS} \]

13.2.5.2 Usage of the *pi*-passive  
Another issue that needs to be addressed in the context of putative passive constructions concerns the properties of the verbs eligible for passivization with respect to degrees of semantic transitivity according to Hopper and Thompson (1980). As a cross-linguistic rule of thumb, high-transitivity clauses lend themselves to passivization most readily (Hopper and Thompson 1980: 292–3). As a matter of fact, the examples given so far almost exclusively contain verbs which are high in Hopper and Thompson’s transitivity parameters (kinesis, agency, telicity, volitionality, and affectedness of the object): *kté ‘to kill’, *aphá ‘to hit’, *thebyá ‘to eat, devour’, and *paslóháq Ø-ilpéya ‘to push/knock down’. The only verb in the above examples that can be regarded as lower in semantic transitivity is *khuwá ‘to chase*. Below, more examples with both high- and lower-transitivity verbs are given:

High transitivity:

(73)  
\[ \text{Siná ki hé wichíšala Ø-kág-a-pi.} \]  
\[ \text{blanket} \]  
\[ \text{def that girl 3SG.PAT-make-PASS} \]  
‘That blanket was made by girls/a girl/the girl(s).’

(74)  
\[ \text{Joe wíšya ektá Ø-dya-pi.} \]  
\[ \text{Joe woman there 3SG.PAT-take to-PASS} \]  
‘Joe was taken there by women/a woman/the woman/the women.’
Argument dereferentialization in Lakota

(75) Wichäša hé mathó Ø-yahláka-pi/ỷubláza-pi/ksúyeya-pi.
man that bear 3SG.PAT-bite-PASS/tear apart-PASS/hurt-PASS

‘That man was bitten/torn apart/hurt by bears/a bear/the bear(s).’

The following examples with low-transitivity verbs are not grammatical:

(76) ∗Wichäša hé mathó Ø-waqyaka-pi.
man that bear 3SG.PAT-see-PASS

‘That man was seen by bears/a bear/the bear.’

(77) ∗Wichäša hé mathó na-Ø-h.’ú˛-pi.
man that bear STEM-3SG.PAT-see-PASS

‘That man was heard by bears/a bear/the bear.’

There is a tendency of admitting the -pi-passive preferably with high-transitivity verbs. This observation is in line with the behaviour of passive constructions at the cross-linguistic level.

Evidence of the use of the -pi-passive in contexts involving other types of argument, such as non-3rd person patients, is at best shaky. The same speaker provided diverging grammaticality judgements on the following two examples, which share the same basic structural profile. In (78), a singular translation, ‘a/the bear’—which, according to what has been said above, is the main clue to analysing the -pi-construction as passive-like—is possible, whereas in (79) the singular translation ‘a/the girl’ is ungrammatical.

(78) Mathó ni-yúblaza-pi.
bear 2SG.PAT-tear apart-PASS

‘You were torn apart by bears/a bear/the bear(s).’

(79) Wichícal’a a-má-pha-pi.
girl STEM-1SG.PAT-hit-PASS?

‘Girls hit me.’

The pi-passive can be used with both animate and inanimate agents. Examples with inanimate agents are given in (80) and (81):

(80) Mathó ki ḟyaq Ø-kat’á-pi.
bear DEF rock 3SG.PAT-kill-PASS

‘The bear was killed by a/the rock/rocks.’

(81) Thípi ki thätéqqa Ø-ihágya-pi.
house DEF hurricane 3SG.PAT-destroy-PASS

‘The house was destroyed by a/the hurricane.’

Acceptability of inanimate agents in pi-passives provides an additional criterion by means of which this construction can be set apart from its putative historical source, the plural marker -pi, since the semantic scope of the latter is limited to the
coding of animate arguments only. Because of its incompatibility with inanimate agents, -pi cannot be interpreted as a plural agent marker in examples like (80) and (81). Thus, in these cases, -pi must be considered a passive marker. For further hints that the Lakota -pi-construction is a passive, see Rood and Taylor (1996: 464).

The pi-passive in historical perspective

The pi-passive can be viewed as an innovative construction that has been formed in response to the influence of English syntax on Lakota. At least, the existing descriptions of the language do not include any data on passive-like constructions in which the agent can be overtly expressed, as is the case in the examples given above.

This hypothetical course of events would accord with similar developments in other languages (cf. Heine and Kuteva 2002: 235–7): that a morpheme which indicates agent or subject plurality develops into an impersonal, then into a passive with the agent unmarked, and ultimately to a construction which allows an overtly specified but pragmatically backgrounded agent. However, data from the closely related Siouan language Omaha make a different approach to the historical source of the pi-passive appear at least equally feasible. There is reason to hypothesize that the original (and still basic) meaning of the morpheme is ‘marked (unexpected) focus’, and that it has retained that meaning in the impersonals and passives we find in both Lakota and Omaha today. In addition, it has evolved from that meaning to indicate subject plural and, at least in Omaha, subject focus. The evidence for this proposal stems from comparing the morpheme’s Lakota distribution with that of a related morpheme in Omaha.

As described in Eschenberg (2005), Omaha also exhibits multiple functions for the cognates of the Lakota pi-morpheme. Like Lakota, it can mark either plural agents or impersonal constructions that translate into English most comfortably as passives. In addition, however, in Omaha the morpheme is used to indicate that the subject of a verb (either transitive agent or the only argument of an intransitive) is ‘on stage’, i.e. an important character in the narrative at that point in the story. The so-called ‘plural’ morpheme thus occurs with either singular or plural subjects when they are prominent. This is reminiscent of the proximate/obviative distinction found in Algonquian languages, but it is also different because more than one NP at a time may be ‘proximate’ in Omaha. Whereas the object is focused by the morpheme that marks 3rd person plural in the passive/impersonal construction in both languages, this Omaha development uses that same morphology to focus on subjects.

Since the morphemes in the two languages are phonologically cognate, it seems safe to assume that their ancestor in the proto-language from which they stem had a meaning which could evolve into agent plural marking, as well as into both agent or subject focus and object focus, and that it might have any one of these functions as its starting point. If, as is usually assumed, that starting point is plurality of subject, the path to the Lakota situation is easy to imagine, as stated above. But how does one go from plural agent to focused agent? Eschenberg proposes that
the development path is plural subject > object focus > backgrounded subject > subject focus, but the last step in this progression is not an evolutionary one; it is rather a complete reversal of functions. Such a development does not seem likely without considerable intermediation of some sort.

Starting with the idea that this morpheme meant 'marked focus', however, we can describe non-contradictory evolutionary paths leading to all the modern functions. One must first accept the idea that the neutral or unmarked focus in a transitive clause is on the agent or subject; we think that is well established and will not take the space to defend the claim here. If that is accepted, the existence of a morpheme to reverse the natural or expected focus placement becomes plausible, and we have the impersonal and passive meaning as the fundamental one for that morpheme. Positing a 'marked focus' meaning for that morpheme implies that its presence will signal either agent suppression or object prominence, i.e. the meanings we now see in impersonals or passives. Next, 'marked focus' could well come to mean 'defocus the expected focus'. If agents are defocused by default by the presence of a focus marker on non-agents, the identity of the agent might become so blurred and irrelevant—as in English impersonal they—that the 'marked focus' marker is reanalysed as plural agent. Next, 'plural agent' becomes 'plural subject', and we have accounted for everything except the Omaha subject focus construction. That would seem to be an extension in another direction from the original 'marked focus' meaning; one would go from 'marked focus' to simply 'focus' for the intransitives, and then to 'focused participant' for transitive agents. Plurality in this scenario is not on the path to the focus constructions, but a separate development, and there is no need to introduce and then erase the concept of number in the meaning of the morpheme.

If any reader has doubts about whether this proposal for the proto-morpheme and its development is realistic, please note the following. The Caddoan language Wichita, spoken in Oklahoma (also a semantic alignment language), has a morpheme with precisely these properties. The pronoun -iy- means fundamentally 'focus on 3rd person patient'. It is used for the singular subject of stative verbs (other persons use morphemes identical with those for transitive objects), for the subject of any verb if the object is the main character in the discourse, and as one of two ways of marking the plural agent of active and transitive verbs. This pronoun cannot occur in the same verb with any other pronoun (so it cannot be used to mark the object of a verb with a 1st or 2nd person subject), and it cannot be marked for number except in the stative verbs, where patient plural morphemes may be used with it. That the meaning 'focus on the patient' is the basic one is indicated by the way it interacts (or not) with other pluralizing morphemes, and by the fact that transitive verbs with this pronoun are ambiguous: they may have a plural agent, or they may have the meaning of 'agent is unimportant'. Thus, t-iy-kaʔacs, for example, means either 'They are eating it' or 'S/He is being eaten' (idiomatic for 'S/He has cancer'); contrast t-i-kaʔacs 'S/He is eating it'. In paradigm or list elicitation, fluent speakers often glossed this pronoun as 3rd person to 3rd person,
with the agent rendered 'the other guy' or 'somebody', or with contrastive stress on the 'him' of a 'He is verbing HIM' translation. Here is a clear case of a patient-focus morpheme being extended to cover the 3rd person plural agent function. We propose that something similar happened in the history of Siouan.

To sum up, if the focus-marking function of the Omaha cognate of Lakota -pi 'plural, passive, impersonal' can be proved to date back to Proto-Siouan times or at least a remote stage in the development of Siouan languages, it can be concluded that the -pi-passive in Lakota might not be an innovative but instead a quite ancient construction. Thus, the possibility arises that in Siouan, which has had a semantic alignment system for as far back as we can reconstruct such things, that split has coexisted with a passive or at least passive-like construction for a very long time.

13.3 Theoretical discussion of the data

One of the basic tenets of Role and Reference Grammar (e.g. Foley and Van Valin 1984, Van Valin 1985, 2001) is the postulate of a typological distinction between role-dominated and reference-dominated languages. This dichotomization evolves around the ways in which the syntax of individual languages is organized; the notion of 'pivot', which describes an argument type that controls clause-internal as well as clause-external syntax in a given language, is a central component of this approach.

Languages in which discourse factors are syntacticized in clause-internal grammar, i.e. languages with a PrP [= pragmatic pivot, R.P. & D.R.], are termed reference-dominated languages, and languages which do not have this syntacticization, i.e. languages with a SmP [= semantic pivot, R.P. & D.R.] only or no pivots at all, are labeled role-dominated languages. (Foley and Van Valin 1984: 123)

In the Role and Reference Grammar framework, Lakota has been identified as a role-dominated language. Grammatical constructions whose occurrence is intimately connected with pivot use are passives and antipassives. Traditional Indo-European-style passives are said to fulfill two pragmatic functions, foregrounding and backgrounding.

Passives which serve to remove the actor from the core of the clause are backgrounding passives, whereas those which function to permit a non-actor to occur as PrP are foregrounding passives. Foregrounding passives are normally found in languages which have PrPs, i.e. in languages in which the choice of pivot is governed by discourse facts...Backgrounding passives, on the other hand, are not so constrained and occur in both reference-dominated and role-dominated languages. (Foley and Van Valin 1984: 168)

Backgrounding functions to impose a structurally as well as pragmatically peripheral status on an argument—in the case of the passive, on the agent/actor. The backgrounded argument is either demoted to an oblique or eliminated from
the clause; either way, the pragmatic salience of its referent is reduced or lost. Foregrounding, on the other hand, elevates an originally less salient participant to a position of greater salience. A frequent structural correlate of this process is promotion, i.e. a shift from a syntactically less central to a more central case, such as from accusative to nominative in accusative systems, when patients/undergoers are to be coded. In a German passive clause, for instance, the original patient/undergoer, marked by the accusative in the active clause, receives nominative case marking. Thus, the nominative case in German can be interpreted as indicating pragmatic salience or foregrounded status.

Van Valin (1985: 368) argues that ‘[t]here is no Indo-European-style passive construction in Lakhota; that is, there is no construction in which the U [= undergoer, R.P. & D.R.] appears as the derived subject of a detransitivized verb with the A [= actor, R.P. & D.R.] either in an oblique phrase or deleted’.

According to Van Valin (p. 368), the closest equivalent to a passive in Lakota is the agent-suppressing impersonal construction dealt with in section 13.2.1. For the purpose of illustration, example (19) is repeated here for convenience:

(82) \textit{Thaló ki hé Ø-kablá-pi.} \\
\textit{meat \, DEF \, that 3SG.PAT-slice-AGIPS} \\
\textit{‘The meat was sliced.’}

Van Valin (1985: 368) analyses this construction as containing an explicit agent/actor specification by means of the element -pi; in section 13.2.1, this element was shown to occur in contexts in which such an interpretation does not apply. Rather, an interpretation as a more ‘abstract’ passive or, at least, impersonal marker, which lacks reference to any of the arguments involved, is in order in such cases.

To further evaluate the structural make-up and function of the Lakota passive with overt agent, as exemplified by (57)/(65), it is worth dealing in more detail with the syntactic and pragmatic status of the agent/actor and patient/undergoer. The question about potential foregrounding processes is particularly relevant here.

It is difficult to elucidate further the pragmatic status of the agent in examples like (57)/(65), repeated here for convenience:

(83) \textit{Wichás.a ki mathó Ø-kté-pi.} \\
\textit{man \, DEF \, bear \, 3SG.PAT-kill-PASS} \\
\textit{‘The man was killed by the/a bear/bears.’}

The agent phrase mathó ‘the/a bear/bears’ is not case-marked; although lack of case marking in Lakota is a feature common to valence-bound or core arguments, not every argument that is not valence-bound is case-marked in this language. Thus, absence of case marking does not provide further clues regarding the pragmatic status of the agent in -pi-‘passives’.

The structural criterion of word order sheds more light on this issue, supporting the hypothesis that the patient is foregrounded in examples like (83). The
word order ‘patient preceding agent’ in constructions of this type appears to be irreversible. And there is, in fact, some evidence that clause-initial position of nominal arguments is linked with high discourse salience in Lakota. Basically, word order (SOV) is utilized for the coding of case relations in Lakota. However, word order becomes available as a device for expressing pragmatic categories, such as foregrounded status, in cases in which the context or the semantic profile of the NPs themselves removes any ambiguity regarding the referential identity of the agent and the patient in a transitive clause. Thus, in (84), the patient NP lé wíchó’oyake ki ‘this story’ must be interpreted as more salient than the agent NP atéwayne kihá ‘my father’.

(84) Lé wíchó’oyake ki lé até-wa-ye kihá o-má-ki-yake.
    this story  DEF this father-1SG.AGT-have as DEF STEM-1SG-BEN-tell

‘This story was told to me by my father.’

If word order is indicative of foregrounding in Lakota, then the patient, which in putative passive constructions like (83) invariably precedes the agent, can be interpreted as foregrounded.

Aside from the observation that Lakota speakers regularly and spontaneously translate the -pi-construction by an English passive, another argument in favour of the hypothesis that the Lakota passive functions to foreground the patient can be derived from the very fact that this construction exists. It has often been claimed that case marking in semantic alignment languages is exclusively sensitive to role semantics—pragmatic categories such as topicality, focus, or foreground status are not seen as components of the functional load of the marking formats for arguments in such languages, insofar as basic grammaticalized marking formats are concerned. It goes without saying that any language, including semantic alignment languages, should be equipped with structural devices for marking certain pragmatic concepts, such as topic or focus, via left dislocation and other techniques, which can be used in an ad hoc manner when expression of such concepts is desired. The assumption that pragmatic categories ‘do not matter’ in the argument-marking system of the semantic alignment language Lakota begs the question of why Lakota is equipped with the -pi-construction exemplified by (83), whose expressive value with respect to role semantics is equivalent to that of a canonical active transitive clause because in both cases, both the agent and the patient are present. Of course, synonymy exists at all levels of language organization, and presumably in syntax as well, so that there may be constructions which are identical or at least quasi-identical with regard to meaning and syntactic function. Nevertheless, it appears counterintuitive to claim that there is absolutely no difference in meaning between a -pi-passive with an overtly expressed agent phrase and the corresponding active transitive construction.
The emergence of active/stative alignment in Otomi

ENRIQUE L. PALANCAR

14.1 Introduction

In this chapter, I show that Otomi intransitive verbs display both an agent/patient and an active/stative semantic alignment patterns. Agent/patient alignments have been widely described in the literature (to mention just a few, Merlan 1985, Van Valin 1990, Mithun 1991, Dixon 1994). Less is known about the active/stative type of alignment in languages such as Guaraní, as proposed in Mithun (1991). The goal of this chapter is twofold: first, I characterize the two alignment splits in this language, and secondly I describe the active/stative alignment in more detail in order to propose how this alignment may have emerged diachronically.

Otomí is spoken in Central Mexico and it belongs to the Otopamean group of Otomanguean languages. The label ‘Otomí’ in fact refers to a number of closely related languages. The phenomenon I study here is found in all varieties. Every language, however, has its own distribution of the phenomenon I study here, and differences in the treatment and scope of semantic alignments both on the lexical level and on the structural level are expected. While I also use other sources, most of the data in this chapter come from the variant spoken in the village of San Ildefonso Tultepec, in the state of Querétaro, Mexico. I refer to this dialect as San Ildefonso Otomi.

This chapter is organized as follows. In the following section, I present a brief overview of the grammatical features of San Ildefonso Otomi which are relevant to understand the object of study. Next, in 14.3 and in 14.4, I describe the semantic alignments found in the language. In 14.5, I study a subclass of relevant verbs pertaining to the active/stative alignment system, which I call ‘stative verbs’. In

1 A growing number of speakers of Mesoquital Otomi prefer to call their language ‘Hñähñu’, among other spellings. Speakers of San Ildefonso Otomi do not recognize the Mesoquital word as a representative name for their own language, ‘Hñoñoñ’ being their native denominative. ‘Otomí’, on the other hand, is a term which enjoys some history in linguistics and anthropology, and at present it is a denomination not regarded to my knowledge as a pejorative term by the majority of native indigenous people belonging to this ethnic group.
14.6. I introduce a special grammatical form of transitive verbs, labelled ‘stative form’, which I claim is related to stative verbs. In 14.7, I compare stative verbs and stative forms, and in 14.8 I present my proposal regarding the diachronic emergence of the active/stative alignment in Otomi. The chapter concludes in 14.9.

14.2 Overview of the language

San Ildefonso Otomi is a language with a basic SVO word order, although VS is also common in intransitive clauses. While noun phrases (NP) are not case-marked, the verb agrees with subject by means of proclitics which also encode Tense/Aspect/Mood (TAM). Examples in (1) illustrate three different proclitics: $dâ$ in (1a) encodes 1st person and completive past; $dî$ in (1b) encodes 1st person and present tense. Present tense is used to express habitual aspect. The zero morpheme in (1b), associated with the verb $hˇô$, encodes 3rd person and present tense:

\[ \begin{align*}
1. & \text{a. } dâ = hûd-i \\
& \text{Pâ } dâ = kô’-i - O-i \\
& \text{1.PST= sit.down.F-F to 1.PST=look-3OBJ-F}
\end{align*} \]

'I sat down to look at it.'

b. \( \text{yà nixî = r } \) hyéthê \( dî = tsi-Ø = hê \)

\( \text{ko yà in-té } \) \( O = hô \) \( \text{ya zà} \)

\( \text{because PARTICLE NEG-what 3.PRES= ingest-3OBJ= PL.EX} \)

'Ve don’t even eat mushrooms now, because there are no trees.'

A paradigm of the most common verbal proclitics found in San Ildefonso Otomi is given in Table 14.1.4

First and 2nd person objects show suffixal agreement, as in (2). Third person object is encoded by a zero morpheme, as illustrated in (3):

\[ \begin{align*}
2. & \text{a. } yà \text{ gá = tôn-k-i, } \text{pê } ga = peñg-i \\
& \text{PARTICLE 2.PST= win.S-1OBJ-F but 1.IRR= return-F}
\end{align*} \]

\[ \begin{align*}
3. & \text{gá } = \text{ tsix’-i } \\
& \text{1.IRR= take.animate.S-2OBJ-F}
\end{align*} \]

'You’ve beaten me (this time), but I’ll come back to get you.'

---

2 In a number of dialects, the imperfect is a distinct morph (e.g. $dî$-má˛-nu \[1.PRES-IMPERF-see\] 'I was seeing (it)'; from Toluca Otomi: Lastra 1992: 23).

3 I have used the convention of Hekking (2002) for the transcription of Otomi, with two innovations: $f$ for $l/ř$ and $b$ for a salient aspiration existing in San Ildefonso Otomi. Deviations from the IPA convention are: $^\prime \text{h}$ = $\text{p}^\text{b}$; $j$ = $\text{b}^\text{h}$; $y$ = $\text{j}$; $x$ = $\text{j}$; $tx$ = $\text{k}^\text{j}$; $n$ = $\text{j}$; $r$ = $\text{ř}$; $ř$ = $\text{ř}$; $a$ = $\text{l}$; $q$ = $\text{Ø}$; $g$ = $\text{l}$; $\text{q}$ = $\text{l}$; $y$ = $\text{l}$. Umlaut indicates nasalization (i.e. $\text{a}^\text{n}$; $\text{l}$; etc.). Otomi has three tones: ascending $\text{ˇv}$; high $\text{´v}$; and low, which is not represented. Given that tone differences between San Ildefonso Otomi and Mezquital Otomi are minimal, I have consulted the excellent transcriptions in Wallis (1956) and Hernández Cruz et al. (2004) for uncertain cases in San Ildefonso Otomi. Conventions: (-) morpheme boundary in source; (,) obscure morpheme boundary in source; (.) absence of morpheme boundary in source; < > non-concatenative morpheme; = clitic.

4 Paradigms can be very complex in other dialects (cf. Voigtlander and Echegoyen 1985).
Table 14.1. Verbal proclitics in San Ildefonso Otomi

<table>
<thead>
<tr>
<th>Subject + TAM</th>
<th>1st</th>
<th>2nd</th>
<th>3rd</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present (habitual)</td>
<td>di=</td>
<td>gi=</td>
<td>Ø/ bi ij=</td>
</tr>
<tr>
<td>Imperfect (past)</td>
<td>n-di=</td>
<td>n-gi=</td>
<td>mi=</td>
</tr>
<tr>
<td>Irrealis (future)</td>
<td>ga=</td>
<td>gi=</td>
<td>da=</td>
</tr>
<tr>
<td>Past (completive)</td>
<td>dä=</td>
<td>gä=</td>
<td>bi=</td>
</tr>
<tr>
<td>Perfect</td>
<td>(x)tà=</td>
<td>(x)kä=</td>
<td>(x)i=</td>
</tr>
<tr>
<td>Pluperfect</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(3)  ýá hín-té di=tô-Ø- j= ma bojó

PARTICLE NEG-what LPRES=win.b-3OBJ-B=1POSS-money
'I don’t earn my money.'

In Table 14.2, I present the object suffixes. Besides the use of affixes, a 1st and 2nd person object may optionally be cross-referenced in the verb by additional object enclitics, as illustrated in (4):

(4)  ýá gá=h<y> and-g-a=thó=gí

PARTICLE 2.PST=<NPS>SEE-3OBJ-B=DEL=1OBJ
‘You caught me red-handed.’ ~ (Mex. Sp.) ‘Me cachaste.’
(lit. ‘You just saw me.’)

Doubling object agreement by means of both an affix and an enclitic is very common in San Ildefonso Otomi. In principle, the enclitic must always co-occur with the suffix, but there are exceptions. When a verb has a disyllabic root, i.e. it does not have a transparent stem formative, e.g. künt’éi ‘keep a jealous eye on someone’, it cannot receive the object suffix, and object cross-reference is thus exclusively encoded by means of enclitics (e.g. mí= künt’éi=gí [3.IMPERF=keep.a.jealous.eye.on=1OBJ]) ‘He was keeping a jealous eye on me’ vs. *mí = künt’éi=g-a=gí). This distribution is, however, somehow reversed with stative verbs, which is a class of verbs highly relevant for this study (see 14.5). Not only can stative verbs receive the enclitic alone, regardless of the morphology of the root, but the use of the suffix alone is ungrammatical, unless in combination with the enclitic.6

In Palancar (2006), the combination of suffix plus enclitic was erroneously analysed as a case of doublet object morphology (e.g. -gagi). The present analysis is superior, as example (4) illustrates: here the delimitative clitic =thó occurs between the personal morphemes, implying that there is no morphological unity between -g-a and -gí. A similar phenomenon concerning double agreement marking has been described in Donohue (2003) for Skou, a non-Austronesian language spoken at the central north coast of New Guinea. In this language, verbs agree with subject by means of proclitics, but most commonly, consonantal prefixes are also used along with the proclitics.

5 In Palancar (2006), the combination of suffix plus enclitic was erroneously analysed as a case of doublet object morphology (e.g. -gagi). The present analysis is superior, as example (4) illustrates: here the delimitative clitic =thó occurs between the personal morphemes, implying that there is no morphological unity between -g-a and -gí. A similar phenomenon concerning double agreement marking has been described in Donohue (2003) for Skou, a non-Austronesian language spoken at the central north coast of New Guinea. In this language, verbs agree with subject by means of proclitics, but most commonly, consonantal prefixes are also used along with the proclitics.

6 The rationale behind this special distribution may lie in the fact that stative verbs do not have the inflectional contrast denoting free/bound from that other verbs have (see 14.5.3).
Table 14.2. Object markers in San Ildefonso Otomi

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>-k/-g</td>
</tr>
<tr>
<td>2nd</td>
<td>-'</td>
</tr>
<tr>
<td>3rd</td>
<td>-Ø</td>
</tr>
</tbody>
</table>

14.3 Semantic alignment in Otomi

Most intransitive verbs in San Ildefonso Otomi agree with S by means of the same proclitics used to encode A in transitive verbs, and differently from O, which shows suffixal agreement. This pattern, which is illustrated in (5) and (6), can be described as showing an accusative morphological alignment. Examples in (5) show the use of the proclitics di= and da=. The former is a marker of present tense and first person S, the second of irrealis and third person S. Both markers are used to encode A in (6):

(5)  S  a. ya = di = za = 'wâ = dî = 'bû = hû

particle 1.pres=be.agreeable=here 1.pres=live/be.s=pl.(incl)

'We are fine here, living together.'

b. Ø = hût = 'û = go_gé = da = <g> = ą = 'ā

3.sg=be.only=3sg rel.cleft(3) 3.irr=<nps>=remain.s=3sg

'That is the only thing that remains.'

(6)  A  a. nú = gá = di = pâ = Ø = nú = 'û = ya ts'oî
def=sg 1.pres=sell-3obj def=dem.pl pl pot

'I’m selling these pots.'

b. da = xu't'-O-a = nô = r = ts'oî

3.irr=wash.dishes-3obj-b=def.sg=sg pot

'She’ll wash up the pot.'

However, a number of intransitive verbs in Otomi index S by means of the same morphology encoding O in transitive clauses. The phenomenon is illustrated in (7), which may be compared to the examples in (4) and the morphemes in Table 14.2 above:

(7)  a. da = <z> = ō = k-a = gi

(3.irr=<nps>burn.s=1obj-b=1obj)

'I will/would get burned.'

b. hi-mî = pâ = 'a = 'i

neg-(3.)imperf=be.hot-2obj-b=2obj

'You weren’t hot.'

I use here the tripartite characterization of grammatical relations in Dixon (1994), which distinguishes three main categories: A, O, and S.
I call the intransitive verbs in (5) the A-verbs, while the verbs showing the object marking of S in (7) are called the O-verbs. O-verbs in (6) are inflected for TAM by means of proclitics, just like other verbs in the language. However, given that verbal proclitics in San Ildefonso Otomi cannot express TAM separately from grammatical person, O-verbs are always inflected by means of a neutral (or dummy) 3rd person proclitic. This may be seen in the glosses of (7), where the reference to this 3rd person has been placed in parentheses.

Given that the language has zero pronominal anaphora for a 3rd person object, as illustrated in (3) above, the semantic alignment shown in (7) is covert when the intransitive verb in question is inflected in 3rd person, as indicated in (8):

(8) pe' 'ingi Ø = káhi-Ø,
    kd = Ø = n-jú-Ø

but NEG \text{pres}=\text{be.tasteful-3obj} \text{intens=(3)} \text{pres}=\text{st-be.bitter-3obj}

‘But it isn’t tasteful, it’s quite bitter.’

14.4 Types of semantic alignment in Otomi

Otomi is an interesting language typologically because intransitive verbs in this language display the two types of semantic alignment proposed in Mithun (1991): agent/patient and active/stative. While A-verbs remain the same in the two alignments, the language has two types of O-verb. In this section I describe these two alignments separately, but I will focus later on the study and emergence of the active/stative alignment.

14.4.1 The agent/patient semantic alignment

A great number of intransitive verbs in Otomi display an agent/patient alignment. While the class of A-verbs here is enormous, the class of O-verbs is rather limited, containing a handful of verbs in each dialect. I call the O-verbs in this type of alignment ‘patient verbs’, as a way to distinguish them from the O-verbs in the active/stative alignment. For the variant of San Ildefonso Otomi, I have been able to identify only ten patient verbs so far, but there may well be more. Such verbs are given in (9):

(9) Patient verbs in the agent/patient alignment

a. tsá bi  ‘be/get tired’
b. pó bo  ‘be/get wet’
c. ‘yehmi ‘be/get drenched’
d. tsót‘i  ‘be/get burnt’
e. níh ti  ‘get one’s hair stand on end’
f. títs‘i  ‘be/get better (from an illness)’
g. tsá yé  ‘be/get worse (from an illness)’
h. sólo = së/hônt‘å  ‘be just/only’ (i.e. ‘it’s just/only)
i. t‘axkë 8 ‘be/get white/pale’
j. hât‘i  ‘be/get pale’

8 Both verbs t‘axki ‘be/get white’ and hât‘i ‘be/get pale’ are used of people when they get pale from an illness or from fright. The verb t‘axki is an inchoative verb related to the stative t‘axi ‘be white’.
The verbs *tsabi* 'be/get tired' in (9a) and *póbo* 'be/get wet' in (9b) have the relics of allomorphs of an old object marker *-b*, whose reflexes in the modern languages have been reanalysed as dative markers (cf. Palancar 2004a). Historically, the class appears to have been larger. Urbano (1990[1605]: 43) mentions another four lexical items in passing: *piti* (<piti>) 'be/get afraid'; *kwent'i* (<cuentti>) 'swoon'; *hiungi* (<hümuej>) 'be/get glad' (lit. 'sit-heart/belly'); and *hwäts'i* (<huättsi>) 'tremble'. All these verbs depict events that participants have but little control over. Lack of control is an important semantic feature in the emergence of similar splits in other languages (cf. e.g. Mithun 1991, Merlan 1985), and it could well have been the defining semantic feature of patient verbs in earlier stages of the language. Nowadays, many A-verbs depict uncontrolled events too. This may be seen in the non-exhaustive list in (10):

(10) A-verbs depicting uncontrolled events

- a. *seya* 'be/get angry'  
- b. *jóhyá* 'be/get happy'  
- c. *túnthå* 'be/get hungry'  
- d. *túthe* 'be/get thirsty'  
- e. *tági* 'fall down to a different plane'  
- f. *hwangi* 'fall down on the same plane'  
- g. *tsó* 'fall into a hole'

Similar agent/patient alignments have been described for other Otomanguean language groups, such as in Amuzgan and Chocho-Popolocan (for Amuzgo, see Smith-Stark and Tapia García 2002; for Chocho, Veerman-Leichsenring 2000; for Popolocan, Swanton 2005; and for Tlapanec, Wichmann 1996). It remains an open question whether similar semantic alignments exist in other families of this macro-phylum.

### 14.4.2 The active/stative alignment

Intransitive verbs in Otomi also display an active/stative alignment, similar to that found in other languages from South America such as Guaraní (Mithun 1991, Velazquez-Castillo, this volume) and Kurripako (Danielsen and Granadillo, this volume). The class of O-verbs in this active/stative alignment is much larger than the class of O-verbs in the agent/patient pattern we saw in the previous section. I will refer to the O-verbs in this alignment as 'stative verbs'. Stative verbs depict property concepts, which are the type of concept commonly encoded with adjectives in other languages (e.g. the property of being 'red', 'big', 'long'). A number of verbs of this class are given in (11) (see Appendix 14.1 for a non-exhaustive list of such verbs, and Palancar 2006 for more details):

(11) A few O-verbs of the active/stative alignment

- a. *pídi* 'be thick'  
- b. *kú* 'be wet'
Active/stative alignment in Otomi

A number of A-verbs also express property concepts which are ascribable to human beings, even though a number of them may be extended to characterize animals too. A relevant list of such A-verbs in San Ildefonso Otomi is given in (12):9

(12) Dimension
dójá  ‘be big’  (also Human)
tx’ajo  ‘be small’  (also Human)
yá’pu  ‘be far away’

Time/Age
tsu  ‘be old (woman)’  (only Human)
xíta  ‘be old (man)’  (only Human)
tsj’unt’u  ‘be young (man)’  (only Human)
ntsjuntsi  ‘be young (woman)’  (only Human)

Propensity
nzá tho  ‘be beautiful’  (only Human)

Physical
tse = tho  ‘be strong’  (only Human)

The last two verbs in (12) can also be inflected as stative verbs, but when they do so, they depict an inherent property. This is shown in (13a) if compared with (13b):10

(13) a. A-verb  
\[
dí = nzá tho
\]
\[
1.prs=be.beautiful
\]

‘I look fine (because something becomes me).’

\[
dí = tse = tho
\]
\[
1.prs=be.strong=del
\]

‘I’m strong (I feel so).’

b. O-verb  
\[
xi = nzá tho = gí
\]
\[
3.perf=be.beautiful=1obj
\]

‘I’m beautiful (by nature).’

\[
xi = n-tse = gí
\]
\[
3.perf=st=be.strong=1obj
\]

‘I’m strong (by nature).’

9 The verbs xíta  ‘be old (for a man)’, tsj’unt’u  ‘be young (for a man)’, and ntsjuntsi  ‘be young (for a woman)’ are derivations from the nouns xíta  ‘old man’, tsj’unt’u  ‘young man’, and ntsjuntsi  ‘daughter’. The lexeme tsu receives low tone when it is a verb ‘be old (for a woman)’ but high tone when a noun ‘old woman’.

10 Similarly, the verb solo=se [be.only=alone] ‘be just/only’ in (8g) may also inflect both as an A-verb and as an O-verb, but without an apparent semantic distinction, e.g. (patient verb) mísolo=se=gí=gí =gí [(3).imperf=be.beautiful=1obj] vs. (A-verb) n_di=solo=se [imperf.1.prs=be.only=alone] ‘I was alone/it was only me’.
A summary of the semantic alignment patterns

A summary of the semantic alignment patterns found in Otomi is given in the above figures. Figure 14.1(a) represents the body of intransitive verbs in the language, which has two main classes: the O-verbs and A-verbs. Figure 14.1(b) represents intransitive verbs that display an agent/patient alignment. Recall that the O-verbs in this alignment pattern are called ‘patient verbs.’ Figure 14.1(c) shows the intransitive verbs of the language that display an active/stative semantic alignment. In this alignment pattern, O-verbs are called ‘stative verbs’ (O-verbs considered here do not include patient verbs).

At least one O-verb in San Ildefonso Otomi may be inflected both as a patient verb and as a stative verb, with a slightly different meaning. The verb is hˇat’i ‘be/get pale,’ and is shown in (14):

\[(14)\]

a. Patient O-verb

\[\text{xi}=h<y>\hat{a}t-k-a=gí\]

(3.)PERF=\text{<NPS>\text{be/get.pale.s-1OBJ-b=1OBJ}}

‘I’m pale (as a result of a long illness).’

b. Stative O-verb

\[\text{xi}=\hat{n}-\hat{a}t-k-a=gí\]

(3.)PERF=\text{ST-be/get.pale.s-1OBJ-b=1OBJ}

‘I’m pale (as a result of a fright).’

14.5 A closer view of stative verbs

Apart from the encoding of their semantic argument as O—the defining feature of the class—stative verbs have three further characteristics that differentiate them clearly from most A-verbs: (a) they use a special paradigm of proclitics; (b) they have a morphologically conditioned nasal prefix; and (c) they lack a morphosyntactic bound form. I study each of these features separately in the following subsections. For each case, I additionally present another realm of the grammar where each characteristic is also found, but without an apparent or obvious connection with stative verbs. In my opinion, the idiosyncrasies of stative verbs point to a complex historical relationship with those areas of grammar.
where the characteristics are found; they reveal the origin of these verbs and, ultimately, provide hints concerning the emergence of the active/stative semantic alignment in Otomi.

14.5.1 The TAM inflection of stative verbs

Stative verbs inflect in the imperfect and the irrealis with a special set of verbal proclitics. Such proclitics can also be used with all the A-verbs that express property concepts (PC) in (12) above. The pattern is shown in the following examples. The proclitics már = in (14a) and dar = in (16a) express the imperfect and the irrealis for stative verbs. As (15b) and (16b) show, these proclitics may be used with A-verbs that express property concepts but not with others, as the impossibility of (15c) and (16c) indicates. Notice that the subject translation is different when a stative verb and an A-verb are used as the proclitics with A-verbs cross-reference subject:

(15) Imperfect Stative verb
   a. már/*mi = n-hêts'i = gi
      (3),IMPERF,PC/3,IMPERF=st-be.high-1OBJ
      'I was tall.'
   PC A-verb
   b. már/mi = dôtâ
      3,IMPERF,PC/3,IMPERF=be.big
      'He was tall.'
   Other A-verb
   c. *már/mi = 'bji
      3,IMPERF,PC/3,IMPERF=be.standing
      'He was standing.'

(16) Irrealis Stative verb
   a. dar/?da = n-hêts'i = gi
      (3),IRR,PC/3,IRR=st-be.high=1OBJ
      'I will/would be tall.'
   PC A-verb
   b. dar/da = dôtâ
      3,IRR,PC/3,IRR=be.big
      'He will/would be tall.'
   Other A-verb
   c. *dar/da = 'bji
      3,IMPERF,PC/3,IMPERF=be.standing
      'He will/would be tall.'

The morphemes már and dar emerged from the reanalysis of clitic morphology involved in the nominal predication construction. In this construction, as illustrated in (17), the predicative NP appears inflected for number and is preceded by a special paradigm of proclitics marking person and TAM, some of which are only used in nominal predication. In (17a) and (17b), the singular enclitic = r is hosted by the proclitics má= and da=. These proclitics encode imperfect and irrealis of a third person.11 The morphological combinations má = r and da = r gave way to the

11 Because they occur in a nominal predication construction, Ecker (1952) and Hess (1968) erroneously treat these markers as inflectional forms of an alleged copular verb.
synchronic proclitics már and dar used with stative verbs. Notice that in (17c), the number of the NP is plural, but the combination má = ya used in this predication was not reanalysed any further. This may be seen in (17d), where a proclitic such as máya does not exist, and the verb receives the proclitic már = with a plural subject:

(17) Imperfect+ SGNP a. má= [r bë]NP
3.IMPERF.NP=SG thief
‘S/he was a thief.’

Irrelais+ SGNP b. da= [r mähjô]NP
3.IRR=SG priest
‘He’ll become a priest.’

Imperfect+ PLNP c. má= [ya bë]NP
3.IMPERF.NP=PL thief
‘They were thieves.’

Stative verb d. nu=ya ndo­már/*máya n-ts’o-Ø
DEF=PL man (3.)IMPERF.PC ST-be.evil-3OBJ
‘The men were evil.’

14.5.2 The nasal prefix of stative verbs

Stative verbs receive a nasal prefix N- I call the ‘stative prefix’. The stative prefix is not used when the verbal stem has a nasal consonant in initial position (e.g. mè ‘be hard’, noho ‘be fat’). The prefix has three allomorphs with regular phonological conditioning: [n-] <n> in (18a) is the default realization of the morpheme; [ni] <ń> in (17b) appears before a glottal consonant ([ʔ, h]); and [m-] <m> occurs before a labial consonant ([p, ph, b]), as in (18c):12

(18) a. bi = <z> i-Ø
 ya t’éi
 nu xi = n-k’angi-Ø
3.PST=<NPS> ingest-3OBJ PL REL (3.)PERF=ST-be.green-3OBJ
‘They ate green wheat.’ (lit. ‘They ate the wheat which is green.’)

b. má = tā-n-hês’i = ‘i
(3.)IMPERF.PC=INTENS-ST-be.high=2OBJ
‘You were very tall.’

c. xi = m-pà-g-a = gi
(3.)PERF=ST-be.hot-1OBJ-B=1OBJ
‘I’m hot.’

The prefix N- is not used in cases where the stative verb receives the same TAM morphology as A-verbs. This happens, for example, in the negative of both the present and the imperfect, where the use of the prefix is ungrammatical. As an illustration consider (19) and (20), which show two circumstances where the

12 When the nasal is attached to a stem with an initial glottal-labial consonant cluster (e.g. [ʔpX] or [hmX], as in ‘bèxiü ‘be dark’ or hmôdi ‘be expensive’), the labial is dropped, and the nasal assimilates the labial feature, as [m-7X] or [m-hX] (e.g. m’-èxiü or m-hôdi).
inflection of a stative verb coincides with that of an A-verb. In such cases, the stative prefix is not used:\(^{13}\)

(19) Negative present

a. A-verb

* híngi Ø=né-i

NEG 3.PRES=dance-F

‘He isn’t dancing.’

b. Stative verb

* híngi Ø=(∗n-)tsé¯-g-a=gi

NEG (3.PRES=(ST-))be.cold-1OBJ-B=1OBJ

‘I’m not cold.’

(20) Negative imperfect

a. A-verb

* hí-mí=né-i

NEG-3.IMPERF=dance-F

‘He wasn’t dancing.’

b. Stative verb

* hí-mí=(∗n-)tsé¯-g-a=gi

NEG-(3.IMPERF=(ST-))be.cold-1OBJ-B=1OBJ

‘I wasn’t cold.’

A few stative verbs denoting colours may appear as nouns designating the colour in question; as nouns, they always bear the stative prefix N-:

(21)

n-t‘áxi ‘white (as a colour)’
n-théni ‘red’
n-k‘áxt’i ‘yellow’
n-k‘ängi ‘blue/green’

14.5.3 The lack of the free/bound form inflectional contrast

The vast majority of A-verbs in Otomi have stem formatives (e.g. tsé+t‘-i ‘get cold’; k‘ö+ts‘-i ‘go to see someone’; zó+n-i ‘weep’ etc.). In Palancar (2004b) I claim that verbs with stem formatives have two inflected forms which I called a ‘free form’ and a ‘bound form’. A free form is phonologically free and is used before a prosodic and/or a clausal boundary. An example is given in (22a). A bound form, in contrast, is phonologically bound and occurs intraclausally, as shown in (22b):

(22)

a. Free form

[da=tsé¯t‘-i]]

3.IRR=get.cold-F

‘It’ll get cold (i.e. something in a fridge).’

b. Bound form

[da=tsé¯t‘-a=ma ngó [Ø da=tá-Ø-i]]

3.IRR=get.cold-B=1POSS meat REL 1.PST=buy-3OBJ-F

‘The meat I bought is going to get cold (i.e. in the fridge).’

\(^{13}\) In contrast, examples (14) and (15) above show two cases where a stative verb does not receive the same inflectional morphology of A-verbs. Notice that the stative prefix was used in such cases.
These forms are not interchangeable. For example, a free form may not be used in the context of a bound form as shown in (23):

\[(da=ts\acute{e}t\acute{e}i \ ma \ ng\acute{o})\]

\[3.\text{IRR}=\text{get.cold} \cdot 1 \text{POSS} \text{meat}\]

Intended reading: ‘My meat is going to get cold.’

On the other hand, stative verbs lack this inflectional contrast, despite the fact that a considerable number of them have stem formatives just like A-verbs (e.g. ts’út’i ‘be thin (wood, thread)’; hé+ts’-i ‘be high’; thé+n-i ‘be red’). Consequently, the /i/ in the citation form of such verbs should be considered part of the root. The pattern is illustrated by the following examples. In (24b) we may observe that the stative verb ts’út’i ‘be thin’ does not undergo the same morphological adjustment as the verb ts\acute{e}t’i ‘get cold’ in (22b). In fact, the use of the expected bound form of such a verb is impossible, as shown in (24c):

\[(24)\]

\[a. \ n\acute{o}=r \ z\acute{a} \ h\text{íngi} \ O=ts'\acute{u}t'\text{-}i\text{-}O \]

\[\text{DEF.SG}=\text{SG} \text{WOOD} \text{NEG} \ (3.)\text{PRS}=\text{be.thin}-\text{3OBJ}\]

‘The tree isn’t thin.’

\[b. \ h\text{íngi} \ O=ts'\acute{u}t'\text{-}i\text{-}O \ n\acute{o}=r \ z\acute{a} \]

\[\text{NEG} \ (3.)\text{PRS}=\text{be.thin}-\text{3OBJ} \text{DEF.SG}=\text{SG} \text{WOOD}\]

‘The tree isn’t thin.’

\[c. *h\text{íngi} \ O=ts'\acute{u}t'\text{-}O \text{-}a=n\acute{o}=r \ z\acute{a} \]

\[\text{NEG} \ (3.)\text{PRS}=\text{be.thin}-\text{3OBJ-B}=\text{DEF.SG}=\text{SG} \text{WOOD}\]

Intended reading: ‘The tree isn’t thin.’

Beside a couple of irregular lexical items, stative verbs are the only verbs that lack the inflectional contrast of verbs with stem formatives. In this respect, stative verbs behave morphologically just like other nouns that also have stem formatives (e.g. ’bos+p-i ‘ash’; t’á+f-i ‘candy’; ó+\text{ng}+n-i ‘poultry’, gít+n-i ‘incense’). Nouns never show the inflectional contrast I have presented in this section.

14.5.4 Summary concerning stative verbs

In the previous sections I have described a number of morphological idiosyncrasies of stative verbs; these verbs are the O-verbs of the active/stative semantic alignment of Otomi. A number of interesting features have been presented: (a) stative verbs receive the proclitics már= and dar= for the imperfect and the irrealis, and such markers originally developed from the inflectional morphology involved in nominal predications; (b) stative verbs receive a nasal prefix, which is used in the cases where the inflectional paradigm diverges from A-verbs; and (c) unlike A-verbs, the stative ones lack the free/bound form inflectional contrast, and they resemble nouns in this respect, at least on morphological grounds.
I propose a hypothesis about how stative verbs may have developed diachronically. This development sheds important light on how the active/stative alignment emerged in the grammar of Otomi. Before the active/stative alignment pattern developed, Otomi had only an agent/patient alignment. I propose here that stative verbs were once former A-verbs that denoted property concepts. In my view of the process, speakers used A-verbs for the linguistic expression of property concepts, but at times they would also use a special ‘stative form’ of such verbs for the same purpose. Such a stative form is described in some detail in the following section. Stative verbs emerged as the outcome of a blending of inflectional features pertaining to the A-verbs involved and of features proper to stative forms. One of the main characteristics borrowed from the latter was the encoding of the semantic argument of the verb as an object. This change motivated the emergence of stative verbs as a subclass of intransitive verbs, which triggered the appearance of a new active/stative semantic alignment in the language.

The purpose of the remaining part of this chapter is to elaborate on this development step by step. In order to achieve such a goal in 14.8, I will have to describe first what a ‘stative form’ is in the following section, and then compare such forms with stative verbs in 14.7.

14.6 Stative forms of transitive verbs

In all dialects of Otomi, transitive verbs have a special form, which has been called ‘participle’ in the Otomian descriptive tradition (cf. Ecker 1952, Hess 1968, Hekking and Andrés de Jesús 1984, Voigtlander and Echegoyen 1985, Bartholomew 2004; etc). I prefer to call such forms ‘stative forms’, more in tune with the description of the verbal system in other Otomanguean languages, such as Zapotec (cf. Pickett et al. 2001).

Stative forms express the state resulting from the action depicted by a transitive verb. Two examples are given in (25). These examples show the stative forms of the transitive verbs ‘ôts’i’ ‘peel’ and ‘ho’ts’e’ ‘cook (by boiling)’ and portray the actions they depict as resulting states:

\[(25)\]
\[
a. \quad n-t-ôts’i \quad \text{ST-IMPER-peel.F-F} \quad \text{‘Peeled.’} \\
b. \quad n-t-ho’ts’e \quad \text{ST-IMPER.cook.by.boiling.F-F} \quad \text{‘Cooked (by boiling).’}
\]

Stative forms are morphologically built on impersonal stems plus the stative prefix N-. I claim that Otomi had an agent/patient alignment pattern before it developed the active/stative one. It remains an open question why property concepts were not at this stage expressed by the patient verbs in the original agent/patient alignment. It seems, though, that the semantics of such verbs was confined to the expression of uncontrolled affecting events, and the crucial point was that properties were not regarded as events.
in (25) are built on the impersonal stems \( t^\text{-}\acute{\text{o}}\text{ts}' \) ‘imper-peel’ and \( t^\text{-}\acute{\text{h}}\text{o}^\text{-}\text{ts}' \) ‘imper-cook (by boiling)’ plus stative morpheme \( N^\text{-} \). However, the morphological process by which stative forms are created is no longer productive synchronically. Consequently, many verbs lack a stative form. The phenomenon is perhaps better understood as lexically conditioned, and the verbal lexemes that have a stative form vary considerably from dialect to dialect.

Given that stative forms and impersonal forms are related, it is relevant first to describe what an impersonal form is. The impersonal form of a transitive verb is used as an exponent of ‘impersonal voice’. Functionally speaking, impersonal voice has the same function as the prototypical passive voice proposed in Shibatani (1985): it serves as a de-agentivizer construction where the agent is not present, because it is either unknown or not relevant. An example is given in (26) with the impersonal form \( \text{tho}^\text{-}\acute{\text{t}}\text{se} \) of the verb \( \text{ho}^\text{-}\text{ts}' \) ‘cook by boiling’:

\[
(26) \quad \text{bi}=t^\text{-}\text{ho}^\text{-}\text{ts}'-\text{Ø}-a=ma \quad \text{jü}
\]

\[
(3.)\text{pst}=\text{imper-cook.by.boiling.B-3OBJ-B=1POSS}\text{ beans}
\]

‘My beans were boiled (by someone whom I don’t know).’ (lit. ‘Boiled my beans.’)

The event in (26) is presented as happening without a specified agent. The construction does not allow the overt expression of an agent, be it an indefinite subject, as in (27a) or an oblique participant, as in (27b):

\[
(27) \quad \text{a.} \quad \ast \text{n’a bi}=t^\text{-}\text{ho}^\text{-}\text{ts}'-\text{Ø}-a=ma \quad \text{jü}
\]

\[
\text{one} \quad (3.)\text{pst}=\text{imper-cook.by.boiling.B-3OBJ-B=1POSS}\text{ beans}
\]

Intended reading: ‘Someone boiled my beans.’

\[
(27) \quad \text{b.} \quad \ast \text{bi}=t^\text{-}\text{ho}^\text{-}\text{ts}'-\text{Ø}-a=ma \quad \text{jü} \quad \text{ká ma bötsi}
\]

\[
(3.)\text{pst}=\text{imper-cook.by.boiling.B-3OBJ-B=1POSS}\text{ beans LOC 1POSS child}
\]

Intended reading: ‘My beans were boiled by my child.’

Crucially, however, the impersonal voice is a subjectless structure that does not promote O to S; the patient remains morphologically encoded as object. This feature will prove to be important for the emergence of stative verbs.

On the other hand, as TAM cannot be morphologically dissociated from grammatical person, the verb in the impersonal voice is inflected with a dummy 3rd person proclitic only to express TAM. As with stative verbs, reference to this dummy grammatical 3rd person has been placed in parentheses throughout, except in example (27).

The following examples show the contrast between the active rendering of the verb \( \text{handi} \) ‘see’ in (28a) and its use in the impersonal voice in (28b). Example (28a) is an instance of the transitive verb \( \text{handi} \) ‘see’, which has two arguments: first, a 3rd person who plays the role of an agent (or an experiencer) and is encoded as subject by means of the proclitic \( \text{bi} = \); and second, a 1st person who
plays the role of a patient (or a stimulus) and is encoded as object by means of both the suffix -g and the clitic =gi. In (28b), in contrast, the verb appears in the impersonal voice, and has the agent argument cancelled, thus having only one argument left: the 1st person. Semantically, the agent is still present—the action is understood as instigated by some agency after all—but the construction just focuses on the resultant state in which the patient is found after the event has taken place. However, this patient is not grammatically treated or raised as subject, but remains an object, just as in (28a), while the construction becomes subjectless:

(28) Active form a. bi=hydr=g-a=gi
   3.PST=<NPS>SEE-1OBJ-B=1OBJ
   'He saw me.'

Impersonal form b. bi=thand-g-a=gi
   (3.)PST=IMPER-SEE-1OBJ-B=1OBJ
   'I was seen.' (lit. 'Saw me.')

On the other hand, speakers use perfect morphology with impersonal forms to encode present time reference, and the reading obtained is clearly one of resultant state:

(29) a. nô=r thuhme yá xi=t-hg\^i=k-Ø-i
   DEF=SG bread PARTICLE (3.)PERF=IMPER.CUI-F-3OBJ
   Resultant: 'Someone has cut the bread/the bread's been cut.'
   (lit. 'The bread, has cut it.')

b. xi=t-hít-\~\-a='í
   (3.)PERF=IMPER-COMFORT-2OBJ-B=2OBJ
   Resultant: 'You've been comforted.' (lit. 'Has comforted you.')

Predicates with stative forms are based on the impersonal voice. Stative forms cancel the presence of the agent completely, and present the participant as just being in a given state. When a verb happens to have both forms, we obtain subtle semantic contrasts. An example of one such contrast is given in (30):

(30) Impersonal form a. hin=txi=k<h>óh-k=gi
   NEG-(3.)PERF=<IMPER>dress.someone-1OBJ-B=1OBJ
   Resultant: 'I'm not (properly) dressed.'
   (e.g., a nurse has left me in such a state).
   (lit. 'Hasn't dressed me').

Stative form b. hin=txi=n-k<h>óh-k=gi
   NEG-(3.)PERF=ST<IMPER>dress.someone-1OBJ-B=1OBJ
   State: 'I'm not (properly) dressed.'
   (e.g., I think that gowns in this hospital are too short).
   (lit. 'Isn't dressed me').

To conclude this section, it is important to mention that a large number of stative forms may be converted into nouns. As nouns, such forms are commonly action
nouns, result nouns, or instrument nouns. A number of examples are given in (31):

(31) Action noun: Mezquital Otomi a. \( n\text{-}<s>\text{a}\ha \) 
\( \text{st-<imper>bathe} \) 
'bathe, bathing'

Mezquital Otomi b. \( n\text{-}t\text{-}\text{'un-the} \) 
\( \text{st-<imper-give.10.3-water.B} \) 
'watering, irrigation'

Result noun: Mezquital Otomi c. \( n\text{-}t\text{-}\text{'ut'i} \) 
\( \text{st-<imper-sprinkle} \) 
'manure'

Instrument noun: San Ildefonso Otomi d. \( n\text{-k<h>\text{'om'i} \) 
\( \text{st-<imper-cover} \) 
'lid, cover'

In summary, stative forms of transitive verbs are intransitive forms that express states. When used predicatively, their notional subject is still encoded as object. Besides, while morphologically based on impersonal stems, these forms receive the same prefix used with stative verbs. In the following section, I shall compare the stative forms with stative verbs in order to show how they can be related.

14.7 Comparison of stative verbs and stative forms

14.7.1 Similarities between stative verbs and stative forms

Both stative verbs and the stative forms of transitive verbs are intransitive verbal forms with only one argument. In both cases, this argument is encoded in the verb by means of object suffixes. Semantically, both forms express states. Morphologically, they use the perfect proclitic \( xi= \) to express a present state of affairs, and they both use the stative prefix \( N\text{-} \). All these similarities may be observed in (32):

(32) Stative form: a. \( xi=n\text{-}t\text{-h\text{'ot-k-i} \) 
\( (3.)\text{perf=st-<imper-comfort-1obj-f} \) 
'I\’m comforted.'

Stative verb: b. \( xi=n\text{-h\text{'ets'}i=g}i \) 
\( (3.)\text{perf=st-be.high=1obj} \) 
'I\’m tall.'

Finally, although with lexical restrictions, both forms may convert into nouns and may be used as such. Examples are given in (33):

15 Most abstract nouns in Otomi were once derived from stative forms, but nowadays many of these nouns are no longer recognized by speakers as being related to their cognate verbal lexemes. Additionally, when a verb lacks a stative form, the impersonal form may be used for similar purposes.

16 The examples from Mezquital Otomi come from Bartholomew (2004).

17 The morphology is obligatory for stative verbs but optional for stative forms.
Active/stative alignment in Otomi

Noun from a stative form: \( n-t\text{-hèni} \)
- \( n-t\text{-ts<'}\text{-é-zà} \)
  - ST-IMPER-SAW
  - 'saw'
Noun from a stative verb: \( n-t\text{-i} \)
- \( n-t\text{-dixi} \)
  - ST-BE.SPICY
  - ST-BE.WHITE
  - 'chili'
  - 'white (the colour)'

Having presented the similarities in this section, I introduce the main differences in the following section.

14.7.2 Differences between stative verbs and stative forms

Stative verbs and stative forms show a number of differences. While stative forms are always related to synchronically attested transitive verbs, stative verbs are an independent class of verbs and do not derive from any other verbs, transitive or intransitive.

Similarly, in predication, stative forms use the same proclitics used with other A-verbs, while they cannot co-occur with the typical proclitics employed with stative verbs. This is shown in (34):

(34) Imperfect

Stative form: a. \( mi/\text{már}=n-j\text{-jùn}-\text{Ø}-i \)
  - (3.)IMPERF/(3.)IMPERF.POCL=ST-GRIND.CORN.F-3OBJ-F
  - 'It (the corn) was ground.'

Stative verb: b. \( *mi/\text{már}=m-pá-g\text{-a}=\text{gi} \)
  - (3.)IMPERF/(3.)IMPERF.POCL=ST-BE.HOT-1OBJ-B=1OBJ
  - 'I was hot.'

On the other hand, stative forms have the inflectional contrast between a free and a bound form, while stative verbs do not. This is shown in the following examples. Examples in (35) show that stative forms have this inflectional contrast like other A-verbs, while examples in (36) illustrate that stative verbs lack such a contrast:

(35) Stative form

a. \( ma\text{ suni} \)
  - (IPOSS CORN.DOUGH (3.)IMPERF=ST-(IMPER.)GRIND.CORN.F-3OBJ-F
  - 'My corn-dough was ground.'

b. \( mi=n-jùn\text{-Ø}-a=\text{ma} \)
  - (3.)IMPERF=ST-(IMPER.)GRIND.CORN.B-3OBJ-B=IPOSS CORN.DOUGH
  - 'My corn-dough was ground.'

(36) Stative verb

a. \( ma\text{ zà\text{-már}=n-hèts'i-Ø} \)
  - (IPOSS TREE (3.)IMPERF.POCL=ST-BE.HIGH-3OBJ
  - 'My tree was tall.'
Finally, on the morphological level, stative forms are always used with the nasal prefix in all grammatical contexts, unlike stative verbs, which do not use it in cases where the TAM inflection coincides with A-verbs. This is shown in (37). In (37a), the stative prefix is not used when a stative verb receives the same proclitics as other A-verbs. However, in similar circumstances, the prefix is obligatory with a stative form, as in (37b):

(37) Stative verb: a. \( \text{híngi } \text{Ø}=(*m)-\text{pá-g-a=gi} \)  
\( \text{NEG} \quad \text{pres}=(\text{st})-\text{be.hot-1OBJ-B=1OBJ} \)  
'\( \text{I'm not hot.} \)'

Stative form: b. \( \text{tó'be híngi} \)  
\( \text{NEG} \quad \text{pres}=(\text{st})-\text{be.hot-1OBJ-B=1OBJ} \)  
'\( \text{It isn't cooked yet.} \)'

14.7.3 Summary of the comparison
A summary of similarities and differences seen in 14.7.1 and 14.7.2 is presented in Table 14.3. Despite the differences, the morphological similarities between

<table>
<thead>
<tr>
<th>Features</th>
<th>Stative forms</th>
<th>Stative verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Similarities</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intransitive</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>O encoding of semantic argument</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Use of perfect conveys present</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Semantic states</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Convert into nouns (restricted)</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Differences</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Derived forms</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Stative prefix in all contexts</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Inflectional contrast (bound/free)</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Use of ( \text{mí} = ) for Imperfect</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>(like A-verbs)</td>
<td>(\text{YES only in negative})</td>
<td></td>
</tr>
<tr>
<td>Use of ( \text{már} = ) for Imperfect</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>
stative verbs and stative forms are too great to be coincidental. In the following section, I propose that stative verbs were once former A-verbs whose inflection incorporated a number of features from stative forms, giving rise to a new class that would motivate the emergence of the active/stative semantic alignment of Otomi.

14.8 The emergence of the active/stative alignment in Otomi

In the previous sections, I have shown that Otomi intransitive verbs display both an agent/patient and an active/stative semantic alignment. I have studied the latter in detail, and have compared the O-verbs in this active/stative semantic alignment with a special stative form which transitive verbs have nowadays. In this section, I make a proposal about how the active/stative alignment emerged in Otomi. The dynamics in question are represented in Figure 14.2.

The extreme left of the figure represents a stage at which intransitive verbs already displayed an agent/patient semantic alignment—very much like the one found today, but perhaps having more patient verbs than now. At this stage, a number of ‘agent’ verbs denoted property concepts. To make the next stage possible, we have to imagine a situation in which intransitive verbs could have stative forms just like the transitive verbs of today—something which is not possible now. In this hypothetical situation, a number of A-verbs that denoted property concepts had such stative forms. These stative forms functioned in most cases as nouns portraying the properties in question, just as stative forms of transitive verbs may now convert into action nouns. The residue of this stage is the handful of stative

![Figure 14.2. Development of the semantic alignment of intransitive verbs in (San Ildefonso) Otomi](image-url)
verbs which by conversion can be used as nouns, as illustrated in (20) above for San Ildefonso Otomi, repeated here as (38):

\[(38)\]
\[
\begin{align*}
n-t’áxi & \quad \text{‘white (as a colour)’} \\
n-théní & \quad \text{‘red’} \\
n-k’áxt’í & \quad \text{‘yellow’} \\
n-k’ángi & \quad \text{‘blue/green’}
\end{align*}
\]

At this point in this process, speakers began to use the stative forms of such A-verbs denoting property concepts as an alternative way to predicate the properties in question. Considering the A-verbs in (11) which still denote property concepts, it seems that stative forms were generally used when the property in question was an inherent property or was not typically ascribable to a human (or animate) being. This use must have become very common, to the extent that speakers reanalysed these stative forms as being a part of the inflectional paradigm of A-verbs denoting the same properties.

When the reanalysis was complete, a separate class of stative verbs had emerged in the system, displaying a unified inflectional paradigm, but with rather mixed origins. Stative verbs inherited from stative forms the fundamental feature of encoding their unique semantic argument as object. In this way, with the rise of stative verbs a new semantic alignment emerged in the language for intransitive verbs.

As already mentioned, stative verbs emerged as a rather mixed morphological class from the blending of different paradigms. On the one hand, stative verbs received from A-verbs the features they share now with them, as in (39a). On the other hand, the rest of their inflectional idiosyncrasies stem from their origin as stative forms. The list in (39) shows the inflectional spaces in the paradigm of stative verbs in San Ildefonso Otomi that stem from either former A-verbs or stative forms:

\[(39)\]
\[
\begin{align*}
a. \text{From A-verbs} & \\
& \begin{array}{ll}
\text{Imperfect (negative polarity)} & mi \\
\text{Present (negative polarity)} & Ø \\
\text{(Marginally) Irrealis} & da \\
\end{array} \\
b. \text{From stative forms in predication} & \\
& \begin{array}{ll}
\text{Perfect for present (positive polarity)} & xí=N- \\
\text{Present (with intensifiers)} & Ø=N-^{18} \\
\end{array} \\
c. \text{From stative forms as nouns in predication} & \\
& \begin{array}{ll}
\text{Imperfect (regardless of polarity)} & már=N- \\
\text{Irrealis (regardless of polarity)} & dar=N- \\
\end{array}
\]

\[^{18}\text{With the intensifiers ťá- ‘very’ and kā= ‘rather’, the verbal proclitic used is zero, but the stative prefix N- must still be used (e.g. Ø=ňá-š-pá=‘-a=š) (3.Officers=intensify-st-hot-2obj=2obj) ‘you’re very hot’.)} \]
All Otomi languages have stative verbs, but they differ greatly as to what the outcome of the diachronic blending is, suggesting that the development in Figure 14.2 started in the proto-language but was developed separately in each individual sub-group. In the cases where a given inflectional feature emerged from a stative form, we now see the use of the stative prefix $\textit{N}$- proper of stative forms.

Moreover, the two last properties inherited from stative forms in (39c) reveal that stative forms were used as nouns in predication in the imperfect and the irrealis, regardless of the polarity of the verb. In such cases, stative verbs inherited the special proclitics used in the nominal predication construction. The relevant examples are repeated in (40):

\[
\begin{array}{ll}
\text{(40) } & \text{STATIVE VERBS} & \text{NOMINAL PREDICATION} \\
\text{Imperfect} & a. \text{ mär=ð-hêts'i=gí } & b. \text{ mû=ð [r be]}_\text{NP} \\
(3.) & \text{IMP.PC=ST-be.high=1OBJ} & 3. \text{IMP.PC=SG thief} \\
& '\text{I was tall.}' & '\text{He was a thief.'} \\
\text{Irrealis} & c. \text{ dar=ð-hêts'i=gí } & d. \text{ da=ð [r md^h_jô]}_\text{NP} \\
(3.1) & \text{IRR.PC=ST-be.high=1OBJ} & 3. \text{IRR=SG priest} \\
& '\text{I'd be tall.'} & '\text{He'll become a priest.'} \\
\end{array}
\]

From being nouns in predication, stative verbs also inherited the lack of the free/bound form inflectional contrast seen in 14.5.3, which other $A$-verbs have. They extended this feature to their whole paradigm.

### 14.9 Conclusion

In this chapter, I have proposed for the first time that Otomi has the two types of semantic alignment pattern described so far in the literature: an agent/patient and an active/stative. For this reason, the language is typologically very interesting, as it is not common to find both alignments in the same system. The goal of the chapter was to advance a proposal for the emergence of the active/stative system, which remains a type of alignment which appears to be less frequent cross-linguistically and is less understood in current linguistic theory. I have proposed that the $O$-verbs in this alignment emerged as reanalysed $A$-verbs through a process of blending the paradigm of $A$-verbs with the one of a special inflectional form I called the stative form, which is now mainly used to denote resulting states of transitive verbs and to convert verbs into nouns. I have shown in detail how this process may have come about in the grammar of San Ildefonso Otomi, while also predicting further differences in other Otomi languages. An agent/patient alignment has been found in a handful of Otomanguean languages; the extent to which this could be regarded a feature of the macro-phylum remains to be seen. Outside Otomi, however, nothing is known about the existence of an active/stative
alignment in other Otopamean languages such as Mazahua and Pame, let alone in another Otomanguean or Mesoamerican languages in general. The present study is a cordial invitation to linguists working on these languages to contribute their findings to this fascinating area of the grammar.

Appendix 14.1 Stative verbs in San Ildefonso Otomi

<table>
<thead>
<tr>
<th>Dimension</th>
<th>Meaning</th>
<th>(Human)</th>
</tr>
</thead>
<tbody>
<tr>
<td>hêts’i</td>
<td>‘be high/tall’</td>
<td></td>
</tr>
<tr>
<td>xidi</td>
<td>‘be wide’</td>
<td></td>
</tr>
<tr>
<td>ndô’yô</td>
<td>‘be thin (for an animal)’</td>
<td></td>
</tr>
<tr>
<td>xîni</td>
<td>‘be thin (for a thing)’</td>
<td></td>
</tr>
<tr>
<td>ts’(t)i</td>
<td>‘be thin (for wood, thread)’</td>
<td></td>
</tr>
<tr>
<td>pêdi</td>
<td>‘be thick’</td>
<td></td>
</tr>
<tr>
<td>ndô</td>
<td>‘be thick (for wood, thread)’</td>
<td></td>
</tr>
<tr>
<td>mâ</td>
<td>‘be long’</td>
<td></td>
</tr>
<tr>
<td>nôho</td>
<td>‘be fat’</td>
<td></td>
</tr>
<tr>
<td>Physical property</td>
<td></td>
<td></td>
</tr>
<tr>
<td>‘ôdi</td>
<td>‘be rough’</td>
<td></td>
</tr>
<tr>
<td>hû</td>
<td>‘be heavy’</td>
<td></td>
</tr>
<tr>
<td>pà</td>
<td>‘be hot’</td>
<td></td>
</tr>
<tr>
<td>tsê</td>
<td>‘be cold’</td>
<td></td>
</tr>
<tr>
<td>k’a</td>
<td>‘be wet’</td>
<td></td>
</tr>
<tr>
<td>xâ</td>
<td>‘be humid’</td>
<td></td>
</tr>
<tr>
<td>‘ônsi</td>
<td>‘be dry’</td>
<td></td>
</tr>
<tr>
<td>tsê_tho</td>
<td>‘be insipid’</td>
<td></td>
</tr>
<tr>
<td>‘û</td>
<td>‘be sweet’</td>
<td></td>
</tr>
<tr>
<td>‘ûxi</td>
<td>‘be salty’</td>
<td></td>
</tr>
<tr>
<td>‘îxi</td>
<td>‘be sour’</td>
<td></td>
</tr>
<tr>
<td>jú</td>
<td>‘be bitter’</td>
<td></td>
</tr>
<tr>
<td>‘î</td>
<td>‘be spicy/hot’</td>
<td></td>
</tr>
<tr>
<td>kâhi</td>
<td>‘be tasteful/flavoury’</td>
<td></td>
</tr>
<tr>
<td>tûdi</td>
<td>‘be soft/weak’</td>
<td></td>
</tr>
<tr>
<td>tsêdi</td>
<td>‘be strong/firm’</td>
<td></td>
</tr>
<tr>
<td>mê</td>
<td>‘be thick (i.e. for a liquid)’</td>
<td></td>
</tr>
<tr>
<td>mê</td>
<td>‘be hard’</td>
<td></td>
</tr>
<tr>
<td>nôksi</td>
<td>‘be stale’</td>
<td></td>
</tr>
<tr>
<td>‘êtsxi</td>
<td>‘be dark’</td>
<td></td>
</tr>
<tr>
<td>ndû-pä</td>
<td>‘be lukewarm’</td>
<td></td>
</tr>
<tr>
<td>ndû-tsegî</td>
<td>‘be weak’</td>
<td></td>
</tr>
<tr>
<td>xô</td>
<td>‘be stinky’</td>
<td></td>
</tr>
<tr>
<td>‘û</td>
<td>‘be painful, be in pain’</td>
<td></td>
</tr>
<tr>
<td>hêri</td>
<td>‘be noisy/talkative’</td>
<td></td>
</tr>
<tr>
<td>ñô_ríhi</td>
<td>‘be fast, quick, light-footed’</td>
<td></td>
</tr>
<tr>
<td>Colour</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pôthi</td>
<td>‘be black’</td>
<td></td>
</tr>
<tr>
<td>po’t’si</td>
<td>‘be dark coloured/navy blue’</td>
<td></td>
</tr>
<tr>
<td>t’âxi</td>
<td>‘be white/silvery’</td>
<td></td>
</tr>
</tbody>
</table>
(Continued)

<table>
<thead>
<tr>
<th>Word</th>
<th>Meaning</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>théñi</td>
<td>'be red'</td>
<td></td>
</tr>
<tr>
<td>k'áxt'í</td>
<td>'be yellow'</td>
<td></td>
</tr>
<tr>
<td>k'ángi</td>
<td>'be blue/green'</td>
<td></td>
</tr>
<tr>
<td>hwé't's'i</td>
<td>'be golden'</td>
<td></td>
</tr>
<tr>
<td>káfë</td>
<td>'be brown'</td>
<td></td>
</tr>
<tr>
<td>éösá</td>
<td>'be pink'</td>
<td></td>
</tr>
<tr>
<td>gwéro</td>
<td>'be blond'</td>
<td>(Human)</td>
</tr>
<tr>
<td>ts'ò</td>
<td>'be ugly/evil'</td>
<td>(Human)</td>
</tr>
<tr>
<td>nzátho</td>
<td>'be beautiful'</td>
<td>(Human)</td>
</tr>
<tr>
<td>zhëhwa</td>
<td>'be silly (of an animal)'</td>
<td>(Human)</td>
</tr>
<tr>
<td>hó</td>
<td>'be good'</td>
<td>(Human)</td>
</tr>
<tr>
<td>hmòdi</td>
<td>'be expensive'</td>
<td></td>
</tr>
</tbody>
</table>
15

Voice and transitivity in Guaraní

MAURA VELÁZQUEZ-CASTILLO

15.1 Introduction

This chapter focuses on the relevance of canonical voice categories for the grammatical system of the South American language Guaraní, and claims that the closest equivalents to voice alternations in this language respond to spatial relations between events and event participants. The constructs at issue, previously labelled as ‘reflexive/passive’ and ‘inverse’, are sensitive to viewpoint and levels of proximity between the predicate and its arguments within inactive situations, and to deviations from the default outward directionality of active events.

Grammatical voice is generally understood as modelling the particular mapping patterns of semantic roles onto grammatical relations, primarily the traditional relations of subject and object. Voice categories (direct active, passive, antipassive, inverse, and middle) have been said to be based on a transitive event type, where alternations are viewed as subject–object prominence asymmetry patterns (e.g. Kemmer 1993 and references therein). However, Klimov (1979: 330) observes that semantically aligned systems are not based on the transitive event type as the primary building block, and consequently lack voice distinctions that are based on transitivity relations (Wichmann, to appear, reports cross-linguistic confirmation). Another challenge to the transitivity-based view of voice phenomena is found in Maldonado (1999), who challenges the assumption that the canonical transitive event is the sole conceptual starting point for the understanding of middle voice phenomena.

1 Guaraní—now generally considered a classic example of the semantic alignment type—is a language of the Tupí-Guaraní stock. The variety examined here is spoken in Paraguay. The data mainly comprise examples from written narrative texts; a small number of examples were constructed, using my native speaker intuition. Examples appear in the orthographic conventions adopted by the Paraguayan Ministry of Education. Most characters have the expected phonetic values, but note that $y$ is an unrounded back vowel [$\ddot{y}$], and $ch$ represents a [ʃ] sound. In most words the stress falls on the last syllable; the stress is orthographically marked only when the word deviates from this pattern. Nasality is marked only in stressed vowels, where this feature is contrastive. Nasality spreads both progressively and regressively from stressed nasal vowels, affecting unstressed vowels and neighbouring consonants. Consonant alternations that should be expected on account of nasal harmony are: $nd/\ddot{n}$, $mb/m$, $j/\ddot{n}$, $p/m$, and $p/mb$. 
Building partially on Klimov’s observations on the subject of grammatical relations, Velázquez-Castillo (2002) argues that the morphosyntax of Guaraní is not organized around the traditional grammatical relations of subject and object, but rather in terms of two spatially configured grammatical relations: (active) event sources and (inactive) event sites. The event source is a participant construed as the initiator or origin of a dynamic event, which may or may not have control over the situation. The event site is a participant construed as containing the situation denoted by the predicate.

I examine two sets of constructs that have been analysed as voice alternations. The so-called ‘inverse’ system (Payne 1994b) involves the 1 > 2 > 3 person hierarchy, and amounts to favouring the inactive participant over the active one for cross-referencing, when the former outranks the latter. My position here is that the inverse category in this language—if there is one—is rather covert, since it involves forms that are not exclusive to the inverse function. There is also an intriguing prefix r-, traditionally described as a ‘relational’ marker, which Payne analyses as an inverse marker. I argue, contra Payne, that r- cannot be considered a straightforward inverse marker, because the whole range of this marker’s use encompasses far more than transitivity-based inverse configurations. The second set of constructions involves the prefix je-, traditionally described as a passive and/or reflexive marker. I show, on the basis of a significant number of non-transitive situations that involve the use of je-, that the so-called ‘reflexive’ and ‘passive’ interpretations of this marker do not constitute its central semantic values, and that je-marked predicates can be analysed as embodying transitivity-independent notions that are more akin to current conceptions of the middle voice.

My general point is that analysing these constructions as inverse, passive, or reflexive constitutes an unwarranted imposition of a transitivity-based model on a language whose systemic foundation is not directly derivable from transitivity relations. I begin with an overview of the basic semantic alignment features of Guaraní.

15.2 Semantic alignment in Guaraní

Guaraní conforms to a number of Klimov’s predicted correlations: the alienable/inalienable opposition, the inclusive/exclusive distinction in 1st person pronouns, the prevalence of aspect over tense, and the tendency for incorporation. There is rich inflectional and derivational verb morphology with relatively specific semantics, while nouns are morphologically simpler. The semantic alignment of this language is most directly manifested in two sets of cross-referencing pronominal prefixes for the coding of active and inactive events (Table 15.1). Velázquez-Castillo (2002) analyses active prefixes as referring to event participants that are construed as sources of change, and as representing a distal relationship between the predicate and its argument. Inactive prefixes occur in predicative as well as
postpositional and nominal possessive constructions. They are analysed as referring to participants which are construed as the physical or psychological sites or settings for the unfolding of a particular event or relation. They represent a generally proximal relation between the verbal, nominal, or adpositional predicative element and its argument. As will become clear later, the involvement of inalienable possession and Speech Act Participants (SAPs) within inactive situations effects an additional level of proximity within the relation.

As Table 15.1 shows, inactive prefixes correspond almost exactly to free pronouns for the 1st and 2nd person, but this is not the case for the 3rd person. Most Tupí-Guaraní languages seem to lack independent 3rd person pronouns, but always have local pronouns. As a result, there has been some variation when it comes to the treatment of inactive 3rd person prefixes. Payne (1994b), following Jensen (1990), treats i- (also realized as, iij-/iñ-) as a 3rd person inactive marker. My own work has followed a similar line in the past, but I have come to agree with Seki (1990) that, on distributional grounds, i-( iij-/iñ-) is better analysed as an allomorph of the ‘relational’, rather than as a 3rd person inactive prefix; but while Seki treats h- and i- as relational 3rd person markers in Kamaiura, I treat relational morphs in general as unspecified for person, the main conditions for variation being morphological class or predicate-argument proximity level. What is important is that, in my analysis, there is no 3rd person inactive prefix for Guaraní, and i-( iij-/iñ-) is a relational inactive marker (rel) of unspecified person, which fills the slot for the missing inactive 3rd person prefix.

There are additional differences between 1st and 2nd person on the one hand and 3rd person on the other that point to a special status of SAPs. There is, for example, a marked preference for predicate-level (i.e. prefixal) realization of core

---

2 In Guaraní the original demonstrative ha’e functions synchronically as an independent 3rd person pronoun.

3 The ‘relational’ or ‘linker’ r- appears only with inactively marked units, and divides the stem inventory in two large classes: a vowel-initial class marked with r/-h-, and a consonant initial class that is zero or i-marked. An analysis of i- as an allomorph of the relational r- is justified for Kamaiurá on distributional grounds (Payne 1994b: 320), and the same is true for Guaraní.
1st or 2nd person inactive arguments, in contrast to 3rd person inactive arguments, which are frequently coded in a PP, outside the predicate. Also significant is the existence of two 1st person plural pronouns, set apart by whether or not they include the 2nd person, as well as an additional pair of special portmanteau prefixes, ro-/po-, which are used when the 1st person acts on the 2nd person; a different treatment from two-participant situations involving other grammatical persons.

15.3 The Guaraní ‘inverse’

Like most other voice phenomena, the inverse is understood in terms of the relationship between the canonical event participants of a transitive clause: the agent and the patient, and is defined by event flow directionality vis-à-vis these participants. The inverse is taken to signal a reversion from the ‘normal’ event flow with respect to a participant hierarchy, most commonly: 1 → 2 → 3. A direct event flow is from a SAP to a 3rd person (1 → 3), or (2 → 3), while the inverse marks a 3rd person acting on a SAP (3 → 1), or (3 → 2), or the 2nd person acting on the 1st person. Like the passive, the inverse is described as a device to signal increased patient prominence, but unlike the passive, it does not involve agent suppression or a grammatical promotion of the patient (Payne 1994b: 317). The centrality of transitivity is maintained throughout the literature, although one also finds passing notes to the effect that inverse morphs may appear in non-transitive contexts.

15.3.1 Inactively marked transitive clauses

The Guaraní ‘inverse’ is implicit in the cross-referencing system, and it amounts to categorizing as inactive bivalent clauses involving directionality towards an SAP. Assuming a personal hierarchy 1 → 2 → 3, an active bivalent predicate selects an inactive cross-referencing prefix over the active one when the former outranks the latter (Payne 1994b). Payne explains the Guaraní ‘inverse’ in terms of transitivity, although she also opens the possibility of expanding her analysis beyond strictly transitive contexts.

Below is the type of transitive data used in support of the existence of an inverse system. Sentence (1) is direct (1 → 3). The prefixes on the predicate register the active and the inactive participants, in that order. The same pattern is observed when both participants belong to the 3rd person. Sentence (2)
represents a canonical inverse situation \(3 \rightarrow 1\): two participants involved in an event that flows in the inverse order from that of the personal hierarchy. The active participant (\(Toma\)) is not cross-referenced on the predicate, which takes a single inactive 1st person prefix. This is the case whenever, the configuration is of the form \(3 \rightarrow \text{SAP}\). Similarly, when \(2 \rightarrow 1\), only the 1st person gets registered on the predicate: \(nde \text{che-nupã} [\text{you} \text{1in-hit}] \text{‘you hit me’}\).

\[
\begin{align*}
(1) & \quad \text{Che a-i-nupã Jose-(pe)} \\
& \quad \text{I act-rel-hit Jose-loc} \\
& \quad \text{‘I hit José.’}
(2) & \quad \text{Toma che-nupã.} \\
& \quad \text{Tomas 1in-hit} \\
& \quad \text{‘Tomás hits me.’}
\end{align*}
\]

The inactive cross-referencing in \(2\) has to do with viewpoint, in the sense of DeLancey (1981). The speaker, who is also the event target, describes the event from his/her own location in space. According to DeLancey (1981: 638), the prototypical viewpoint, i.e. the spatial location from where a given scene is viewed, is the site of the speech act: ‘when an SAP is also a participant in the event reported, then the most natural viewpoint for the sentence is with the SAP.’ The viewpoint concept lends substance to the otherwise unclear common observation that the inverse treats SAPs as ‘proximal’, and the 3rd person as ‘distal’ (e.g. Dahlstrom 1986: 71–2, Gildea 1994: 203–5).

Following DeLancey’s line of thought, my suggestion for the case of Guarani is that a situation that targets an SAP effects a switch in viewpoint from the event source to the event target. The question that arises is how this switch in perspective motivates the inactive marking of these otherwise active predicates. First, if we consider that an SAP is a participant close to EGO, the speech act defines a personal sphere identified with the self, much like body-part relations do. Second, with the viewpoint switch, the situation is viewed ‘from within’ this personal sphere, rather than as projected forward from the source. This configuration of inclusion in the sphere of EGO conforms to that of inactively marked ‘site-participants’ in monovalent predicates. In this regard, the treatment of SAPs in inverse event configuration resembles that of body-part relations, also morphologically classified as inactive. The notion of maximal proximity, shared by these two relation types, will be relevant to the proposed function of the prefix \(r-\), to which I now turn.

15.3.2 The \(r\)-formative: an inverse morph?

The prefix \(r-\), common to several Tupi-Guarani languages, has eluded a satisfactory analysis to date. Traditionally it is known as the ‘relational’ or ‘linking’ morph. For some analysts it is simply a morphological class marker, or a ‘morphological transition’ prefix, standing between other prefixes and vowel initial-stems (Dooley 1998: xi–xii). Velázquez-Castillo (1996) suggests that \(r\)-marks inherently relational stems, while Payne (1994b) argues for an inverse analysis. I show here that the overall distribution of \(r\)-militates against a strictly inverse analysis, and that a transitivity-based analysis obscures rather than clarifies its function within the
overall semantic alignment of this language. I claim that the traditional analysis of \( r \)- as a relational marker is essentially correct, and that the prefix marks maximally proximal inactive relations.

Example (3) replicates the inactive ‘inverse’ pattern described above: only an inactive 1st person prefix appears on the predicate, even though the active 3rd person \( Toma \) is overtly expressed. In this case, there is also an \( r \)- that appears after the inactive 1st person prefix. It is on the basis of examples such as (3), where \( r \)- would be marking a \( (3 \to 1) \) configuration, that Payne articulates the main argument for her inverse analysis: ‘it is clear that when the “linking/relation” prefix occurs, it occurs in addition to at least 1st and 2nd person Set 2 [inactive] prefixes’ (Payne 1994b: 321). The similarity of this pattern to the Algonquian inverse leads Payne to conclude that \( r \)- is a ‘fairly straightforward case of an inverse marker.’

(3) \( Toma \) che-\( r \)-eka.

\( Toma \) in-\( n \)-REL-SEEK

‘Tomás seeks me.’

In (4), where -eka ‘to seek’ combines with a 3rd person active prefix, the \( h \)- prefix appears instead of \( r \)-. For this reason, stems like -eka are commonly known as ‘alternating stems,’ and \( h \)- as an inactive 3rd person prefix. On the surface, then, the occurrence of \( r \)- seems indeed sensitive to the presence of inactive 1st and 2nd person prefixes. There is, however, a difference between (3) and (4) which will turn out to be important: in (4) the patient Jose-pe is expressed outside the predicate, while in (3) the inactive argument is coded on the predicate by a pronominal prefix. This, rather than the fact that the patient belongs to the 3rd person, will turn out to be more important for understanding the absence of \( r \)- in sentences like (4).

(4) \( Toma \) o-h-eka Jose-(pe)

\( Toma \) 3ACT-REL-SEEK Jose-LOC

‘Tomás seeks José.’

Beyond sentences like (3), the range of uses of \( r \)- does not correspond to an inverse function. An inverse situation rests on two crucial elements: (1) the presence of two unequally ranked event participants, and (2) event directionality towards SAPs, i.e. a 1st or 2nd person transitive patient. Neither of these two elements is consistently present in the uses of the Guaraní formative -\( r \), and to understand why this is the case, it is important to consider that \( r \)- appears only with ‘alternating’ stems, and that these are not limited to transitive predicates, as shown in Table 15.2. In fact, the largest group of \( r \)-carrying stems consists of nouns of the inalienable class: kinship and body-part terms. Thus, even under the (reasonable) assumption that inalienable terms express a special type of predication, they would be monovalent and, therefore, not transitive predicates. Another group of \( r \)-carrying stems are monovalent inactive predicates designating physical or emotional conditions. Finally, the \( r \)-bearing bivalent predicates are not necessarily defined by high transitivity
Table 15.2. Different types of r- carrying stems

<table>
<thead>
<tr>
<th>Stem type</th>
<th>No. of items</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nouns (inalienable class)</td>
<td>60</td>
<td>56</td>
</tr>
<tr>
<td>Monovalent predicates (inactive class)</td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td>Bivalent predicates</td>
<td>17</td>
<td>16</td>
</tr>
<tr>
<td>Total</td>
<td>106</td>
<td>100</td>
</tr>
</tbody>
</table>

Given this distribution pattern, the first difficulty for an inversion analysis is that r- does not consistently appear in canonical inverse situations. Sentence (2), for instance, has a canonical inverse configuration, but features no r-. The absence of r- in countless inversely configured transitive sentences would be inexplicable under this analysis. A second, related difficulty is the large number of constructions where r- could not possibly involve a topicality ranking between two arguments. Example (5) features r- in combination with a monovalent inactive predicate, (6) shows r- in a nominal construction, and (7) with the postposition -ehe/-ese ‘about.’ In (5), the single argument of the predicate, -etia’e ‘cheerful,’ is marked on the predicate with an inactive 1st person prefix, and also stated separately for emphasis. Note that the predicate -asy takes the prefix h- instead of r-, and that it involves a 3rd person argument (Maria), stated outside the predicate. Whenever h- occurs on the predicate, a non-emphatic NP or PP is present outside the predicate. The argument, then, is not expressed by a personal prefix attached to the predicate, and this is a key difference between h-asy and r-etia’e.

5)  Maria h-asy ha che katu che-r-eti’a’e a-pu’â-vo.
   Maria rel-ill and I emph in-rel- cheerful iact-get=up-ser
   ‘Maria is ill, and I (on the other hand) am cheerful as I got up.’

6)  Ahâniri, che-r-a’y
   No in-rel-son
   ‘No, my son.’

7)  O-ñ-akã-rapu’ã o-ma’è h-ese...
    3act-mid-head-lift 3act-look rel-about
    o-jere o-ma’è ta’yra r-ehe.
    3act-turn 3act-look son rel-about
    ‘He raised his head and looked at him... turned and looked at his son.’
A third problem is that the stipulated condition of directionality towards an SAP for the use of r- does not hold. Example (7) above features both the h- and r- forms of -ese/-ehe, and both instances involve a 3rd person target. In nominal possessive constructions, such as (8), r- combines regularly with 3rd person possessors. In (9), r- appears in a noun–verb combination, featuring a bivalent predicate which does not involve directionality toward an SAP. The noun mba'e appears immediately before the bivalent predicate -eka 'to seek.' Clearly, the presence of r- in these examples is not motivated by grammatical person.

(8) Mokõi Lope r-a'y-re ndaje o-u raka'e…
Two López REL-son-PST EV 3ACT-come DISTANT-PST
‘Two of Lopez’ sons had come, I am told.’

(9) o-i-kuua-ma Karai Poli mba'e r-e-ka-pa o-u hina.
3ACT-REL-know-PFV Mr. Poli thing REL-search-INT 3ACT-come PROG
‘Mr Poli already knew in search of (what) thing he was coming.’

Since r- appears in contexts where directionality as a function of a topicality competition between two participants is completely irrelevant, and since ‘the notion of the inverse is crucially dependent on its occurrence in transitive clauses’ (Payne 1994b: 335), an inverse analysis of r- needs to explain its occurrence in intransitive structures, and those in which r- appears with a 3rd person argument. Payne stipulates that possessed Ns and adpositions are predicate-like elements that carry an argument ‘in a state of being in a relationship’ to the possessed N or the adposition, which endows it with a (P)atient-like quality (Payne 1994b: 336). By this account, the occurrence of r- is motivated by ‘P-oriented’ constructions in general, of which inverse situations are a special type. This would explain why r- appears only with inactive prefixes. Additionally, she suggests that r-, initially a nominal marker, is in the process of being reinterpreted as an inverse marker (Payne 1994b: 313, 336).

Payne is certainly one of the few analysts who attempts to motivate the transitive and non-transitive occurrences of putative inverse markers. The problem is that the concept of P, as originally conceived, is tied to the syntactic configuration of bivalent predicates, and any extension of this concept to include an undefined class of non-transitive ‘patients’ renders the term meaningless. Such an extension would necessitate careful elaboration, based on the particular relations that underlie the semantic alignment of this language.

The putative diachronic explanation is also problematic. If r- were in the process of being reinterpreted from a nominal marker to a specific inverse morph, occurrences of r- in non-inverse configurations must be treated as remnants of earlier historical stages. However, the putative inverse function of the prefix is, numerically at least, vastly exceeded by other, ‘remnant’ functions.

A strictly inverse analysis of the prefix depends entirely on transitivity relations, unwarrantedly privileging one context of occurrence over the others. An analysis
which rests primarily on transitive data, with minimal attention to non-transitive contexts, cannot accurately capture the shared features of the transitive and the intransitive uses, and obscures any understanding of the semantic core of the prefix. But what is the semantic/functional core of this prefix, and how does this core relate to all the contexts in which it occurs?

From what we have observed, the traditional characterization of r- as a ‘relational’ prefix does not seem at all far-fetched. The prefix combines with units that belong to lexico-semantic categories involving a proximal relation of some kind. Taking Payne’s suggestion that possessed Ns (especially inalienable Ns) and adpositions are predicate-like units that carry an argument, I posit that r- marks proximal argument-predicate connections in inactive situations. A proximal relation may involve objective physical proximity, as in part–whole relations, or a construed proximity between the predicated event or relation, and EGO. A proximal construal involving an event/relation and EGO requires a point of view which is situated in the deictic centre: r-marked bivalent predicates locate their point of view within the speech act, and have a 1st or a 2nd person argument expressed on the predicate.

All the patterns of occurrence of r-, including those which fall outside the inverse function, can be motivated under a relational analysis. In my analysis, both r- and h- are person-neutral forms, indicating degree of argument-predicate connection, r- being reserved for situations of close connection and h- for situations without such close connection. That r- does not require association with an SAP is not a problem, because closeness to EGO is only one type of possible proximal relation. The personal ranking observed in transitive clauses is the manifestation of a deictic point of view, a broader phenomenon which, although relevant to the meaning of r-, is not directly involved in it. This is evident when we consider that the h-/r- alternation is not motivated by grammatical person; instead, it correlates neatly with the position of the argument vis-à-vis the stem. The r- form appears whenever the argument occurs immediately before the stem, regardless of grammatical person. The h- version, on the other hand, appears when the referent of the argument is available in the local discourse context, even in the same sentence as the predicate, but not stated immediately before the predicate. The reason why h-, not r-, appears frequently in association with the 3rd person is that a 3rd person possessor or argument is not necessarily expressed next to the predicate, as a 1st or a 2nd person possessor or argument is.

The fact that the use of r- does not require a transitive configuration and a particular event flow directionality with respect to two participants does not constitute a problem either, because proximal relations are not inherently tied to transitivity. We have seen that the set of r-carrying (alternating) stems cuts

5 In a construction such as: ro-h-ova-mokã-ta [1–2 ac-face-dry-fut] ‘I will dry your face’, the h- cannot possibly be a 3rd person prefix, because there is no 3rd person involved. This is congruent with analysing h- as an allomorph of the relational. H- separates the personal prefix from the predicate when there is a less-proximal relation of the predicated situation to EGO.
across lexical categories and includes one- and two-place predicates, as well as nouns.

In sum, I have demonstrated that the relation designated by r- includes inverse situations, but it is not confined to these. I propose that the use of r- relates to a distinction between levels of proximity within inactive situations, and that this is the operative factor in the r-/h- alternation. I now turn to the second set of Guaraní constructions commonly analysed in terms of voice alternations, namely the so-called 'reflexive/passive' je- constructions.

15.4 The prefix je-: a reflexive/passive marker?

The prefix je-/ñe- has been described as a reflexive and/or passive marker—both functions crucially hinging upon transitive events—even though je- normally combines with monovalent and bivalent predicates alike. In (10) je- appears with a bivalent predicate and is the translational equivalent of a reflexive or passive.

(10) Toma o-je-japi.
    Tomás 3act-je-shoot

   ‘Tomas was shot or shot himself.’

The predicate carries active morphology (o-), which makes sense on account of lexical aspect, given that the process is inherently dynamic under both readings. But normally the source of a two-participant event is marked as active, and the target as inactive. In the reflexive reading, the active marking is justified, because the single event participant counts as the event source. In the passive reading, however, the participant indexed with o- is not interpreted as the event source, but as the target. The participant roles in these two readings are different, and correspond to the two roles normally present in a two-participant predicate. The fact that the same (active) marking is used for both these participant role types constitutes a problem for a transitivity-based analysis. This is compounded by the fact that je- itself is compatible with either active or inactive marking. Although inactive predicates normally do not combine with je-, as shown in (11) (where the prefix is represented by its allomorph), the prefix can combine with the inactive relational marker i- when it attaches to postpositions. The postpositions -rehe ‘about/at/against’, -pe ‘in, into’ and -hegui ‘from’, combined with je-, yield: i-je-he ‘about, at himself’, i-ju-pe ‘to himself’, and i-je-hegui ‘from himself’.

(11) *Toma i-ñe-kane’dô.
    Toma  rel-je-tired

   ‘Tomás himself is tired.’

In a split-self reflexive sentence, such as (12), the active o- on the predicate and the inactive i- in the postpositional phrase distinguish participant roles in a typical two-participant event. But such reflexives are highly marked and not available in
most uses of je-, as we will see; (12) could be a response to the question 'Who did it?' following a potentially ambiguous utterance like (10).

(12) Ha‘e o-japo i-je-he.
He 3ACT-do REL-je-TO
‘He did it to himself.’

The active marking of the featured participant in the je-construction is reminiscent of the behavior of absolute patients, which are patients of monovalent verbs, like kái 'burn', designating a spontaneous, self-contained event, viewed independently from an external source (Velázquez-Castillo 2002). Like absolute patients, je-marked predicates feature only one participant—an oblique agent of the type ‘by X’ is not possible. But unlike absolute patients, the participant marked on a je-bearing predicate is not always interpreted as the event-internal source.

The problematic nature of the transitive analysis of je-as either a passive or reflexive marker will become increasingly clear as we examine cases of je-marked verbs where transitivity relations are not centrally involved, and where neither reflexive nor passive readings are obtainable. My proposal is that je-signals the cancellation of the default outward projection (i.e. outward directionality) inherent in an active predicate. In combination with a postposition, je-cancels an expected outside source. This schema is akin to some current, transitivity-independent conceptions of the middle voice. Maldonado (to appear) defines the middle voice as encompassing those actions, events, or states pertaining to the subject’s own sphere, a voice category that contrasts with the active-direct voice in that the process remains in the subject instead of being projected to another participant. He questions the prevalent view that the main conceptual (and historical) source of the middle is the reflexive and thus, ultimately, a transitive event type (see Kemmer 1993 and references therein). He argues that events focused on the subject’s sphere are basic constructs in their own right and need not develop from the reflexive. Understood in this way, the notion of a middle event ties neatly into the distinction between centripetal (an action that is circumscribed to a participant) and centrifugal (an action directed outwards from a participant) event configurations suggested by Klimov (1979) for semantically aligned systems. Given this similarity, and the familiarity of the middle voice as a grammatical category, I will use the term ‘middle voice’ (mid) to designate all instances of the je-prefix.

15.4.1 Bivalent predicates interpreted as single-participant events

This section examines je-marked bivalent predicates expressing body-move and body-care events. The participant is invariably human or at least animate, and is interpreted as both the source and target of the event; thus, a passive interpretation is ruled out. In a transitivity-based model, the availability of a split-self reflexive reading would indicate the dual-participant nature of the event, but I will show that such true reflexive reading is not available either. In (13), the motion
verb *hekýi* ‘pull’ depicts a participant, Nika, who pulls himself away from a blow aimed at him from another participant. Nika’s body, as an undifferentiated whole, performs a self-propelled, non-translational motion. In (14) the participant’s body rotates, self-impelled, on its own axis. Note, that although there is a causative *mbo*-, no split-self reading is available: ‘*o-mbo-jere i-jehe*.

(13)  
...*o-ha’ã Nika-pe ha péva o-je-hekýi chugui.*  
3ACT-aim Nika-in and this 3ACT-MID-pull from him  
‘...he aimed at Nika and he (Nika) pulled himself from him (dodged him).’

(14)  
...*o-poí-je in-iru-gui ha o-ñe-mbo-jere he’i i-chupe…*  
3ACT-let-go-ev 3IN-partner-from and 3ACT-MID-Caus-turn  
tell REL-him  
‘He let go of his partner and turned around and told him...’

Body-moves often feature a participant who performs an involuntary or partly conscious movement involving only a part of his/her body (usually coded by means of noun incorporation), either to change posture or to signal a particular mental state. In (15), Losánto looks up and, hand on waist, waits defiantly in response to a threat from a drunken man. The moves, confined within the bounds of the participant’s body, express his mental/physical state, and are clearly not projected towards a second participant. But note that a split-self reflexive reading is not available: ‘*o-joko i-ku’a i-jehe*.

(15)  
*O-je-sa-upí-re jeko o-je-ku’a-ko Losánto o-ha’drô.*  
3ACT-MID-eye-lift-when ev 3ACT-MID-waist-hold Losánto 3ACT-wait  
‘As he looked up, hands on waist, Losanto set out to wait, [so I’m told].’

One often finds *je-* combined with the causative *mbo*-, which normally derives a bivalent predicate out of a monovalent one. In (16), *je-* combines with a *mbo*-derived bivalent predicate and an incorporated body-part term, to designate a self-care event. This event type does not have the traits of a regular transitive verb, since it invariably involves a single participant who acts deliberately on its body, with no sense of an external Causer. Expectedly, no split-self reading is available.

(16)  
...*o-ñe-hova-mo-kā mante i-pañuélo-pe.*  
3ACT-MID-face-Caus-dry frequently REL-handkerchief-in  
‘He frequently dried his face with his handkerchief.’

Another common use of *mbo*-marked predicates, shown in (17), is to express the idea ‘pretend to be or do X’, as in *ñe-mbo-puka* [MID-Caus-laugh] ‘to pretend to laugh,’ and *ñe-my-asê* [MID-Caus-cry] ‘to pretend to cry,’ i.e. a middle rather than a reflexive meaning.

(17)  
Re-ñe-mbo-tavy-sé niko nde.  
2ACT-MID-Caus-ignorant-des EMPH you
‘It is that you want to pretend to be ignorant (or silly).’
*‘It is that you want to be made (or kept) ignorant.’

One probable reason why je- has long been taken for a reflexive or passive marker is that many of its uses involve a bivalent predicate. But the situations described by these verbs are often inherently single-participant events, with no sense of an external agent, and limited or no control on the part of the participant. The common occurrence of an incorporated body-part term may have contributed to the confusion of je- with a reflexive or passive marker, since the body-term can be taken to stand for a second participant. But the body-part in these constructions is construed as organically connected to the possessor and virtually identified with it (Velázquez-Castillo 1996: 161–3).

15.4.2 Collective human activity, inanimate event target/site

Sometimes, the je- marked predicate is interpreted to designate a generalized or collective human action, construed around an inanimate entity as the target/site of the event. Since the source is diffused into a generic entity and the putative patient is inanimate, a ‘reflexive’ reading is not available. Example (18) shows a je-marked bivalent predicate with a collective human activity reading. The reciprocal ño- and the totalitative -pa in the postpositional phrase are overt expressions of the collective nature of the event source.

(18) Tembi-puru o-je-puru avei o-ño-ndive-pa
      pat-use 3act-mid-use also 3act-mid-with-tot

‘Tools/utensils were also used among all.’

Under a transitivity model, an agent-defocusing passive with a valence reduction operation would have to be proposed for these verbs; but such an analysis would be unjustified because, as shown by the ‘reflexive’ uses, je- does not always eliminate the event source. Unlike an absolute patient, construed as a spontaneous event source, the participant of a ‘reflexive’ (i.e. middle) event is construed as delimiting the sphere of realization of an active event, which may or may not include the source. Moreover, the same collective human activity reading is also available in the case of inherently agentless constructions, as will be explained.

There is another type of collective human activity construction where je- appears with a monovalent predicate. These are single-participant constructions with no potential underlying asymmetry between a putative subject and object.

6 The reciprocal jo-/ño-, which clearly shares phonic substance with je-, occurs with active predicates and postpositions but, unlike je-, is obligatorily marked as active. The active marking makes sense when one considers that jo- refers to a dynamic unit of interacting participants. Because jo- denotes a situation where two active participants are involved in a bidirectional action, i.e. without the unidirectionality and internal asymmetry observed in transitive situations, the two participants are construed as a unit.
In these cases, the event is construed around an explicitly stated site or location. In (19), the main focus is on the actual occurrence of the event within the specified site, Kañada. Since the site cannot be interpreted to be a target, no patient-focusing passive reading is possible. An agent-suppressing rather than patient-promoting passive analysis would not be possible either, because inherently agentless predicates, like mano 'to die' and a 'to fall', can easily combine with je- to obtain the same reading, as shown in (20).

(19) Kañada-pe o-je-jerok̖ y kāda pyhare.
Kañada-in 3ACT-MID-dance every night
‘Kañada is danced in every night’ or ‘In Kañada there is dancing every night.’

(20) Heta gueteri o-ñe-mano-vé-ta nd-o-pá-i-rō la
much even 3ACT-MID-die-more-FUT NEG-3ACT-end-NEG-COND DEF
ñorairō.
fight
‘Even more dying will occur if the fight does not end.’

We have seen, then, that the notion of transitivity and its inherent participant asymmetry are analytically irrelevant in these cases. Section 15.4.3 examines ‘action nominals’, which code active events construed as self-contained spaces with no participant involvement.

15.4.3 Action nominal constructions: self-contained events
Sometimes a je-marked verb is used without a cross-referencing prefix, in which case it functions as a deverbal ‘action nominal’ (Comrie 1976: 178), referring to a type of human activity without association with any specific participant. Verb stems can be either bivalent or monovalent, but in either case, event participants are undistinguishable from the event itself.

Compare the predicative (21) and nominal use (22) of the bivalent verb juka 'kill'. In (21), the je-marked verb carries the customary prefix o-, cross-referencing the patient, ovecha 'sheep'. In (22), je-juka is understood as an argument of the predicate o-i-ko, where it is cross-referenced by o-. This is one indication that je-juka functions as a noun. Interestingly, neither the agent nor the patient appears in the nominal version (22). Although the process designated by action nominals is invariably understood as a type of human action, the source cannot be overtly expressed. Significantly, the nominal version in (22) carries no cross-referencing prefix. The patient can be specified as a type, but only preverbally, as in kure je-juka 'pig butchering' (22). This fixed preverbal position suggests that kure does not count as a true participant (Velázquez-Castillo 1996: 123–7).

(21) O-je-juka pete̱i ovecha ne-santo-ára-rehe.
3ACT-MID-kill one sheep 2IN-saint-day-for
‘A sheep was butchered for your birthday.’
(22) O-i-ko petetje-juka ange pyhure.
   3ACT-rel-exist one MID-kill last-night
   'A killing took place last night.'

Example (23) features a semantically intransitive action nominal involving a bivalent verb stem. The sentence designates an inner state construed as an internal process. The incorporated body-part noun has no specified possessor and cannot be considered a true event participant (Velázquez-Castillo 1996: 144–9).

(23) ... ha'e-kuéra nd-o-i-kuuā-i je-py'a-py.
   he-pl NEG-3ACT-rel-know-NEG MID-chest-press
   ‘... they did not know of worries, (lit.: ‘heart-pressing (things)’)’

In (24) the action nominal derives from a monovalent predicate. The formal and semantic intransitivity of the construction is therefore unquestionable. The preceding sentence is a general statement about the land and the sky, so no potential event source or site is available.7

(24) Yvy ha yva nda-i-jāra-i. Nda-i-pōri je-i-ko vai-rã.
   earth and sky NEG-REL-owner-NEG NEG-REL-exist MID-REL-live bad-FUT
   ‘Earth and sky do not have an owner. There is no (reason) for quarrelling.’

Example (25) also features a monovalent verb stem. Interestingly, (25) contains the suffix -ha, indicating that the deverbal unit construes the event as defining a urination place. The suffix codes a semantic element that ties into the general localist nature of my analysis. Significantly, je-kuaru-ha takes the locative -pe, confirming the spatial conception of the event, and the construction’s nominal nature.

(25) O-tı-gui h-apicha-kuéra-gui
   3ACT-shame-from REL-companion-PL-from
   o-ho o-i-ke je-kuaru-há-pe
   3ACT-go 3ACT-3IN-enter MID-urinate-place-in
   ‘Ashamed because of his companions, he went (and) entered the place where one urinates.’

To summarize, I have shown that (i) there are a number of monovalent je-marked predicates that cannot be analysed as having a transitive base, and (ii) a significant number of bivalent je-marked predicates designate intrinsically intransitive situations, such as body-care and body-move events. I have proposed a unified analysis where je- marks the cancellation of the default outwardly projection of active events, and of an outside source in the relation expressed by postpositions. The event or relation is viewed as circumscribed by a sphere defined by the most

7 Note the presence of the future marker -rã, which, along with postpositions, is a normal feature of non-derived nouns in Guaraní.
proximate entity available in the scene. In the case of monovalent predicates, this can be an explicitly designated place, which defines, literally, the event site. In the case of bivalent predicates involving situations such as body-move and body-care, this locus or site is defined by the main participant’s body. In other bivalent predicates, the patient acts as the event site, and these are the main culprits for the confusion of je- with a reflexive and/or passive marker. My analysis is essentially independent of the concept of transitivity, and is compatible with current characterizations of the middle voice. It provides for the possibility that this construction involve canonical transitive situations, but does not rely on the notion of transitivity.

15.5 Concluding remarks

My overall proposal is that, in place of a traditional grammatical voice, Guaraní has something more akin to a direction system based on the source vs. site participant types underlying the inactive/active distinction. Inactive events are conceived as essentially non-directional and contained within an event site. Active events, on the other hand, have a default outward direction vis-à-vis an event source. A je-marked predicate indicates an unexpected non-directional configuration in an otherwise active situation: it can indicate directionality towards the event source, or lack of an external source in a particular relation that normally has one.

The so-called Guaraní inverse is a deictic category (in the sense of DeLancey 1981) and categorizes as inactive a dynamic event that is configured as follows: (i) the event is directed towards EGO or a participant close to EGO, and (ii) the viewpoint is located in the sphere of EGO. This situation of containment approximates the in-site configuration of an inactive situation. Within inactive situations, an r-marked stem designates maximal proximity between a predicate and a site participant, differentiating it from regular inactive configurations. Maximal proximity is determined by physical proximity, as in part–whole relations, or by conceptual proximity to EGO, when the viewpoint is located in the sphere of EGO.

By showing that the closest equivalents to voice phenomena in Guaraní cut across the transitive divide, this chapter demonstrates that the existing bent in the literature in favour of transitive events as the sole conceptual departure point for understanding voice-like phenomena may not be applicable to semantically aligned languages. The alternative analyses offered do not hinge upon the concept of transitivity, and are based on the semantic relations that underlie the particular system of this language. This analytical approach is relevant for voice phenomena in general, but it is particularly relevant for their equivalent in semantically aligned languages, where the centrality of transitivity relations is questionable.
Agreement in two Arawak languages: Baure and Kurripako

SWINTHA DANIELSEN AND TANIA GRANADILLO

16.1 Introduction

The majority of Arawak languages have been argued to show a semantic alignment (cf. Aikhenvald 1999: 87ff.). Core arguments are cross-referenced on the verb by means of prefixes or suffixes (alternatively, pro- or enclitics). Aikhenvald (1999: 86), describing cross-Arawak patterns, states: 'Verbs typically divide into three classes: transitive (with two core arguments, A and O), active intransitive (with one core argument, S₁) and stative intransitive (with one core argument, S₀).' These verb classes differ in their choice of agreement for subject. For transitive and the so-called 'active' intransitive verbs the subject, A and S₁, is marked by an agreement prefix. The P of transitive verbs is marked by an agreement suffix. The split in the intransitive verbs is apparent from the fact that the subjects of some verbs, the so-called 'stative' intransitive verbs, are marked by a personal suffix identical to that which marks the Ps of transitive verbs (therefore S_P).

Fieldwork on Baure was financed by Radboud University Nijmegen, with technical support and equipment from MPI Nijmegen. Many of the conclusions are owed to comments by Pieter Muysken, Mily Crevels, Hein van der Voort, Katharina Haude, Rik van Gijn, Andreij Malchukov, Katja Hannß, Françoise Rose, and many more. The Baure data have been gathered from the following speakers: Asunta Durán, Dolores Chimanacay, Marcial Chonono, Juana Pinaicobo, Melquiades Durán, Justina Cajareico, Hercilia Chipeno, Ignacio Martínez, Lucio Oní, Guillermina Pinaicobo, Rosalia Pinaicobo, Esteban Chipeno, Julian Imanareico, Eustaquia Chuirpuay, and † Mercedes Peña, whose cooperation is here acknowledged. Field trips to the Kurripako were partially financed by the Tinker Foundation, University of Arizona Social and Behavioral Sciences Research Institute, University of Arizona Joint Anthropology and Linguistics Research Funds, and National Science Foundation Grant BCS-0318762.

1 Alternatively these morphemes are analysed as clitics in some Arawak languages, as in the case of Baure in this chapter.

2 Different terminology has been suggested instead of 'active' vs. 'stative'. Wise (1986: 57) divides verbs in Arawak languages into 'active' and 'absolutive' verbs. Facundes (2000: 274) treats verbs with S₁ marking in Apurinã (North Arawak) only as a subgroup of 'descriptive' verbs. Like Merlan (1985: 325), Facundes uses the descriptive labels 'subjective' vs. 'objective' marking, which are the most neutral.

3 The same personal cross-reference prefixes that mark the subject of transitive and active intransitive verbs (A and S₁) generally also mark the possessor of nouns.
Even closely related languages differ as to which verbs fall in the classes of ‘active’ and ‘stative’ intransitives (Aikhenvald 1999: 86). The choice of the labels ‘active’ vs. ‘stative’ relates to the fact that the split generally involves the parameter ±event (see Table 16.4): stative verbs are non-eventive, such as ‘to be cold’, while ‘active’ verbs might include verbs such as ‘sleep’, ‘be lost’, or ‘sit’. This shows the split in the coding choice for stative intransitive verbs (see also Facundes 2000: 273ff.) and that the split in argument marking is not always semantic, but in some cases has lexicalized into conjugation classes.

Some Arawak languages have a partly fluid semantic alignment system (e.g. Baniwa in Aikhenvald 2001: 175); such a system is characterized by different marking choices for one and the same intransitive verb depending on the semantics of the state of affairs sketched. The split may involve a considerable number or only some of the intransitive verbs. The split systems have only been described in detail for very few individual Arawak languages (mainly for Warekena in Aikhenvald 1998 and Apurinã in Facundes 2000).

In this chapter the agreement systems of two Arawak languages are investigated: Kurripako from the North Arawak group, and Baure from the South Arawak group, chosen for reasons of maximal spread in the family. After having given a general introduction to argument marking systems in Arawak languages, we describe the specific systems of Kurripako (section 16.2) and Baure (section 16.3). Section 16.4 provides additional data on argument marking on predicates derived by attributive and privative prefixes. Section 16.5 presents a comparison of the two languages. In section 16.6 the two languages are discussed from the point of view of semantic alignment. Section 16.7 presents data on fluid semantic alignment in the languages compared. Finally, the results of the comparison are summed up in the conclusion.

16.2 Argument marking on Kurripako predicates

Kurripako is a North Arawak language, spoken by around 10,000 people in northwest Amazonia around the borders of Venezuela, Columbia, and Brazil. The data presented here were collected by Granadillo during short field trips in 2000 and 2001, and through all of 2004. The Kurripakos in Venezuela are mostly bilingual, although there are a few monolingual speakers of Kurripako and a few monolingual Spanish speakers. Kurripako is a polysynthetic, head-marking language with predominantly VOS constituent order. The language has a gender and an extensive noun class system and few grammatical cases, dative -sri, and various locatives. Subjects and objects are indexed by means of affixes, to be described below.

As already mentioned in the introduction, Split-S systems in argument marking on verbs have been noted for different Arawak languages. Kurripako is a prototypical example of a language with a split in the marking of S (i.e. a semantic alignment) based on eventivity.
16.2.1 Subjective marking (Sa); active verbs in Kurripako

Verbs in Kurripako are either transitive or intransitive; there are additionally a few undervived ditransitives and some derived benefactives and causatives. Intransitive verbs are further divided into active and stative subtypes depending on the kind of agreement they take (as described briefly in the introduction).

In Kurripako (K), transitive and active intransitive verbal predicates must carry subject (Sa, A) prefixes (except in negative imperative constructions), which may co-occur with a subject NP. The agreement prefixes mark person, number, and, for 3rd person, gender of the subject. Stative intransitive verbs, in contrast, do not carry personal prefixes, but are instead followed by free pronouns. Example (1) shows a transitive verb, (2) an active intransitive, (3) a stative intransitive, and (4) a possessive NP.

(1) K *nu*-ira-ka *patsiaka.*

`1sg-drink-prog manioc.drink`

'I am drinking manioc drink.'

(2) K *nu*-dia-ka-wa *panti-liku.*

`1sg-return-prog-intr house-loc`

'I am returning into the house.'

(3) K *haamaa-ka bhuua.*

(4) K *nu*-tsinu-ni

`be.tired-prog 1sg 1sg-dog-poss`

'I am tired. ' 'my dog'

First and 2nd person Ps are not marked on the verb. The independent pronouns, undifferentiated for syntactic function, appear after the verb; it is only their position (after the verb) that signals that they are coding a P. Third person objects (singular feminine, singular non-feminine, and plural) may be expressed by either suffixes or nouns (or pronouns). These suffixes do not co-occur with either the pronouns or the nouns, as shown in (5–7).
Agreement in two Arawak languages

(5) K nu-heema kalaka

1SG-barbeque CHICKEN

'I barbecue the chicken (fem.).'

(6) K nu-heema srua

1SG-barbeque 3SG.F

'I barbecue it (fem.).'

(7) K nu-heema-no

1SG-barbeque-3SG.F

'I barbecue it (fem.).'

In Kurripako, a recipient or beneficiary is marked with a dative marker, which appears before the P; it is not grammatical for a P to show agreement on the verb when there is a dative argument in the clause, so it must be expressed with independent pronouns or nominals as shown in (8).

(8) K wa-a-pia li-sru peethe

1PL-give-PFV 3SG.M-DAT manioc.bread

'We gave him manioc bread.'

The 3rd person focus prefix in Kurripako must be used when there is a fronted 3rd person nominal subject. This is exemplified in (9), in which the subject is focused and can be contrasted with (10), which does not have an emphatic reading.

(9) K atsinali i-ira-ka patsiaka

man 3SGFOC-drink-T/A manioc.drink

'The man drinks manioc drink.'

(10) K li-ira-ka patsiaka atsinali

3SG.M-drink-T/A manioc.drink man

'The man drinks manioc drink.'

16.2.2 Sₚ marking: stative verbs in Kurripako

Sₚ marking is marking of the intransitive subject with the affixes otherwise used to show agreement with a P. For Kurripako, this is found with stative intransitive verbs, as shown in (3) above, in contrast to active intransitive verbs, as in (2). The use of the P affixes with these verbs is obligatory; other forms of marking, e.g. an active intransitive verb with an external free pronoun (11), or a stative intransitive verb with a pronominal prefix (12), are ungrammatical.

(11) K *dia-ka-wa hmuu panti-liku.

return-PROG-INTR 1SG house-LOC

'I am returning into the house.'

(12) K *nu-haampa-ka.

1SG-be.tired-PROG

'I am tired.'

In terms of agreement patterns, some intransitive subjects pattern with As and some with Ps. The use of the person markers is exactly the same as on transitive verbs, including the bound forms for 3rd person singular objects.
16.3 Baure

The seriously endangered language Baure belongs to the South Arawak branch of the language family, together with the Moxo languages\(^4\) Trinitario and Ignaciano of Llanos de Moxos in Bolivia. It is spoken in the Bolivian Amazonia towards the border with Brazil. The data were collected by Danielsen during field trips in 2003 and 2004. Baure is a polysynthetic, head-marking language with VSO order, gender, an extensive noun class system, and no case marking for core arguments. Baure uses agreement clitics on possessed nouns and on different kinds of predicates, as will be described in the following sections.

16.3.1 Nominative subjects: verbal predicates in Baure

Baure shows the same nominative-accusative alignment in its agreement affixes for all verbal predicates, with the subject (S, A) showing agreement by an obligatory proclitic\(^5\) which may co-occur with a coreferent NP. The full set of agreement clitics is shown in Table 16.2.

Example (13) shows the clitic on a transitive verb, (14) shows two active intransitive verbs with the same agreement clitic, (15) shows two stative intransitive verbs, and (16) shows that the same clitics are also used in possessive NPs.

\(^1\) The unspecified prefix is presumably only used with bound nouns, but not on verbs.

**Table 16.2. Baure agreement clitics**

<table>
<thead>
<tr>
<th>Person</th>
<th>Nominative</th>
<th>Accusative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sing.</td>
<td>Pl.</td>
</tr>
<tr>
<td>1</td>
<td>ni-</td>
<td>vi-</td>
</tr>
<tr>
<td>2</td>
<td>pi-</td>
<td>yi-</td>
</tr>
<tr>
<td>3m</td>
<td>ro-</td>
<td>no-</td>
</tr>
<tr>
<td>3f</td>
<td>ri-</td>
<td>no-</td>
</tr>
<tr>
<td>UNSP(^6)</td>
<td>e-</td>
<td>–</td>
</tr>
</tbody>
</table>

(13) B \(vi=nik\) mos

\(\text{ipl}=\text{eat sweet corn}\)

‘We eat sweet corn.’

(14) B \(vi=sim\) \(ač\) \(vi=kowyo-po\) \(ač\) \(vi=imok\)

\(\text{ipl}=\text{arrive}\) and \(\text{ipl}=\text{bathe-prfLX}\) and \(\text{ipl}=\text{sleep}\)

‘We arrive, we take a bath (bathe ourselves), and we sleep.’

---

\(^4\) Moxo or Mojo are alternative ways of spelling the same referent, in English and Spanish. The Moxo languages Trinitario and Ignaciano are very closely related.

\(^5\) In Baure, unlike Kurripako, the agreement markers have been analysed as clitics. This type of clitics has to be considered as on the borderline between affixes and clausal clitics.
Agreement in two Arawak languages

(15) B vi=ha’ino-wo aĉ vi=imoko-s’-ino-wo
   1PL=be.tired-COP and 1PL=sleep-aprx-subj-COP
   ‘We are tired and (we are) sleepy.’

(16) B vi=tovian
    1PL=neighbour
    ‘our neighbour’

Transitive verbs show agreement for Ps by means of an enclitic. Unlike the nominative agreement seen in (13–16), accusative markers may not co-occur with an object NP; cf. (17) and (18). Free personal pronouns are used for special emphasis, as in the exceptional preverbal use of nti ‘1sg’ coding the subject in (17), and cannot replace agreement on the verb.

(17) B nti’ ni=komorik p-a-š simori
    1SG 1SG=kill one-clf:animal-one pig
    ‘I killed one pig.’

(18) B heni, ver ni=komorikie=ro
    yes PVF 1SG=kill=3sg.m
    ‘Yes, I already killed it.’

Baure ditransitive verbs (which are mainly derived) allow overt agreement for both objects; the clitic showing agreement with the recipient precedes that of the patient, as demonstrated in (19) and (20). 6

(19) B pi=pa=ni=ro, ni=pa=pi=ro
    2SG=give=1SG=3SG.m 1SG=give=2SG=3SG.m
    ‘You give it to me, I give it to you.’

(20) B ni=wo’ik-ino=pi=ro
    1SG=butcher-ben=2SG=3SG.m
    ‘I butcher it for you.’

16.3.2 Accusative subjects: nonverbal predicates in Baure

In Baure, stative nonverbal predicates (not to be confused with stative verbal predicates) take the accusative clitics that are used for object agreement on transitive verbs; cf. (17) and (18). In order to be used as a predicate with agreement, the nominal or adjectival base must be derived with the copular suffix, which functions as a kind of verbalizer. In addition to or instead of the copula, other verbal morphemes may be attached to the base.

6 There are restrictions on the combination of agreement clitics that may appear; generally, two ro-/ri 3sg (m/f) clitics are not permitted in series, presumably for phonetic reasons.
(21) B monˇci-wo=ni n=asore-he-wo
    child-cop=1sg 1sg=be.strong-distr-cop
    ‘When I was a child I was very (completely) strong.’

It may be added here that nouns and adjectives can function as predicates without person cross-reference, when the subject is explicit and no aspect or time specification is necessary, being clear from the context. In that case a nominal or adjectival predicate is juxtaposed with the subject nominal, as demonstrated in (22). Example (22) shows the completely unmarked nominal predicate monˇci ‘(be a) child’.

(22) B ...koˇce  te  ni=šir monˇci, ti ni=hin monˇci napiri’
    because dem1.1sg=son child dem1.1sg=daughter child also
    ‘...because my son is (still) a child, and my daughter is also (still) a child.’

Adjectival predicates may be treated similarly. Example (23) shows stative predicates based on the adjective mehewkon ‘bad’. In both cases the adjectival base functions as predicative and displays the morpheme -wapa ‘change of state’ suffixed. The derived verb does not show agreement in the first clause (mehewkon-wapa) due to the presence of an explicit nominal subject te peˇcpi’ ‘your roof’. In contrast we find the same predicate in the second clause with agreement (mehewkon-wape-ro) and no explicit subject. The accusative clitic -ro ‘3sg.m’ in mehewkon-wapero ‘it is bad’ refers to the subject ‘your roof’, which is a non-feminine noun.

(23) B mehewkon-wapa te  p=eˇcpi’, hen, mehewkon-wape=ro.
    bad-cos dem1.2sg=roof yes  bad-cos=3sg.m
    ‘Your roof is already bad (falling apart), yes it is bad.’

In contrast, all verbs referring to states as well as weather verbs in Baure are marked nominatively, cf. nasorohew ‘I was very (completely) strong’ in (21) and in the verbs for ‘to be cold’ in (24) and (25).

(24) B ni=mane-wapa.
    1sg=be.cold-cos
    ‘I am cold/ I got cold.’

(25) B nokope’ ro=tokenoko-wo te ahikowon.
    yesterday 3sg.m=be.cold-cop dem1.m morning
    ‘Yesterday the morning was cold.’

16.3.3 Accusative subjects in Baure
We have shown that Kurripako shows a split of argument marking on intransitive verbs, just as it has been suggested for Arawak languages in general; cf. section 16.1
Agreement in two Arawak languages

above. Baure, on the other hand, distinguishes verbal from nonverbal predicates by different argument marking. There is, however, a small subgroup of frequently used verb-like predicates in Baure with accusative agreement. An exhaustive list of the known members of this class is kwo- ‘exist, be’ (copular predicate), koehoe- ‘give birth’, and eto- ‘finish’. This kind of accusatively marked predicate can be intransitive, as in (26), but also transitive, as in (27) and (28), and double argument marking on a transitive predicate may occur (29).

Intransitive accusative predicates

(26) B kwo=ni ne’ pi=weri-ye.
exist=1sg here 2sg=house-LOC
‘I am here in your house.’

Even though the following two clauses seem to represent transitive predicates, it is not clear if the NP pinonev ‘twins’ in (27) can really be analysed as a P, because it cannot be replaced by a personal enclitic, a feature of other Ps. The subject NP in (27) on the other hand, can be pronominalized.

Transitive accusative predicates

(27) B koehoe=ri pino-nev.
give.birth=3sg.f twin-pl
‘She gave birth to twins.’

(28) B ver eto=ni to ni=vesa-ˇc.
finish=1sg art 1sg=read-nom2
‘I already finished (my) reading.’

Double marking on a transitive accusative predicate

(29) B aˇc ver eto=ro=ni.
and PFV finish=3sg.=-1sg
‘And I already finished it (the weeding and cleaning of the field).’

In order to decide whether we are dealing with verbs in examples (26–29) it is important to find out more about their origin; thus, in the following we shall tentatively offer some hypotheses in this direction. The copular predicate kwo- ‘exist, be’ is composed of the attributive prefix kO- and the copular suffix -wo (see example (26)). It patterns like a predicate with a nominal base, in that it appears with the copular suffix, and in that the subject shows accusative agreement.

The predicate koehoe- ‘give birth’ has also been derived by means of the attributive prefix kO- and another element, presumably a noun ‘-ihoe’. Even though both predicates seem to have evolved from a combination with the attributive

7 In Baure there are numerous words whose stems are nothing more than grammatical morphemes in idiosyncratic (lexicalized) combinations.
prefix, koehoe- ‘give birth’ can be transitive. Another interesting point is that the attributive generally functions as a verbalizer (as shown in section 16.5 below), and triggers subjective marking, as in (34). Many languages, however, exhibit special verbs or constructions for the predicate ‘give birth’ and for copulas. In addition, the lack of control of the subject in both cases may have led to the accusative construction. It is not clear, however, whether the predicates are best interpreted as exceptional verbs (possibly showing the last relics of the Arawak semantic alignment system), or should be regarded as instances of exceptional non-verbal predicates.

The predicate eto- ‘finish’ also functions as an aspectual preverbal particle for terminative, just like the adverb ver ‘already, PERF, which also functions as the preverbal particle for perfect. 8 The predicate base eto- may also have been a temporal adverb *eto with the meaning ‘at the end, over’. The predicate eto- ‘finish’ patterns like nonverbal predicates with two exceptions: first, there is no obligatory verbal morpheme before the person enclitic; secondly, the predicate can be transitive and may take two accusative clitics, for both subject and object, where the clitic for the object precedes that of the subject.

The predicate eto- ‘finish’ behaves more like a verb than the other accusative predicates. It seems reasonable to assume that it has been grammaticalized from a non-verbal predicate into the present form, which can get double marking of constituents like ditransitive verbs. It cannot be regarded as a typical modern Baure verb, however, because of the accusative marking strategy.

In sum, Baure shows two different kinds of argument marking for semantic subjects, one marking for subjects of verbal predicates and another for subjects of non-verbal predicates.

16.4 Attributive and privative derivations

A common feature of the Arawak languages is the presence of an attributive prefix (ka- in Kurripako, kO- 9 in Baure) and a privative prefix (ma- in Kurripako, mO- in Baure). In Kurripako, both prefixes may derive stative verbs from nouns and in both cases the derived verb takes S by marking, as shown in (30) and (31).

\( \text{(30)} \) K ka-ipe-ka kutsi \( \text{ATTR-meat-T/A pig} \)

\( \text{‘The pig is fat.’} \)

\( \text{(31)} \) K ma-ipe-ka kutsi \( \text{PRIV-meat-T/A pig} \)

\( \text{‘The pig is thin.’} \)

When a derived verb takes a prefix, person markers cannot appear. We hypothesize that this is because there is only one prefix position in Kurripako verbs. This

---

8 When eto is used as a preverbal particle for terminative, the verb is not in a subordinate construction; this contrasts to the use of eto- as a main predicate in (35), where the second predicate is marked by the action nominalizer, showing that it enters into a subordinate construction.

9 The capital O refers to the weakness of this vowel, which may be dropped or may change according to vowel harmony.
hypothesis is further supported by data from the imperatives. Unlike other verbs, imperatives do not have any T/A marking. Positive imperatives carry 2nd person prefix markers, either singular or plural, but this is not true of negative imperatives (prohibitions).

Negative imperatives appear with a privative marker ma- and a restrictive suffix -tsa. When these affixes are added, the normal agreement suffixes are not found and the subject can only be expressed by means of an independent pronoun, just like Ps. 

(32) K ma-pita-tsa!
   PRIV-bathe-RES
   'Don’t bathe!'

This can be considered evidence for the presence of only one prefix slot and the necessary Sp marking of some of the stative verbs. When one of these verbs has a second argument, this second argument is assigned an oblique case, such as locative as seen in example (33).

(33) K ka-ako-ka hmaa kuripako-liku
   ATTR-word-T/A 1SG Kurripako-LOC
   'I speak Kurripako.'

In Baure the privative prefix mO- ‘without’ derives negative opposites mainly of attributive nominals (very often in the function of modifiers) which otherwise are marked by the attributive prefix kO- ‘with’. In contrast to Kurripako, attributive predicates are marked nominatively, as in (34); like Kurripako, the derived privative counterparts are marked differently, with the accusative clitics. Example (35) shows that privative predicates in Baure behave like nonverbal predicates. It can be argued that the attributive functions as a verbalizer, while the privative derives nouns (or adjectives).

(34) B ver ri= k-avinon
   PFV 3SG.F=ATTR-husband
   'She is already married.'

(35) B m-avinon-e=ri?
   PRIV-husband-LK=3SG.F
   'Is she unmarried?'

One difference with respect to other nonverbal predicates is that privative predicates do not need the copula morpheme before the person enclitic, but instead the linker -a11 (which in 35 has undergone a morphophonological change -a > -e). Example (35) can be compared to (21) above, in which the base of the nonverbal predicate was a noun. The predicate in (21) can only get the person clitic attached when it is preceded by a verbal morpheme, such as the copula suffix -wo. Example

10 Modifiers or adjectives in Baure all show nominal properties. They can, however, be distinguished from nouns morphologically and syntactically.
11 The linker -a has a very wide range of functions, among others connecting stative bases (verbal or nonverbal) with non-stative suffixes; it follows incorporated nouns or classifiers and may occur in noun compounds.
(23) also exhibits an example of a nonverbal predicate based on an adjective, showing the same characteristics.

Note the derivational differences between Kurripako and Baure in the use of these prefixes. Whereas in Kurripako the attributive or privative prefixes derive stative verbs with objective marking, in Baure the privative prefix derives nonverbal predicates taking objective marking and the attributive prefix derives verbal predicates taking subjective marking.

16.5 Comparison

As seen earlier in Tables 16.1 and 16.2, both Kurripako and Baure show verbal agreement for core arguments but differ with respect to details of the inventory.

Note the paucity of object suffixes in Kurripako compared to Baure (Aikhenvald 1999 lists four Arawak languages that have completely lost their agreement suffixes). In Kurripako, agreement suffixes are used only for P and S_p. When there is both a recipient and a patient, the recipient is indicated by a dative marker and the patient follows the recipient. This is the only order allowed; it is the same as the order of the enclitics in Baure (cf. examples (19) and (20) for Baure and (8) for Kurripako).

There are significant differences in the agreement patterns of intransitive sentences, however. Kurripako separates active subjects from stative subjects by marking the first group similarly to an A and the second group similarly to P. Baure, in contrast, marks the subjects of all verbal clauses identically, using nominative agreement, regardless of whether they are S or A. It is only subjects of nonverbal predicates that have argument markings like P. The split in this case is between verbal and nonverbal predicates and not between different kinds of intransitive verbs. The question then becomes whether these two different cases can both be considered semantic alignment.

16.6 Semantic alignment?

Mithun (1991) explores the semantic bases of the phenomenon of split alignment systems, and presents three prototypical cases that have different semantic features triggering the different agreement markers used with different types of intransitive verbs. Table 16.3 summarizes the features, provides examples of typical verbal meanings, and shows the distribution of agreement that differentiates each of the three language types. Stative-active languages are those that follow the first pattern, exemplified here by Guaraní. Guaraní differentiates intransitive verbs for lexical aspect, specifically for eventhood (Mithun 1991: 523) by marking two different cases. Intransitive event verbs are marked by agreement pattern ‘I for active or grammatical agent’ while all ‘event’ verbs are marked by agreement pattern ‘II for stative or grammatical patient’. Examples (36)–(40) parallel A–F
**Table 16.3.** Summary of agreement marking (Mithun 1991: 524)

<table>
<thead>
<tr>
<th></th>
<th>Guarani</th>
<th>Lakhota</th>
<th>Central Pomo</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. +event 'jump', 'go', 'run'</td>
<td>I</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>+P/E/I*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+control</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. +event 'hiccough', 'sneeze', 'vomit'</td>
<td>I</td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>+P/E/I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>−control</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. +event 'fall', 'die', 'slip'</td>
<td>I</td>
<td>II</td>
<td>II</td>
</tr>
<tr>
<td>−P/E/I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>−control</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D. −event 'reside', 'be prudent', 'be patient'</td>
<td>II</td>
<td>I</td>
<td>I</td>
</tr>
<tr>
<td>+P/E/I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+control</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. −event 'be tall', 'be strong', 'be right-handed'</td>
<td>II</td>
<td>II</td>
<td>I</td>
</tr>
<tr>
<td>−P/E/I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>−control</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>−affect</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>F. −event 'be sick', 'be tired', 'be cold'</td>
<td>II</td>
<td>II</td>
<td>II</td>
</tr>
<tr>
<td>−P/E/I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>−control</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>+affect</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* P/E/I stands for performed/effected/instigated. The symbols I and II identify the case used with such verbs in each language—I for active or grammatical agent, II for stative or grammatical patient’ (Mithun 1992: 523).

in the table (with D missing), supporting the claim that the Kurripako semantic alignment is based on eventhood.

(36) K **mu-ito kenke-riku.** (cf. A in Table 16.3)

1sg-go 1sg-foc/go field-LOC

'I go to the field.'

(37) K **hlia hmapeni i-katha leche.** (cf. B in Table 16.3)

3sg.m child 3sgfoc-vomit milk

'The child vomits the milk.'

(38) K **heri i-hiwa hiipai-naku.** (cf. C in Table 16.3)

horsefly 3sgfoc-fall dirt-LOC

'The horsefly falls on the dirt.'

(39) K **halipa hlia-hi juan.** (cf. E in Table 16.3)

be.tall 3sg.m-dem juan

'Juan is tall.'
Examples of verbs of type D are not included because they are absent from the Kurripako database. There is the verb 'to live', but it is not clear that this is the same as 'to reside'. The main difficulty in trying to find other examples that could be substituted for 'to be prudent' and 'to be patient' is that the assignment of the semantic feature of control (which is the critical feature in this case) is one that is very much tied to cultural values and very difficult to discern if the right question has not been asked or the critical contrast is not present in the data. That said, it is still possible to fit Kurripako into one of the three types without this set of data because of the distribution of the argument marking. Examples for sets A, B, and C all pattern together having prefixes, whereas examples for sets E and F pattern together having suffixal patient type marking. So this means that it can only fit into the first distribution within this typology, and we would expect to find that verbs of type D also have their semantic subject marked like transitive objects.

In Baure an intransitive S argument shows the same nominative agreement that characterizes the A of a transitive verb. It is, however, possible to have accusative marking for predicates of type D–F on the base of adjectives or nouns. This is exemplified in the following alternative examples:

(40)  K  hnete haamaa kepinaa hlia.  (cf. F in Table 16.3)
     dem be.tired be.full 3SG.M
     ‘Then he was tired and full.’

(41)  B  aĉ teĉ šiye pikor, pikoro-wo=ro  (cf. D in Table 16.3)
     and dem2.M fox rascal-cop=3SG.M
     ‘And that fox is rascally, he is a rascal.’

(42)  B  to eton ě-işiê, ě-işiê-wo=ri  (cf. E in Table 16.3)
     art woman big-size big-size-cop=3SG.F
     ‘The woman is tall, she is tall.’

(43)  B  aĉ mavi-wape=ro  (cf. F in Table 16.3)
     and sick-cop=3SG.M
     ‘And he is (seriously) ill.’

All the predicates in (41)–(43) are constructed on the basis of adjectives, and nonverbal predicates take accusative marking in Baure. The results of comparison are summarized again in Table 16.4. As is common among Arawak languages, Kurripako follows the pattern of an active-static alignment system based on eventhood, whereas Baure shows a nominative-accusative system in its verbal clauses.

Given the data presented above, we argue that Kurripako has a semantic alignment system in which eventhood is the relevant semantic feature that gives rise to a different marking system for stative intransitive subjects vs. agents and active intransitive subjects. The split in Baure is between verbal and nonverbal predicates. Eventhood is still a semantic feature that groups the nonverbal predicates
### 16.7 Fluid-S marking on verbs

In Kurripako, there is one pair of verbs that can be described in terms of fluid semantic alignment. The same verb base may take either $S_a$ and $S_p$ marking with different meanings, as seen in (44) and (45).

| Table 16.4. Summary of agreement strategies in Kurripako and Baure (cf. Table 16.3 and Mithun 1991: 524) |
|---|---|
| Kurripako | Baure |
| A. +event 'jump', 'go', 'run' | I$^a$ | I |
| +P/E/I +control | | |
| B. +event 'hiccough', 'sneeze', 'vomit' | I | I |
| +P/E/I | | |
| −control | | |
| C. +event 'fall', 'die', 'slip' | I | I |
| −P/E/I | | |
| −control | | |
| D. −event 'reside', 'be prudent', 'be patient' | II | I/(II) |
| +P/E/I +control | | |
| E. −event 'be tall', 'be strong', 'be right-handed' | II | I/(II) |
| −P/E/I | | |
| −control −affect | | |
| F. −event 'be sick', 'be tired', 'be cold' | II | I/(II) |
| −P/E/I | | |
| −control +affect | | |

$^a$ Here I stands for $S_a$ marking, and II stands for $S_p$ marking.

Together, but it is secondary to the word class assignment, since non-eventive verbal predicates show nominative agreement. We conclude that Baure cannot be considered an example of semantic alignment, since the differential marking of subjects of intransitives has no semantic basis, but is determined by the syntactic category of parts of speech. Yet the system may have evolved from one of semantic alignment. Possibly stative intransitive verbs and nonverbal predicates were earlier marked in the same way. Such a system could easily change into the current one if all intransitive verbal predicates became regularized. The lack of historical and comparative data, however, makes this hypothesis impossible to test at present.
In Baure there are also some traces of a fluid system in argument marking. There is at least one example of a verb base that can have both nominative and accusative marking with slightly different interpretations. This is the verbal base -moro’in(o)- ‘be sick of/be thirsty’, which can either take regular nominative marking (46), in which case it means ‘be sick of’, or accusative marking (47), in which case it is found to mean ‘to be thirsty’.

46. B ver kač=hi ri=nik, ri=nik, ri=nik, 
PFV GO=QUOT 3SG.F=eat 3SG.F=eat 3SG.F=eat
nka kač=hi ri=moro’ine=ro=hi. 
NEG GO=QUOT 3SG.F=be.sick.of=3SG.M=QUOT
‘She started to eat and eat and eat and she didn’t get sick of it.’

47. B ikarek morö’ino-wape=ri =hi 
thus be.thirsty-COS=3SG.F=QUOT
ač ri=kač-po-w wapoeri-ye=hi.
and 3SG.F=GO-PREFLX-COP RIVER-LOC=QUOT
‘Thus she got thirsty and went to the river.’

These examples, even if they are few, provide some evidence of an earlier system of at least partial fluid semantic alignment. It is possible that there were more verbs behaving in this way, but there are no historical data to ascertain this. The presence of the phenomenon in two geographically separate and genetically distant languages suggests that fluid marking was a characteristic of proto-Arawak.

16.8 Conclusions
Kurripako and Baure differ in their predicate argument marking. Kurripako follows the prototypical active-stative split as described for Arawak languages in general, whereas Baure predicates can be divided into verbal and nonverbal syntactic classes. The accusative marking in Baure was presumably originally a feature of stative predicates, but nowadays it is only nonverbal predicates which exhibit this agreement pattern. Some traces of Fluid-S making, combined with minor patterns in the dataset, hint at a formerly richer class of accusative predicates, possibly

12 It is possible that the base of the accusative-marking predicate is in fact the adjective "moro’in ‘thirsty’—adjectives are treated like other nonverbal base, showing accusative marking. The final syllable -m(o) may also be reanalysed as the nominalizer -no ‘NOMi’.
including verbs, and suggest that a system of semantic alignment similar to that of Kurripako might be reconstructed for Baure. Most contemporary Baure verbs are not marked accusatively, however. Kurripako also presents traces of fluid semantic alignment, suggesting that this was a feature of proto-Arawak (contra Aikhenvald 1999).

It is our hope that the data provided here from two distantly related and geographically disparate Arawak languages will further our understanding of this language family in particular and of argument marking systems in general. Given the relatively shallow time depth of the Arawakan family, which is supposed to have broken up around 3,000 years ago (Zucchi 2002), our study suggests that the active/stative distinction may be typologically unstable (see Mithun, this volume, for a similar observation with respect to semantic alignment in North America).
Affectededness and viewpoint in Pilagá (Guaykuruan): a semantically aligned case-marking system

ALEJANDRA VIDAL

17.1 Introduction*

Pilagá is a South American language spoken in the province of Formosa, Argentina, by approximately 6,000 people. Pilagá enclaves are located in the Chaco region, the most linguistically diverse area in the country, in terms of Native American peoples and languages. The Chaco region extends over three countries, Argentina, Paraguay, and Bolivia.

The purpose of this chapter is to present the analysis of a semantically aligned system that shares some characteristics with both agentive/patientive case marking and active vs. middle voice systems. However, as it happens with other agentive systems, for many verbs the assignment of a particular stem to a prefix seems to be lexicalized. This phenomenon was discussed by Mithun (1991) for North American languages and also for Guaraní. Moreover, Pilagá case marking exhibits a number of oddities (e.g. the encoding of trajectory) which are discussed here with a view to a typological comparison with other systems described in the literature.

The chapter is organized as follows. In section 17.2 I present the two sets of verbal prefixes which constitute the case-marking system and show how they work formally with different verb classes. In section 17.3, I discuss situation types where either sets may occur. In section 17.4 I briefly discuss the interaction between case

* I would like to acknowledge all the Pilagá speakers who in one way or another made possible the writing of this chapter. I am particularly thankful to Ignacio Silva, Federico Pérez, and José Miranda, who participated in sentence elicitation at different stages of my research. I also thank the editors for their constructive comments and suggestions. Financial support was received from the National Science Foundation through a pre-dissertation improvement grant (1996–7), the Graduate School at the University of Oregon (1996), the Instituto de Asuntos Indígenas, Argentina (1997–8), and the Fundación Antorchas, Argentina (1998–2000).
Affectedness and viewpoint in Pilagá

Table 17.1. ‘Agentive-patientive’ case-marking pronominal prefixes in Pilagá

<table>
<thead>
<tr>
<th>Set A</th>
<th>Set B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>-s</td>
</tr>
<tr>
<td>2SG</td>
<td>aw-</td>
</tr>
<tr>
<td>3SG</td>
<td>d-</td>
</tr>
<tr>
<td></td>
<td>t-</td>
</tr>
<tr>
<td></td>
<td>i-</td>
</tr>
<tr>
<td></td>
<td>w-</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>1PL</td>
<td>s--...-s-o-sa</td>
</tr>
<tr>
<td></td>
<td>s--...-q(a)q</td>
</tr>
<tr>
<td></td>
<td>s--...-qa</td>
</tr>
<tr>
<td></td>
<td>s--...-t(a)</td>
</tr>
<tr>
<td>2PL</td>
<td>aw--...-i</td>
</tr>
<tr>
<td></td>
<td>aw--...-e</td>
</tr>
<tr>
<td></td>
<td>aw--...-q(a)q</td>
</tr>
<tr>
<td></td>
<td>aw--...-q(æ)q</td>
</tr>
<tr>
<td>3PL</td>
<td>d--...-q</td>
</tr>
<tr>
<td></td>
<td>t--...-d</td>
</tr>
<tr>
<td></td>
<td>i--...-d</td>
</tr>
<tr>
<td></td>
<td>ø--...-d</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

marking and voice, arguing that according to our analysis this is not a voice-based system. Final remarks are found in section 17.5.

17.2 Case marking

In Pilagá, pronominal prefixes constitute the only case-marking device, since the language has neither case markers on nouns or free pronouns nor adpositions for nominals. Set A and set B index the subject and a separate third set of prefixes marks the object of transitive and ditransitive verbs. The pronominal affixes are shown in Table 17.1. Here the following conventions are used. Affixes to the left of a row of dots are prefixes and affixes to the right are either suffixed or infixed to the root. Prefixes represent person and suffixes number. Allomorphs are separated by ∼. Variation among affixes is neither phonologically conditioned nor, apparently, governed by structural properties of the predicates.

Case marking is partially semantically based. Set A prototypically references a performer or a source, with or without control. Set B prototypically references subject participants affected by the event, or which have no control over it. To
some extent Pilagá set A and set B forms index the semantic role of transitive and of intransitive subjects. The first represents participants conceived as semantic agents, whether the clause is transitive or intransitive; the second set represents subject participants acting as semantic patients or as experiencers, again in both intransitive and transitive sentences. Although this schema accounts for the vast majority of the prefix choices on Pilagá verbs, the assignment of a particular case to a verb is still lexicalized.

A separate set of prefixes mark the object of transitive and ditransitive verbs. Object prefixes can combine with either set A or set B prefixes. From a semantic point of view, object prefixes generally encode a human participant who receives the action denoted by the transitive verb. For this reason, I term this participant Dative. Unlike subject prefixes, which can co-occur with a pronoun or full NP, object prefixes and full pronouns or lexical NPs are mutually exclusive.

The object markers are presented in Table 17.2. They share some similarities with possessive prefixes, and partially also with set B prefixes (Table 17.1), particularly for the 2nd person singular. Note that the paradigm only provides procliticized forms for the singular. First and 2nd person plural are expressed through full pronouns exclusively:

<table>
<thead>
<tr>
<th>Person</th>
<th>物 marker</th>
</tr>
</thead>
<tbody>
<tr>
<td>1SG</td>
<td>yi-</td>
</tr>
<tr>
<td>1PL</td>
<td>qom-qi</td>
</tr>
<tr>
<td>2SG</td>
<td>am-</td>
</tr>
<tr>
<td>2PL</td>
<td>am-qi</td>
</tr>
<tr>
<td>3SG</td>
<td>ø-</td>
</tr>
<tr>
<td>3PL</td>
<td>ø-</td>
</tr>
</tbody>
</table>

As noted, the object is indexed morphologically if it is a speech act participant, i.e. a 1st or a 2nd singular person. Third persons singular and plural are always zero. Allomorphs for 1st person singular are lexically assigned. Historically, possessive prefixes and patientive pronominal prefixes seem to be the source for object prefixes in Pilagá, and, like object prefixes, 1st person possessive allomorphs are also lexically assigned.

Besides the specific set of prefixes for 1st and 2nd person singular, full pronouns after the verb (hayim ‘I’, am ‘you’) can be used to fulfil the object function, instead of the prefix. Nevertheless, for the coding of plural participants there are no prefixes. Only full pronouns qom-qi ‘we’, am-qi ‘you (pl.)’ are used, which also surface to the left of the verb. In Pilagá there is only one set of independent pronouns: hayim ‘I’, am ‘you’, qom-qi ‘we’, am-qi ‘you (pl.)’. These are used irrespective of whether the participant is a subject or an object.

Thus, to mark the object participant either a pronoun or a prefix is used, but not both. In (i) I provide examples illustrating object prefixes in combination with set A and set B subject prefixes.
Affectedness and viewpoint in Pilagá

A separate subsystem of number affixes is used to encode number of the object. These forms are always suffixed. They occupy the last position in the verbal complex and are optional.

Although object number suffixes have scope over the object participant of a transitive verb, a few instances where the object number suffixes have scope over the intransitive subject have also been found in the data.

Object number suffixes can co-occur with noun phrases or free pronouns. Pilagá has three morphemes for object number: -a ‘singular’, -to ‘paucal’ (referring to a small group of referents), and -lo ‘plural’ (referring to a larger number of entities than those included in the paucal category). Although not difficult to elicit, the paucal number suffix is a very low-frequency morpheme. Usually the plural suffix assumes the pluralizing function, no matter whether there are two or more entities involved.

17.2.1 Overall distribution of case marking

I will start by discussing the distribution of subject case marking for the first two groups of verbs, i.e. those that can only occur with either set A or set B only.

Set A and set B are formal labels for classes of forms. I have avoided the use of semantic-sounding terminology such as active/stative or agentive/non-agentive because in Pilagá, for many verbs the grammatical coding does not always align with semantics of agentivity/non-agentivity.
17.2.2 Set A-only verbs

In the Pilagá case-marking system, the Agent does not always exercise control over the action; rather, the Agent seems to be a performer or a source with or without control.

The verbs presented in (2) are events characterized as activities (Vendler 1967). These verbs take set A case marking only.

(2) a. se-taqta-tak
   1a-speak-prog
   ‘I am speaking.’

b. se-seta-ya
   1a-criticize-hab
   ‘I criticize.’

c. s-talaq
   1a-shout
   ‘I shout.’

d. s-ae-ta
   1a-go/move-dir.out of
   ‘I come out.’ (surge, e.g. out of the water)

e. s-ae-yi
   1a-go/move-dir.downtowards/inside
   ‘I go.’ (e.g. to the forest)

f. s-atfija-qi-ye’
   1a-procede-dir.thither
   ‘I come from.’ (e.g. a far place)

g. s-atfijyaq-ya
   1a-procede-dir.downtowards
   ‘I descend from.’

h. s-aw’o
   1a-make
   ‘I make/prepare.’

i. se-ke’le
   1a-eat
   ‘I eat.’

j. se-wala’a
   1a-play
   ‘I play.’

Verbs in (3) involve a physical activity of some sort where the subject is definitely an Agent, according to our definition of Agent as a performer (this even goes for 1sg tijaqa ‘descend (genetic lineage)’, which should be taken as a metaphorical extension of the motion verb tjiya ‘proceed from a location’). But as soon as one starts looking at other verbs that fall into this group, problematic cases appear. For instance, mental activity predicates are also found in the class of set A-only verbs, as seen in (3).

(3) a. s-ateto-n
   1a-know-nprog
   ‘I know (a person or place).’

b. sa-yate-n
   1a-know-nprog
   ‘I know.’ (somebody told me)

In fact, some set A-only marked predicates in Pilagá are verbs of perception and verbs of involuntary bodily processes (4a, b). It could be argued that verbs of perception and verbs of mental processes are not subject to control because of a certain lack of volition on the part of the subject participant, as in (3) and (4). In (4), verbs like ‘cough’ or ‘scratch’ address a more spontaneous behaviour than, for instance, ‘play’, ‘eat’, or ‘shout’ given in (2). But in all such cases, the participant is marked by a person prefix from set A.
Affectedness and viewpoint in Pilagá

(4) a. se-sena\text{nā-n}  
\text{1A-scratch-NPROG}  
'I scratch (myself).'

b. sa-qaeyoso\text{–n}  
\text{1A-cough-NPROG}  
'I cough.'

c. s-awe\text{–t}e-\text{e}\text{xet}  
\text{1A-observe-EPVOW-DIR}  
'I observe.'

d. se-lota-p\text{e}\text{yə}  
\text{1A-look-HAB}  
'I see something.'

Up to this point, set A-only case-marked verbs might suggest that one general semantic feature for this class is that the subject participant is understood as a performer or as a source, with or without control. However, in Pilagá other verb types, such as states, also appear in this class. Verbs in (5), which are all states (in particular, emotion predicates) where the only participant is more an Undergoer than an Agent, show that 'control' does not play a role in the choice of case-marking here.¹

(5) a. s-e\text{k}\text{o}n  
\text{1A-headache}  
'I have a headache.'

b. s-akiko  
\text{1A-sad}  
'I am sad.'

c. s-el\text{w}\text{a}k  
\text{1A-sick}  
'I am sick.'

d. sa-soqowat  
\text{1A-be hungry}  
'I am hungry.'

Therefore the examples provided in this section show that set A can occur with different verb classes, and that many of the verbs can be used transitively, by instantiating a patient or a dative object through an object prefix or a postverbal noun phrase. I provide more examples of verbs included in the set A-only case-marked group in (6).

(6) a. Physical activity verbs: -\text{l}e\text{w} 'to die', -\text{l}a\text{t} 'to kill' (trans./refl.), -\text{o}p\text{i} 'to carry water', -\text{f}\text{i}y\text{a}f\text{o} 'to get', -\text{s}\text{o}n\text{a} 'to stick/nail something on the floor', -\text{p}\text{o}y\text{a} 'to obtain/get for oneself', -\text{t}\text{e}t\text{e} 'to prepare (for instance, food or fire)', -\text{a}l\text{i}k 'to eat', -\text{k}e\text{c}\text{e}t\text{e}/\text{k}i 'to feed' -\text{e}p\text{e}(t) 'to fish', -\text{k}\text{e}t\text{a} 'to point out', -\text{a}q\text{a} 'to throw', -\text{p}\text{e} 'to sit/move downwards', -\text{s}\text{e}\text{t} 'to move/go', -\text{t}o\text{f}e 'to sleep', -\text{l}e\text{k}\text{e}t\text{e} 'to mix', -\text{w}o\text{k} 'to sharpen'.

b. Mental and motion predicates: -\text{w}"\text{r}a\text{t} 'to observe', -\text{y}a\text{k} 'to wait', -\text{li} 'to hear', -\text{p}\text{e}t\text{a} 'to think'.

c. Statives/positionals: -\text{f}t\text{i\text{y}}\text{a}(/t) 'to be bent', -\text{s}\text{o}d\text{d}i 'to be kneeling down', -\text{k}o\text{a} 'to be asleep', -\text{n}e\text{t}a 'to reside/live/be', -\text{w}a\text{n}a 'to have'.

d. Speech act verbs: -\text{se}\text{e}t 'to criticize', -\text{s}e\text{t} 'to tell/narrate', -\text{t}a\text{q}a 'to speak'.

e. Translational motion verbs: -\text{p}ae 'to go to', -\text{k}o\text{e}(t) 'to fly', -\text{y}o\text{q} 'to carry'.

¹ With regard to the treatment of emotions in some South Asian languages, DeLancey (1984: 10) argues that predicates of emotions are considered 'agentive' or not according to the speaker’s assessment of the degree of control which the subject exercised or could have exercised.
17.2.3 Set B-only verbs

A small group of verb roots take set B prefixes exclusively. The core of the verbs in this set are intransitive predicates where the participant is an Undergoer. Examples are found in (7).

(7) a. \( \text{ñ}i\text{-ono} \alpha \text{k} \)  
\( 1b\text{-be.happy} \)  
\( \text{‘I am happy.’} \)

b. \( \text{ñ}i\text{-dof}-i\text{-ji} \)  
\( 1b\text{-be.sincere-epvov-compl} \)  
\( \text{‘I am sincere.’} \)

Some verbs that describe body postures such as ‘standing’, ‘lying’, and ‘sitting’ also pattern with non-agentive/patientive marking (8). (However, as noted above, ‘to be bent’ and ‘to be kneeling down’ pattern with set A case marking.)

(8) a. \( n\text{-tsate}-\text{tapiji} \)  
\( 3b\text{-stand up-\text{dur}} \)  
\( \text{‘He is in the process of moving to an upright position.’} \)

b. \( \text{ne-na-\text{tapiji}} \)  
\( 3b\text{-lie-\text{dur}} \)  
\( \text{‘He is in the process of moving to an extended (=lying down) position.’} \)

c. \( \text{ne-s’o\text{o}-\text{tapiji}} \)  
\( 3b\text{-sit-\text{dur}} \)  
\( \text{‘He is in the process of moving to a downwards position.’} \)

Most spontaneous bodily processes, also called ‘reaction verbs’ by Klaiman (1991: 119), have a human participant and select a prefix from set B. (But note that, exceptionally, the verb ‘cough’ in (4b) is a set A case-marking-only verb.)

(9) a. \( \text{ñ}i\text{-awek} \)  
\( 1b\text{-breathe} \)  
\( \text{‘I breathe.’} \)

b. \( \text{ñ}i\text{-laya\text{‘i}k} \)  
\( 1b\text{-menstruate} \)  
\( \text{‘I menstruate.’} \)

c. \( \text{ñ}i\text{-oye-n} \)  
\( 1b\text{-cry-asp} \)  
\( \text{‘I cry.’} \)

d. \( \text{ñi\text{-qase-n}} \)  
\( 1b\text{-sneeze-asp} \)  
\( \text{‘I sneeze.’} \)

If we assume that the core of the B-only group are intransitive undergoer predicates, some exceptions to the core meaning of the class are found. In particular, the next set (10) contains activity verbs where the participant can only be marked by set B forms, but this is not in any given state and does not undergo a change of state. Considering the verb’s lexical semantics, there is no apparent explanation for the fact that verbs in (2) co-occur with forms from set A, and verbs in (10) co-occur with forms from set B exclusively.

(10) a. \( \text{ñi-yom} \)  
\( 1b\text{-drink} \)  
\( \text{‘I drink.’} \)

b. \( \text{ñi-alogte-n} \)  
\( 1b\text{-dominate-asp} \)  
\( \text{‘I dominate (somebody).’} \)
Affectedness and viewpoint in Pilagá

Other verb roots included in the set B-only group are:

(11) a. Physical activities: *pet* ‘to shave (trans./refl.)’, *wan*e ‘to meet’, *toyoxe* ‘to spit’, *mateda* ‘to pull out/up’, *wana* ‘to catch’, *ato* ‘to recollect’, *loon* ‘to swim’, *tfele* ‘to have a bath’, *kese* ‘to joke’/‘to copulate’.

b. Positionals: *nae* ‘to be laid down’.

c. Mental and emotion predicates: *kian* ‘to listen’, *towe* ‘to remember’.

Even if the group of verbs that co-occur with prefixes from set B exclusively constitute a smaller group, one can still argue that some of these are activity verbs that, under some circumstances, can be used transitively or intransitively, where transitivity has no relation to the choice of prefix set. For example, *pet* ‘to shave’ can be used intransitively with a reflexive meaning (12a), or transitively (12a) with the addition of a patient object. But there is no change in the prefix class whatsoever: Also, as in (10a–c), a noun phrase can make explicit the affected participant or patient object, but even in those cases this does not affect the choice of prefix set. A verb like *aloq(te)* may be used transitively in (10b), or intransitively, with a reflexive suffix as in *judoqtel*tat ‘I dominate myself’, again with no change in the prefix class.

(12) a. *pi-petak*
   1B-shave
   ‘I shave myself.’

b. *yi-an-petak*
   1O-2B-shave
   ‘You shave me.’

c. *so*l siya*a*awa so*vote n-atfel
   CLF person before 3B-have.a.bath
   ‘The person had a bath.’

d. *so*l siya*a*awa n-atfel*a*yan da?* ko*vat
   CLF person 3B-have.a.bath-TR CLF.vert.extended son
   ‘That person bathed his son.’

e. *pa-kia*la-lat*
   1B-listen-REFL
   ‘I listen to myself.’

f. *yi-na-kia*la-n-a
   1O-3B-listen-NPROG-OBINUM.SG
   ‘He listens to me.’
In sum, much of the case marking in Pilagá does appear semantically determined, but the fact that many of these verb stems can appear with only one particular prefix set suggests that the system is very lexicalized.

I will now discuss the 3rd group of verbs, which constitutes the vast majority of Pilagá verb stems. This is the group that can occur with either set A or set B markers.

17.2.4 Alternating set A and set B case marking

The possibility for many verb stems to be assigned to one case or to the other suggests that Pilagá subject case marking reflects an agentive/patientive system. If that were so, one would expect that one class of prefixes might embody a participant who is in control of the event described by the verb, while the other class might encode a participant who is not in control. Thus, there should be basically one class containing verbs that correlate with agentivity, and another class correlated with stativity or non-volitionality.

However, I have shown above that in Pilagá, participants coded by set A forms do not always exhibit control or instigation, and that participants marked by set B forms do not always demonstrate affectedness and non-instigation. Verbs such as ‘to be sad’, ‘to be sick’, or ‘to cough’ occur with set A case marking, whereas the subject of verbs such as ‘to be sincere’ and ‘to sneeze’ are coded with set B prefixes. In order to proceed with the discussion on the alternation of case marking in Pilagá, it is important to examine what kinds of semantic contrast arise when both sets of prefixes are possible.

17.3 Alternating case marking in Pilagá: situation types

In Pilagá, set A and set B can occur with the vast majority of verb stems. The case-marking variation shows functional similarities among all the constructions encoded by the same set of prefixes. Apart from some lexicalized idiosyncrasies, the generalization for this large class of verbs holds that set A aligns with induced and non-reflexive events, and in general with events where the participant exhibits intentionality when performing an action. Conversely, set B encodes resultative events, reflexive events, and, in general, events where the participant does not exhibit deliberation when performing an action.

17.3.1 ‘Induced’ vs. ‘spontaneous’ events and ‘resultative’ states

Two distinctions that set A vs. set B case marking reveals in Pilagá are (a) ‘induced’ vs. ‘spontaneous’ events (or what has been termed the ‘causative/inchoative’ opposition in the literature, cf. Haspelmath 1993, inter alia); and (b) ‘induced’ vs. ‘resultative states’. The connection between ‘spontaneous’ events and ‘resultative states’ will be clarified shortly.

2 ‘Event’ is a cover term for any verb class.
Both members of a causative/inchoative verb pair express the same basic situation (generally a change of state, more rarely an atelic process), and differ only in that the causative verb meaning includes an agent participant who causes the situation, whereas the inchoative verb meaning excludes a causing agent and presents the situation as occurring spontaneously.

DeLancey (1984) proposes that agentivity is a superordinate category that includes ‘causation’ (which in turn subsumes ‘volition’ and ‘proximate causation’ entailed by participating in an event with external consequences). Thus, a volitional transitive event (e.g. ‘X breaks the cup’) is characterized by a two-stage causation scheme; whereas a non-volitional event (e.g. ‘the cup broke’) will involve one causal chain. On the other hand, a volitional intransitive event (e.g. ‘I jump’) involves one causal vector ‘in which the act of volition causes the act’ (DeLancey 1984: 8), and a non-volitional intransitive event (e.g. ‘X died’) does not define any causal chain.

Croft (1994: 91–3) reconsiders the causation scheme elaborated by Delancey (1984) and restates it in somewhat different terms. He proposes that ‘causative’, ‘inchoative’, and ‘stative’ are part of an idealized ‘causal chain’ containing three separate segments, i.e., cause-become-state. He argues that, since simple events are endpoint-oriented, verbs may encode just the last segment (i.e. stative), or the second and the last (i.e. inchoative), or all three segments (i.e. causative).

We may apply this schema to the description of Pilagá case marking as follows. When an event is conceptualized from its starting point (or source: DeLancey 1990), this situation is encoded by means of set A case marking (recall our characterization of Agent in Pilagá as a performer or source). However, when the event is conceptualized as a change of a state (i.e. the second segment, according to Croft) or as the outcome of an event (i.e. the last segment in Croft’s model), the speaker chooses set B case marking.

In the third group of Pilagá verbs, ‘causative’ vs. ‘stative’, and ‘causative’ vs. ‘inchoative’ meanings correlate with alternating case marking, as shown in (13a, b) and (13c, d).

(13) a. s-ewat-eye
1A-open-dir.forward
‘I open (e.g. the door).’

b. jnewatayi
1B-open-compl
‘It (e.g. the door) is open.’

c. [yitiyiyat]
[yi-i-tfyia-t]
10-3A-tight-prog
‘They/it pinch(es) me.’

d. sa-tfyia-lo
1A-tight-objnum.pl
‘I tighten them.’

e. n-tfyia-atlat
3B-tight-refl
‘They shrank.’
It is, however, important to underscore that suffixes also play a role in shaping either causative, inchoative, or stative verbal meanings. For example, in (13), not all causatives or all statives receive the same type of verbal suffix; rather, it is the verb's lexical semantics that motivates the occurrences of the specific suffixes. 'Open' (13a, b), on the other hand, is marked with an aspectual suffix -(ta)yi 'completive' in the stative.

Nevertheless, it is important to note that for some verbs the causative/inchoative and the causative/stative alternations do no correlate with case marking changes. (14b) is an example where the stative is an adjectival form, with no case markers added; that it is not a verb is shown by the fact that the non-finite qaqata 'to be dry' is the same for 1st and 2nd person. The verb 'break' in (14c, d) shows a somewhat similar situation. Case marking does not make any contribution to the causative vs. inchoative distinction, since in both meanings the subject is grammatically categorized as an Agent. Here, what marks the difference between the causative and the inchoative construction is the aspectual marker -yi 'completive' in (14d). Conversely, note that in (14c) for the causative meaning the verb stem has a directional suffix -yi 'DIR.downwards'. However, for 'break' the difference between the causative and inchoative does correlate with case marking, as (14c, d) show. Case marking proves to differentiate the process vs. the state, as the addition of (14e) demonstrates:

\[
\begin{align*}
(14) & \quad \text{a. se-qayat} & \quad \text{b. hayim ø-qaqa-ta} \\
& 1A\text{-dry} & \quad \text{PRO.1SG 3A\text{-dry}-RES} \\
& 'I dry.' & \quad 'I am dry.' \\
& [dolatfiyi] & \quad [d-ola-t-iyi] \\
& 3A\text{-break-VFT-DIR.downwards} & \quad 3A\text{-break-COMPL} \\
& 'He breaks (something).' & \quad 'It broke.' \\
& \text{c.} & \quad \text{d.} \\
& [d-ola-t-iyi] & \quad \text{3A\text{-break}} \\
& 3A\text{-break} & \quad 'It is broken.'
\end{align*}
\]

Sometimes different stems are used for causative vs. inchoative meanings (this is what Haspelmath (1993: 92) calls 'suppletive' alternations for causative/inchoative verb pairs). Some examples are given in (15). (15a–c) show that the causative/inchoative reading for the concept 'lose' does not depend upon case marking. (15b) represents the accomplished process with a verb stem like som 'to get lost', while (15a) with the stem gem 'to lose' presents the event from the point of view of the entire process. But most important here is that in both examples the participant is marked by set A forms. (15c, d) constitute another instance where the inchoative and the causative concepts are not derived from the same stem, and where case marking also does not participate in coding the causative/inchoative semantic distinction.
(15) a. se-gem
    1A-lose(TR)
    'I lost something.'

b. se-somak
    1A-lose(INTR)
    'I got lost.'

c. yu-lat
    3A-kill
    'He kills/someone.'

d. so'ote yi-lew
    already 3A-die
    'He died already.'

17.3.2 Events with high vs. low intentionality

With a few verbs denoting mental and speech predicates, the choice of one set
vs. another set of prefixes may indicate the degree of intentionality of the subject
participant. Set B indicates that his/her intention is not deliberate, while set A
marking indicates it is more purposeful. See (16) for an illustration.

(16) a. ni-tʃo'ot
    1B-put forward/tell/present
    'I introduce myself.'

b. an-setʃo'ot
    20BJ-1A-put forward/tell/present
    'I introduce you.'

c. anʃo'ot
    2B-put forward/tell/present
    'Introduce yourself!'/You introduce yourself.'

d. anʃo'ot
    2A-put forward/tell/present
    'Tell something.'

e. jua-nom-ta qall'i nesejem
    1B-know-res 'yesterday he got up
    'I know he got up (I heard or saw him accidentally, I was there).'

f. sa-nom-a
    1A-know-obj num.s g
    'I have knowledge of something (i.e. a person or place because
    somebody told me).'

g. ni-lot-aqi?a
    1B-look-e p v o w -d i r . d o w n w a r d s
    'I look downwards.' (e.g. the subject is up in a tree)

h. selotjiyi
    [se-lot-iyi]
    1A-look-dir.in a straight line
    'I look at something/somebody insistently, without moving my eyes.'

i. naʃo'ot-tapiji
    1B-hide-dur
    'I am hiding (or pretending).'
As shown in (16a, b), the speech verb *tšo’ot* ‘warn’ can be associated with one (16a) or two arguments (16b). This is indicated by a different prefix choice plus the aggregate of an object participant marked on the verb (see *an- ‘2O’* in 16b). When no second human participant is ostensibly involved, the degree of intentionality on the part of the subject participant is still indicated by alternating prefixes from set A and set B (16c, d). The verb *noma ‘know’* (16e, f) also exhibits alternating case marking from both sets; but the aspectual suffix *-ta ‘resultative’, which derives activities into states, is also signalling that the event is seen from the point of view of its results. In such a case the subject participant is marked as an Undergoer (16e), unlike (16f), where case marking indicates a greater involvement. The last pair (16g, h) exemplifies the case of a verb *lot ‘look’* that can take either one (16g) or two arguments (16h) marked on the verb. In (16h), the object participant is zero marked, since it is 3rd person. (16g) indicates that the subject participant is not looking at a specific target or with any intention; whereas (16h) signals that he or she is looking at something or somebody on purpose.

17.3.3 Reflexive/non-reflexive events

The choice may convey a difference in ‘reflexive/non-reflexive’ meaning. A ‘reflexive’ event involves one participant, which stands in an Initiator/Endpoint relation to itself (Kemmer 1993: 52). To obtain the reflexive meaning, *-lP pat ‘reflexive’* may also be added to the root (cf. 17), though this is not necessarily the case. For some verbs, case marking alone suffices to yield the reflexive vs. non-reflexive interpretation (cf. 18). ‘Reciprocality’, indicated through *-at*, is another meaning associated with non-agentive in Pilagá.

The pairs in (17) and (18) contrast ‘reflexive/non-reflexive’, and also illustrate the reciprocal usage of set B marking with a reciprocal suffix *-at*, as opposed to the non-reciprocal action with two participants. (17a) is reflexive, (17b) non-reflexive, and both are marked with set B prefixes. According to the glosses (i.e. ‘look at myself’ in (17a) and ‘realize’ in (17b)), (17b) is also marked with a set B prefix even though non-reflexive. However, the sense is not ‘reflexive’ since, instead of a reflexive suffix, a directional marker shows up on the verb to derive the meaning of ‘to realize’ (lit. look forward). (17c) is a variation on (17a), based on a different grammatical person.

(17) a. *pi-lo-lP at*
   1b-look-refl
   ‘I look at myself (in the mirror).’

b. *pi-lo-lot-eye*
   1b-look-dir.forward
   ‘I realize’ (as if one’s mind opens)/‘I open my eyes (when I wake up).’

c. *an-lo-lP at*
   2b-look-refl
   ‘You look at yourself (in the mirror).’
The following pairs illustrate reflexive/non-reflexive and reciprocal/non-reciprocal constructions in Pilagá.

(18)  

a. se-kode-eye  
    1A-pour-DIR.forward  
    ‘I pour’ (a liquid, on a surface or on somebody).’

b. pi-kode-lat  
    [ji-kod-e-lat]  
    1B-pour-EPVOW-REFL  
    ‘I pour (liquid) on myself.’

c. am s-atlet  
    PRO.2SG 1A-offer  
    ‘I offer you something.’

d. n-atlet-a-lt  
    1B-offer-RECIP  
    ‘We offer (something) to each other.’

e. am sa-wana  
    PRO.2SG 1A-meet  
    ‘I meet you.’

f. n-a-wana-lat  
    1B-meet-RECIP  
    ‘We meet each other.’

g. y-alat  
    3A-kill  
    ‘He kills/ed (someone).’

h. y-alat-e-lat  
    3B-kill-EPVOW-REFL  
    ‘He killed himself.’

While some verbs require a specifically reflexive suffix to express the reflexive meaning, others convey this meaning by means of case marking only. In (19) I provide examples of reflexive/non-reflexive constructions, which are based solely on alternating case marking.

(19)  

a. pet?et  
    1B-fix  
    ‘I fix myself.’

b. set?et  
    1A-fix  
    ‘I fix him/her.’

c. ni?amen  
    3B-spread  
    ‘He/she makes up (e.g. as with facial makeup).’

(d. y-tamen  
    3A-spread  
    ‘He/she spreads something (e.g. grease).’

e. ji-yo  
    1B-clean  
    ‘I clean myself.’

f. si-yo  
    1A-clean  
    ‘I clean somebody/something.’

Most examples in (19) are body-care verbs. The lack of a reflexive marker in (19), as opposed to what we see in (18), supports the idea of Haiman (1983) regarding an economic motivation in language. Since the expectation for washing, shaving, and similar actions is that the person will perform the action upon him- or herself, such verbs do not require an over-reflexive object; this could also explain why grooming or body-care verbs do not require the reflexive marker in Pilagá either. A verb like ‘hide’, as in (19e–g), does not immediately fit into the category of body-care verbs, although it does denote an event where the body is the locus of the action.

One hypothesis that deserves further work is that the verbs in (18) differ from those in (19) in terms of the degree of transitivity involved. The first group of
predicates generally presupposes a second (human) participant and are conceivably higher in semantic transitivity (Hopper and Thompson 1980). Transitivity is downgraded in reflexive constructions (Hopper and Thompson 1980: 277). Thus, in (18), ‘reflexivity’ is less expected and, as such, it is ‘marked’ by choice of prefix set on the verb.

In the next section I will discuss one last property of the Pilagà case-marking system, the encoding of ‘trajectory of the event’.

17.3.4 Trajectory of the event: motion towards/away from a vantage point

For one group of stems that may take either set of prefixes, case marking describes opposed orientations of the spatial trajectory of the subject participant. The group consists of translational motion verbs involving motion to a different location, but also includes verbs that do not literally imply that the subject moves, such as ‘to buy’ / ‘to sell’, ‘to learn’ / ‘to teach’, and ‘to plough’ / ‘to harvest’. Thus, non-motion verbs of this group describe an imaginary trajectory relative to some point of reference.

3 Klein (1981) had proposed the existence in Toba of a similar parameter which she calls ‘directionality’, subdivided into ‘adcorporeality’ vs. ‘abcorporeality’. According to Klein, this parameter accounts for the distribution of subject prefixes for all verb stems. In a similar vein, Velázquez-Castillo (this volume) proposes a semantic interpretation of events marked as respectively active vs. inactive in terms of ‘centrifugal’ (directionality towards the event source) vs. non-directional or ‘centripetal’ events.
Note that in (20a–j) the stem form is clearly irrelevant for the choice of the prefix and that motion verbs (20a–j) carry a directional suffix, independently of the fact that the case marking indicates ‘trajectory’ relative to a point of reference. Very importantly, the suffix is also coding path or trajectory. The point of reference coincides with the speaker location in (20a, b) and (20c, d). The gloss in (20e) suggests that the point of reference is the speaker’s location as the subject participant moves within that space. Unlike (20e), in (20f) the speaker moves towards that point of reference, from the starting point, in an upwards direction, as indicated by the directional suffix. In (20i, j) the ‘coming’ or ‘going’ can only be predicated from the stance of the speaker who sees the event participant approaching towards or away from him/her.

Certain motion verbs have co-lexicalized together with a directional suffix, to the extent that their basic stem form now contains a directional suffix. When the form for the verb ‘bring’ or ‘take’ is requested in direct elicitation, speakers will provide the expressions transcribed in (20g, h); but the same root can also combine with other directional suffixes (e.g. sedoyet ‘I receive’, sedoyi ‘I carry something from the inside to the outside’).

Now, the non-motion verbs provided in (21) do not have directional markers attached, though valence suffixes may be relevant to the meaning of these constructions. ‘Sell’/’buy’ exhibit different stems, although the distribution of set A/set B case marking for ‘sell/buy’ respectively holds. In the case of ‘teach/learn’ and ‘plough/harvest’ the verbs differ in case marking, the stem being the same. In these examples, case marking indicates that a trajectory is implied; the only participant is affected when the event is done in his/her own interest, and thus set B case marking occurs (even though the affected participant may also be an agent, as in ‘buy’ or ‘cultivate’). When the subject participant is a performer and a second participant is involved as the affected Undergoer, set A forms occur.

(21) a. ña-pa‘aγen
   1b-learn
   ‘I learn.’

   b. sa-pa‘aγen-t-ʔaγan
       1a-learn-VFF-VAL
       ‘I teach.’

c. ña-poyaq
   1b-cultivate
   ‘I cultivate or sow for myself (i.e. my own field).’

   d. sapoγaʔan
      [sa-poyaq-ʔaʔan]
      1a-cultivate-VAL
      ‘I harvest (i.e. for somebody else).’

e. n-men-ʔaʔ
   3b-give-RECIP
   ‘They give each other (exchange) things.’

   f. de-men-ʔaʔan
      3a-give-VAL
      ‘He/she sells.’

g. hasoʔ perentanaʔae yi-ede-ye
   f.clf teacher 3a-write/carve-LOC.on
   ‘The teacher writes (e.g. something on the blackboard).’
17.4 Discussion: case marking and voice

It has been suggested for Toba (which is Pilagá’s closest relative within the southern branch of the Guaykuruan family) that the cognates of what I have named ‘set A’ and ‘set B’ are part of a voice system and respectively encode the distinction ‘active’ vs. ‘middle’ (Censabella 1997). Censabella analyses all instances of the two sets as marking active vs. middle voice. Her interpretation of middle voice in Toba is based on Kemmer’s (1993) survey of middle voice languages. For Kemmer, middle voice is tied to ‘reflexivity’ (in body action, body posture, and grooming verbs) and ‘reciprocality’, but also to other middle situation types such as spontaneous events (with lack of volitional initiation), passive-middle events (where the external causer is pragmatically de-emphasized: cf. Kemmer 1993: 147), and mental processes, which naturally exhibit low control, will, or instigation. In sum, middle events comprise situations where the initiator is also an endpoint, or an affected entity, or where the event is characterized by a lower degree of participant elaboration (Kemmer 1993: 243). All these functions are part of what set B case marking does in Pilagá. However, in Pilagá, the expression of ‘reflexivity’ is also a function of the reflexive suffix. In those cases where the verb does not necessitate a reflexive marker (i.e. as in 18), set B constructions are semantically, though not structurally, reflexive.

The distribution of case marking in Pilagá demonstrates that most stems of the set A-only group are activities, with or without control. In such cases, the subject participant is mapped onto the set A case marking since he or she is perceived as non-affected. (The grammatical changes that have caused certain states or positional verbs to became part of this group remain synchronically obscure.) The second group, comprising stems which are exclusively set-B marked, is comparatively smaller and also has some exceptions to the core meaning of the group. These verbs are most usually non-actions where the only participant is grammaticized as if it were an Undergoer (a ‘controller’, to use the terminology of Klaiman 1991). The group of verbs that may take either case is large, and pattern semantically with middle meanings, which makes this particular system look voice-oriented.

However, the array of meanings that set B case marking encodes covers other senses beyond what have been characterized as typical middle functions (in particular, ‘trajectory of the event’ with motion-cum-translation and non-motion verbs (a feature that to my knowledge has not been documented for agentive or active systems so far). Furthermore, the distribution of case marking for the first two groups (i.e set A-only and set B-only) is not entirely

h. *hayem n-ede-leye*
   
   PRO:ISG 1B-write/carve-LOC.on
   
   ‘I sign (on a piece of paper).’
predicted. In my analysis, case marking overlaps with middle meanings, but ‘middle’ is a by-product of the set B prefix class in Pilagá. Based on the description provided in the preceding sections, I argue that the Pilagá system is generally organized on the basis of ‘agentivity’, where the set B paradigm semantically includes some middle meanings, such as ‘reflexivity’ and ‘spontaneous events’.

### 17.5 Conclusion

The general characteristics of the Pilagá person-marking system is that choices between sets are based on whether the subject participant is the agentive doer of an activity verb or an affected entity. In the final analysis, however, set A and set B are best regarded as formal labels used to designate each prefix class, which need not imply that the subject argument position is filled by a participant who is invariably perceived as either the semantic Agent or the semantic Undergoer of an event or state. The semantic case roles of Agent and Undergoer are nevertheless necessary in order to account for the fact that many verb stems can alternatively encode the grammatical subject by both sets, leading to different conceptualizations of the same event.

I would like to propose the existence of two semantic parameters in connection with this system, which ultimately seem to trigger the distribution of case marking:

(a) **Viewpoint.** I argue that in Pilagá an event is viewed from the stance of its initial viewpoint or from its endpoint. Pilagá appears to be sensitive to a model of causation (as proposed by DeLancey 1981 and reinterpreted by Croft 1994). This means that ‘causative’, ‘inchoative’, and ‘stative’ indicate different elaboration types of event, and this elaboration exhibits some correlation with pronominal case marking in this language. Similar systems have been documented in the literature. Merlan (1985: 351) notes that in Arikara (Caddoan), inchoativization is marked by a shift in the prefix inflectional class. However, as I said, this correlation is not 100 per cent consistent too, which implies that sometimes the distinction ‘causative’ vs. ‘inchoative’ and/or ‘stative’ does not align with alternating set A and set B in Pilagá.

‘Viewpoint’ also accounts for the fact that pronominal case marking encodes ‘trajectory of the event’, by which participants ‘move’, either in a physical or a metaphorical sense, towards a specific point of reference which may or may not coincide with the speaker’s location.

---

4 The proposal of such an explanatory parameter for semantic alignment systems is not new (cf. DeLancey 1981, which constitutes the first work suggesting a connection between ‘viewpoint’, agent/patient categories, and semantic alignment).
(b) Affectedness. When the event is conceptualized from its endpoint, the subject participant is affected. ‘Affectedness’ is a crucial parameter in voice systems, and this is the reason why ‘middle’ meanings can be obtained from set B case marking. An affected participant furthermore lacks intentionality, volition, or purpose. Some verbs denoting mental actions and states may convey such distinctions by set B forms, although the system also exhibits pervasive leticalization, which results in the case marking by set A forms of certain mental predicates.
References


432 References


(1858) *Versuch einer jenissej-ostjakischen und kottischen Sprachlehre*. St Petersburg.


References


References


References


References


References


GREKOFF, GEORGE (n.d.). ‘Notes on Chimariko, based on the fieldnotes of John Peabody Harrington’. MS, University of California, Berkeley: Survey of California and Other Indian Languages.


References


References


KAVITSKAYA, DARYA (in prep.). 'Accentuation in Western Tundra Nenets'.


References


442 References


References


——— (2004). ‘A grammar of Creek (Muskogee)’. MS, College of William and Mary.


References


References


References


References


References

Tereshchenko, Natalja Mitrofanovna (1956). Materials and Studies on Nenets [Materialy i issledovania po jazyku nencev]. Moscow: AN SSSR.


References


452 References


