Problems of Polysynthesis

Studia typologica

EDITED BY THOMAS STOLZ

Beihefte / *Supplements*

STUF – Sprachtypologie und Universalienforschung
Language Typology and Universals

VOLUME 4

Problems of Polysynthesis

Edited by Nicholas Evans and Hans-Jürgen Sasse

Akademie Verlag

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Problems of polysynthesis /
ed. by Nicholas Evans and Hans-Jürgen Sasse.
– Berlin : Akad. Verl., 2002
(Studia typologica ; Vol. 4)
ISBN 3-05-003732-6

ISSN 1617-2957

Druck und Bindung: GAM Media GmbH, Berlin

Printed in the Federal Republic of Germany

Contents

Introduction: problems of polysynthesis

NICHOLAS EVANS & HANS-JÜRGEN SASSE[1]

1. The challenge of polysynthetic languages

The papers assembled in this volume deal with a range of questions raised for linguistic theory and description by polysynthetic languages. Prototypical polysynthetic languages, found among unrelated language families in such varied parts of the world as North America, Meso-America, Siberia, northern Australia (pace FORTESCUE 1994), and Papua New Guinea, display remarkably similar suites of grammatical characteristics. But, nearly two centuries after HUMBOLDT (1836, especially chapter 17, on Aztec) and KLEINSCHMIDT (1851, on Eskimo) began to make the existence and interest of polysynthetic languages widely known among linguists, languages of this type continue to pose a challenge to every major linguistic theory.

The remarkably similar structures of two unrelated languages from opposite ends of the earth can be used to illustrate the workings of a typical polysynthetic language. Consider the following sentences from Bininj Gun-wok (Gunwinyguan, Northern Australia) and Cayuga (Northern Iroquoian, Great Lakes Region of North America):

---

1 We would like to acknowledge the following people and organizations: the DFG for financially supporting the 1998 workshop, the Alexander von Humboldt-Stiftung for supporting Evans' stay at the Institut für Sprachwissenschaft in Cologne, and the School of Languages, University of Melbourne for financial assistance to Evans for his part of the editing, Brunhild Sieck for organizing the 1998 workshop, the members of the Polysynthesis project at Cologne, in particular Claudia Leto, Johanna Mattissen, and Christoph Benden for their invaluable help with all editorial activities, including the tedious tasks of proof-reading, formatting, correspondence with contributors and publishers, etc., and to the participants in the 1998 workshop for their comments and discussion on the papers presented there. Finally we thank Martin Haspelmath and Thomas Stolz for negotiating the publication of the current volume in the series Studia Typologica.

(1) *Nga-ban-marne-yawoih-dulk-djobge-ng.*
BGW 1sgSUBJ-3plOBJ-BEN-again-tree-cut-PP
'I cut (the) tree/wood for them again.'
'I cut another tree for them.'

(2) *Ę-s-kakhe-hǫna't-á-yęthw-ahs.*
CAY FUT-REP-1sgSUBJ/3plOBJ-potato-SJ-plant-BEN:PUNC
'I will plant potatoes for them again.'

Both these examples illustrate the way in which polysynthetic languages represent, in a single verbal word, what in English takes an entire multi-word clause. They achieve this by using pronominal affixes for some arguments (subject and indirect object in this case), and incorporated nouns for others (object in this case). Some of the arguments are directly subcategorized for by the root (here subject 'I' and objects 'tree/wood' and 'potatoes' respectively), while others are introduced by 'applicative' affixes – here the benefactive affix which, in both languages, adds a beneficiary which gets encoded by pronominal affix in the object slot. In addition, both languages have at their disposal a range of adverbial-type affixes to the verb, which we exemplify here with affixes encoding repetition.

Continuing the theme of how verbal morphology is used to alter argument structure, all of which gets represented on the verbal word, consider two further examples in each language, illustrating the formation of one-word reflexive and reciprocal clauses.

(3) *Nga-bid-djobge-rr-inj.*
BGW 1sgSUBJ-hand-cut-RR-PP
'I cut myself on the hand.'

(4) *A-k-atat-wę'nahs-á-ik.*
CAY FAC-1sgSUBJ-REFL-tongue-SJ-bite:PUNC
'I bit my tongue.' (Lit.: 'I bit myself in the tongue.')

(5) *Barri- marne- dulk-djobge-rr-inj.*
BGW 3plSUBJ:PST-BEN-tree-cut-RR-PP
'They cut tree(s) for each other.'

(6) a. *T-ę-hęn-atat-hǫna't-a-yę́:thw-ahs.*
CAY DU-FUT-3plMASC:SUBJ-REFL-potato-SJ-plant-BEN:PUNC
'They will plant potatoes for each other.'

cf. b. *Ę-hęn-atat-hǫna't-a-yę́:thw-ahs.*
FUT-3plMASC:SUBJ-REFL-potato-SJ-plant-BEN:PUNC
'They will plant potatoes for themselves.'

In (3) and (4), the transitive verb is reflexivized by an affix, creating a one-place verb taking a subject prefix only; an incorporated body-part noun marks the locus of the self-directed action. In (5) and (6a), the subject and derived object both instigate and benefit from a reciprocally-directed action; in both languages this is represented syntactically by suppressing the pronominal slot that would normally express an argument added by a benefactive applicative, adding an affixal marker of the reciprocal, and continuing to allow the encoding of the direct object by incorporated nominal. A typological difference between the two languages here is that in BGW the reflexive and reciprocal affixes are formally and positionally identical, whereas in Cayuga distinct formal techniques are used: the reflexive prefix (cf. (6b)) takes on a reciprocal reading only in combination with the "dualic" prefix. A further difference is that in BGW the reflexive readings (unlike the reciprocal readings) are only available with underived verbs, whereas in Cayuga verbal applicatives may feed the reflexive.

Morphologically complex as these examples are, they do not exhaust the degree to which polysynthetic languages can concentrate the expression of information into the verbal word. Some polysynthetic languages may include several verb roots in the same word, have affixes indicating the shape, spatial layout or distribution of actants, quantification, instrument by which the action is effected (e.g. by the teeth), or conjunctional linkage with other clauses or words, in addition to the many inflectional and derivational categories found in languages with less exuberant morphology. We will add a final example, from West Greenlandic (FORTESCUE 1983: 97), to illustrate, inter alia, morphological integration of adjectival concepts (*-sua(q)-* 'great'), copula, embedding (*-nerar-* 'say that'), affixes expressing speaker's attitude (*-galuar-* 'sure but'), and interclausal relations (*-li* 'but'):

(7)	*Aliiku-sersu-i-llammas-sua-a-nerar-ta-ssa-galuar-paal-li, ...*
WG	entertainment-provide-SEMITRANS-one:good:at-COP-say:that-REP-FUT-sure:but-3plSUBJ/3sgOBJ-but
	'However, they will say that he is a great entertainer, but ...'

Essentially, then, a prototypical polysynthetic language is one in which it is possible, in a single word, to use processes of morphological composition to encode information about both the predicate and all its arguments, for all major clause types (i.e. one-, two- and three-place predicates, basic and derived), to a level of specificity allowing this word to serve alone as a free-standing utterance without reliance on context.

We formulate our definition in this way so as to leave open certain key questions:

(a) should a language be considered polysynthetic if it has multiple agreement but no incorporation, as in Abaza (triple agreement) or Georgian (double agreement)?

(b) should a language be considered polysynthetic if it includes constructions in which certain argument positions are not overtly filled, e.g. certain non-finite constructions?

(c) should a language be considered polysynthetic if certain argument positions require the unification of information from inside and outside the verbal word, e.g. through constructional linking of an affix with an external nominal or noun phrase?

(d) should a language be considered polysynthetic if it exhibits extraordinarily rich combinatory processes in the verbal word (noun incorporation, integration of adverbial, instrumental, spatial, etc. material) but no agreement affixes? Haisla is an example of such a language.

Here we differ from some other investigators, notably BAKER (1996:17-19), who defines polysynthetic in such a way as to answer each of the above questions in the negative.

Our reason for adopting a less restrictive definition than Baker is the following. 'Polysynthetic' is a useful label to the extent that it captures recurrent clusterings of typological characteristics, probably reflecting what NETTLE (1999) calls 'local optima', in which combinations of logically independent structural parameters which happen to work well together form 'basins of attraction' in the complex state space given by the set of all logically possible parameter combinations – they undergo functional co-selection, along lines spelled out by HEATH (1975). An alternate hypothesis, put forward in BAKER's (1996:19-20) Principles and Parameters approach is that they automatically flow from the setting of a single 'macroparameter', which he terms the 'morphological visibility condition'.

Now such a clustering, if indeed it exists, will be comparable in importance for typology to the suite of characteristics found in SOV languages, for example, and likewise call out for a unified explanation. But the empirical question of whether these various characteristics necessarily go together can only be tested by looking at a broader sample of languages that allows the possibility that each parameter is in fact independent. Otherwise the clustering of properties is simply circular, the result of definitionally excluding from consideration any language that fails to meet at least one of the criteria.

The languages discussed in this volume, then, have been selected according to this somewhat more open heuristic definition. Although most of them would be considered polysynthetic by most definitions (Cayuga, Eskimo, Apache and Aztec in the Americas, Bininj Gun-wok, Ngalakgan and Rembarrnga in Australia, and probably Ket in Siberia), we have included two further languages – Georgian and Nivkh – where their position is less clear, in the case of Georgian because of the existence of alternate strategies in which external argument encoding must be used in certain circumstances, and in the case of Nivkh because, although it exhibits many parallels with canonical polysynthetic languages in terms of the complexity of its morphology, the 'dependent-head' synthesis technique it employs does not include the representation of subjects on the verb.

Moreover, by including papers from a wide variety of geographical areas we try to compensate for the distorting effects of concentration on Amerindian languages that characterized prior work on polysynthesis though some areas with polysynthetic languages remain unrepresented in this volume, such as some of the Munda langugages in India, and a number of Papuan languages.

Papers in this volume also raise a number of further issues that are seldom touched upon in the discussion of polysynthesis. These include the different linguistic levels of analysis or representation and their intersection, which in turn give rise to a variety of angles from which polysynthesis can be approached: syntax, morphology, lexicon, etc. It is easy to see that polysynthetic languages constitute a challenge to traditional views on the interplay between morphology and syntax. These views proceed from a European-type notion of sentence, conceived of as a multi-word string whose constituent parts are made up of separate word forms that represent syntactic categories (subjects, objects, attributes, predicates, etc.). Since Humboldt's groundbreaking work on morphological typology in the early 19th century, the standard task of inflectional morphology has always been seen as the marking of syntactic functions; i.e. it is viewed as a grammatical device subsidiary to syntax. However, here we are dealing with languages where basic clause structure is in the morphology itself. How does the morphosyntactic structure of such languages relate to the traditional concept of language levels and to what extent can the problems raised by these languages be handled in the frameworks of traditional morphological and syntactic theories, which rest upon this concept?

Traditional ideas about the division of labor between linguistic components or 'modules' are likewise called into question when polysynthetic languages are looked upon from the angle of the lexicon. On the classical view, the lexical unit of a language is held to be the lexeme, an abstract unit constituted by an inflectional paradigm of word forms based on a root or stem (which may consist of only one form in the case of an isolating language) plus a range of related word senses. Lexemes are usually thought to be inserted, in their appropriate word forms, into the slots opened by syntactic structure. Polysynthetic languages often resist such a treatment. First, it is often difficult to define a pattern of phrase structure, into which lexical insertion could take place. Second, the enormous number of grammatical categories may make it difficult to set up 'paradigms' in the same sense as is done in languages with less spectacular morphology. Third, the tremendous number of combinatory possibilities of morphological material often leads to the lexicalization of single forms of the paradigm or entire subparadigms, giving rise to ambiguities between compositional and non-compositional readings, fuzzy boundaries between productive word-sentences and 'lexemes', etc. In many languages, therefore, problems of determining the lexical unit arise, especially when the languages are highly fusional and roots or stems are difficult to isolate. Further problems of the interaction between lexicon and syntax also arise in the realm of incorporation: incorporating languages differ considerably in their language-specific

constraints on incorporability and hence, in their degree of lexicalization of incorporation. As a result, the polysynthetic lexicon exposes, in a particularly acute form, the gap between the theoretically composable and the idiomatic, a problem that is found in all languages (cf. PAWLEY & SYDER 1983) but which, in non-polysynthetic languages, does not need to be solved before a dictionary can be made.

2. Background to the present volume

Most of the papers contained in this volume were first presented at a workshop entitled 'Problems of Polysynthesis', organized by the present editors and hosted by the Institut für Sprachwissenschaft at the University of Cologne in October 1998, with the support of the Deutsche Forschungsgemeinschaft. At that conference we sought to bring together a number of scholars involved in research on a range of topics concerning polysynthesis. A selection of papers from that conference, originally planned as a guest-edited volume of STUF, appeared through a series of misunderstandings as uncoordinated free-standing papers in the combined third and fourth numbers of the 1999 issue of that journal. Subsequent discussions with the STUF editorial board have enabled us to return to the originally envisaged format in the form of the current volume. This has given us the opportunity to add two further papers from the 1998 conference, those by Dagmar Jung on Apache and by Werner Drossard on Ket, as well as two invited articles, by Rachel Nordlinger & Adam Saulwick on Rembarrnga, and Brett Baker on Ngalakgan, which deal with topics inadequately covered at the original conference. Naturally it has also given the original authors the opportunity to correct and update their earlier papers where necessary. Drossard's paper also incorporates recent research done in the framework of a project on polysynthesis currently carried through at the Institut für Sprachwissenschaft at the University of Cologne and funded by the Deutsche Forschungsgemeinschaft.

3. Overview of papers

The first three papers deal with issues of argument representation in polysynthetic languages, focussing on the related questions of whether verbal slots are the only locus of argument encoding, of what is the semantic status of pronominal markers on verbs, and how pronominal affixes relate to external material.

NICHOLAS EVANS, focussing on Bininj Gun-wok (Gunwinyguan, North Australia) takes on the question of whether bound pronominal affixes really are equivalent to pronouns, acting as the true exponents of the verb's arguments, with external full NPs merely functioning as some kind of adjunct or modifier; this position is well-known through the

work of such authors as JELINEK (1984) and BAKER (1996). In Bininj Gun-wok, Evans shows, bound object pronominal affixes are insensitive to referentiality and definiteness properties of the object, in the same way that subject agreement in German or Latin is insensitive to referentiality and definiteness; this follows from the obligatory nature of such inflections, which deprives them of a definiteness-signalling function (unless a further, indefinite object affix is also present as in Nahuatl and some other languages). Since such insensitivity to discourse status is not a feature of third person personal pronouns it is inappropriate to treat these affixes as the sole pronominal exponents of a verb's arguments – verbs bearing them occupy an intermediate status, in terms of specification of argument properties, between a verb with no argument information (e.g. an English simple past verb) and a verb combined with free pronouns for all its arguments. A number of other polysynthetic languages which appear to display similar characteristics are surveyed more briefly. The need to unify material from external NPs with pronominal affixes in order to produce expressions that are sensitive to referentiality and definiteness means that analysts will get a distorted view of how polysynthetic languages work if they just focus on the spectacularly self-sufficient verbal words seen so often in treatments of polysynthetic languages.

BRETT BAKER's paper deals with another Gunwinyguan language, Ngalakgan, which while paralleling Bininj Gun-wok in many respects introduces a further complication because of the availability of noun-class markers for subject and object on the verb, encoding 'vegetable' and 'neuter' classes corresponding to eponymous categories in the morphology of external nominals and adjectives. Baker shows that there is systematic 'disagreement' between the choice of verbal pronominal prefixes and the corresponding noun classes on external nominals, rather reminiscent of the way noun classifiers in many languages are varied to give different construals of a single entity. For example, in Ngalakgan the vegetable-class noun *mu-jukgul* 'acacia holosericea' will be represented by a vegetable-class verbal pronominal prefix when it is being construed as a source of fish poison or soap, but by a neuter-class prefix when it is being construed as a source of medicine, spears or firewood, or simply as a tree. He concludes that, while first and second person prefixes may plausibly be regarded as the real arguments of the clause, third person prefixes present a more complex picture; he proposes the term 'pronominal generic affix' for the function of the third person non-human prefixes just discussed.

Georgian, with its 1,500 year old written tradition, is probably the language with the longest documented history as a polysynthetic language, though as WINFRIED BOEDER points out in his paper, its argument agreement characteristics have been remarkably stable throughout this time. Boeder's paper focusses on the intricate interrelationships between verbal encoding and encoding by external NPs, made particularly complicated by the need to distinguish two types of morphological non-appearance on the verb. On the one hand there are arguments which are underlyingly present in the verbal morphology but are suppressed from appearing in particular combinations (e.g. the suppression of second person singular subject marking when there is a first person object marker), while on the other hand

there is the total 'disposing of' object encoding in reflexive constructions. Boeder shows how the difference between the two situations can be demonstrated by whether apposition of a NP in the appropriate case is possible: it is possible in the case of 'suppression' but not in the case of reflexives. In his conclusion, Boeder reminds us that the Jelinek / Baker 'pronominal argument' position continues an old tradition of taking the 'the finite verb form as a clause in nuce which can be expanded by adding "appositional" subject and object noun phrases', a position which he finds basically compatible with the workings of Georgian, but nonetheless suggests that other interdependencies between morphology and syntax must be considered.

The next two papers, by Michel Launey on Classical Nahuatl (Aztec) and Johanna Mattissen on Nivkh, examine functional parallels between noun phrases and higher-level units – or, to view it the other way, they show how the richly-coded structures at the level of the clausal predicate are exactly parallelled in the structure of noun phrases. The relevant higher-level units, however, differ in these two papers: it is the verb-clause (i.e. the predicate with all its bound arguments, including the subject) in the case of Classical Nahuatl, but it is the VP (i.e. verb plus object, but not subject) in the case of Nivkh.

LAUNEY's paper makes available in English the view of Nahuatl structure advanced in his important 1994 book *Une grammaire omniprédicative*: that even though nominals and verbals are clearly distinct on morphological grounds, both normally function directly as predicates in the language. Both must be preceded by the determiner *in* before being used as arguments, so that a sentence like *ni-k-itta in kone:-tl* 'I see the baby' (see (1b) of his article) is best understood, structurally, as 'I see him/her, who is a baby'.[2] In the present paper he extends this analysis to take in striking syntactic and semantic parallels between verbs with incorporated nominals and compound nominal structures. Despite this, he points out two important differences: (a) while the head element of such multi-morphemic words can be either a verb (in the case of noun incorporation) or a noun (in the case of compounding), the dependent element can only be a noun, (b) the parallels that are clear in the case of 'absolute' (unpossessed) nouns become less clear once the noun is possessed, reflecting the complex semantics of the possessive argument place.

MATTISSEN's article deals with a typological option normally omitted from considerations of polysynthesis, but well-represented in north-East Asia. Whereas canonical polysynthetic languages represent all arguments on the verb – without regard to whether they are subject or object – Nivkh constructs morphologically complex words by synthesizing head and dependent, construed here to mean the object of a verb but not its subject. Far-reaching internal sandhi rules make the identification of the phonological word in Nivkh relatively straightforward, and the units so defined can be extremely complex, e.g. 'salmon-catches', 'teach our younger sibling!' or 'saw a man building a boat'. Their

2 Note in passing that this is strikingly similar to JELINEK's (1995) analysis of Straits Salish, except that in the case of Aztec, unlike in Straits Salish, there are clear morphological grounds for distinguishing nominal and verbal word classes.

complexity notwithstanding, they are composed according to quite a different principle to that found in standard polysynthetic languages, which Mattissen labels 'head-dependent synthesis': heads of NPs can compound with their dependents (e.g. 'my-long-board', 'those-seven-people'), and verbs can compound with their objects ('people-even-kills') or other complements like subcategorized locatives (e.g. 'needle-inserts'). Unlike with noun incorporation or pronominal affixation in 'normal' polysynthetic languages this process can involve even question words (e.g. 'whom-shot?') or proper names like Xor (Xor-hit-person for 'the man who hit Xor'). However, it cannot extend to subjects – subject pronouns must be free words[3]. After surveying the Nivkh system in detail, Mattissen goes on to look at structural parallels in two other unrelated languages of north-east Asia, Ainu and Chukchi, which she argues exhibit similar structures (though Chukchi, as is well-known, also exhibits subject agreement so in this case we may be dealing with a language that is BOTH polysynthetic and exhibits head-dependent synthesis).

The next two papers, by Dagmar Jung on the Athapaskan language Apache and Rachel Nordlinger and Adam Saulwick on the Australian language Rembarrnga, deal with the issue of interclausal relations, in particular with the question of how polysynthetic languages represent subordination, and whether they permit non-finite constructions. Various authors have noted the lack or weak development of subordinate constructions in polysynthetic languages, especially the paucity or total absence of non-finite constructions. Some (e.g. HEATH 1975, MITHUN 1984) have advanced functional explanations for this, while others (most notably MARK BAKER 1996) claim that the lack of infinitives derives directly from the setting of a macro-parameter, the 'morphological visibility condition', requiring the expression of all arguments on the verb in polysynthetic languages.

JUNG's paper focusses on Apache, which is typical of a head-marking language lacking non-finite constructions (whether it would also be considered polysynthetic depends on whether noun incorporation is taken as a necessary criterion, but certainly it has obligatory representation of arguments on the verb through pronominal affixation). She outlines the structural alternatives to non-finite constructions, which include (a) the direct postposition of locational markers such as the ablative and directional markers to inflected verbs, (b) the use of a 'nominalizing' suffix, roughly translatable as 'the one (who/which)', to convert inflected verbs, simplex clauses and even complex clauses into referring expressions functioning as relative clauses or lexicalized entity descriptions, (c) the use of a linker *-go*, again postposed to fully inflected verbs, to indicate simultaneity and a range of other inter-clausal relationships such as conditionality, chaining of events in a paragraph, habituality (with the verb *ghai* 'years pass' functioning as a quasi-auxiliary in this construction), (d) analytic causatives, (e) complements of desiderative and cognitive verbs, (f) resultatives and

3 The possibility of compounding nominal dependents with heads means that one can form single-word expressions like [boat-make-progressive-person-see] 'saw a man building a boat' which look like subject incorporation. But here the relation of 'man' to 'build' is of participle to head rather than subject to finite verb.

complements of perception verbs. Although most of the constructions discussed arguably involve subordination, none involve any restriction on finite verbal categories of subject/object agreement or aspect, making Apache a clear example of a head-marking language with only finite constructions.

Rembarrnga, discussed in the next chapter, is quite a different case. Though canonically polysynthetic in its structure (it is related to Bininj Gun-wok and Ngalakgan), it has been reported as having infinitive constructions. Its status as a possible counter-example to the claim that polysynthetic languages cannot have non-finite constructions was already discussed by MARK BAKER (1996:475-6), employing data from GRAHAM MCKAY's (1975) Ph.D. thesis. Baker attempts to dispose of McKay's examples of putative infinitives by treating them either as nominalized verbs or as characteristic of an aberrant dialect. Given the inconclusive nature of Baker's counter-proposals, Rembarrnga merits a closer look on the basis of more extended material, and SAULWICK AND NORDLINGER return to this question, adducing new data and discussing various alternative definitions of 'infinitive', both in terms of neutralized TAM distinctions and in terms of suppressed argument agreement, as well as considering the complex issue of homophony between some TAM allomorphs and exponents of the infinitive. Their fine-grained analysis shows that Rembarrnga must be analysed as having two alternative infinitive constructions (one with argument exponence on the verb and one without) despite its canonically polysynthetic structure. Note in passing that Rembarrnga is by no means the only polysynthetic language to do this: Chukchi, for example, has both participles and gerunds, neither of which agree for subject or object.[4]

Given that polysynthetic languages push to the limits the boundary of the grammatical word, it is natural to ask whether they also place strains on the identification of the lexicological word: where words express clauses and even complex sentences, what are the relevant units in the speaker's mental lexicon? This is also a crucial question for those engaged in the lexicography of polysynthetic languages. The next two papers, by Hans-Jürgen Sasse and Werner Drossard, focus on the organization of the lexicon in two polysynthetic languages, respectively Cayuga (Iroquoian) and Ket (Yeniseyan, Central Siberia).

Polysynthetic word-sentences from Cayuga, the focus of SASSE's paper, have already been exemplified in the opening part of this introduction. Examples like (2) do not illustrate clearly, however, the degree to which particular combinations of morphemes in polysynthetic words are productive and compositional. When we selected example sentences for this introduction we wanted to keep their content as parallel as possible, but it turned out that the most direct analogue of the Bininj Gun-wok sentence in (1), with 'tree'

4 See COMRIE (1981:249, 252) for some accessible examples, and BAKER (1996:476-484) for an attempt to account for most of these through a process of 'excorporation' which leaves behind fragments of complex heads, though Baker admits there remain examples he cannot account for in this way, which he suggests result from Russian influence.

incorporated into 'cut', would not actually have an incorporating equivalent in Cayuga, reflecting the fact that languages differ considerably in the constraints they impose on incorporability.

In Northern Iroquoian languages, object-denoting expressions can only be incorporated if they derive from a certain type of root, usually called 'nominal roots'. Otherwise they have to be derivationally 'prepared' by adding one of a specific set of affixes. But this type of derivation is not available to all non-nominal roots, with the result that certain items are simply not incorporable. Moreover, verbs come in two classes, those that do incorporate and those that don't, and even those that do may be subject to lexical restrictions, morphophonemic idiosyncrasies and so forth. In addition, various degrees of semantic lexicalization come into play: in Northern Iroquoian languages N + V compounds very quickly enter the lexicon and become conventionalized, with the effect that ambiguities between compositional and non-compositional readings arise, or the compositional reading gets blocked (on this issue see also MITHUN & CORBETT 1999). Sasse's paper outlines the challenges these problems pose for Cayuga lexicography and lexicology, discussing the difficulties that arise from more morpheme-centred approaches and the various reasons for adopting larger multi-morpheme units, with some speakers going so far as to advocate the full polysynthetic word as the unit for organizing lexicographic entries.

DROSSARD's paper begins by establishing Ket's status as a polysynthetic language, exhibiting subject and object agreement, noun incorporation and numerous adverbial-type affixes to the verb. Ket is particularly interesting for typological studies of polysynthesis in the light of its isolate status and the fact that it does not share its polysynthetic characteristics with any languages of the surrounding area (discounting the other members of the Yeniseyan family, such as Kot, which become extinct over the last two centuries). The main focus of Drossard's paper is on the system of multiple verb stems by which verb lexemes are composed, reminiscent of the complex verb constructions found in such (non-polysynthetic) languages as Jaminjung in Australia (SCHULTZE-BERNDT 2000) and Kalam in Papua New Guinea (PAWLEY 1993). But in Ket, unlike in Jaminjung or Kalam, these multiple verb stems co-occur inside the one polysynthetic word, with varying degrees of compositionality and productivity. Drossard discusses some of the analytic problems that arise from the fuzzy boundaries between the verb stem class and the set of incorporated noun roots, on the one hand, and adverbial affixes on the other. The tentative nature of some of the analyses underlines the urgent need for detailed analysis of the lexicon of this highly aberrant language.

The volume closes with the only paper dealing with diachronic issues. Eskimo was one of the earliest-described and is still perhaps the morphologically richest polysynthetic language in terms of what a single verbal word can express. MICHAEL FORTESCUE examines both the reduction in morphological complexity that has occurred within the Aleut branch of the Eskimo-Aleut family, but also – moving now beyond what the comparative method can currently achieve, and relying on typological comparisons with Ural-Altaic languages in northern Eurasia, such as Finnish and Nganasan – a more speculative scenario as to how the

extreme polysynthetic structures exemplified in Eskimo (and likely to have been present in proto-Eskimo-Aleut) could have arisen. Fortescue argues for a number of interrelated changes, including the development of incorporation with originally free verbs for 'get' or 'have' which would have typically taken indefinite objects, the subsequent incorporation of verb-extending and negative auxiliaries into the verb, and the reinterpretation of originally derivational suffixes as having sentential scope. The lack of cognate morphemes between Eskimo-Aleut and Ural-Altaic means that, although plausible, Fortescue's proposals cannot be regarded as proven, and here it will be interesting to see whether similar processes will be discovered in future work on the diachrony of polysynthesis in those language families whose languages span the typological spectrum from synthetic to polysynthetic, offering the chance to compare both morphological forms and typological features.[5]

ABBREVIATIONS

BEN 'benefactive', COP 'copula', DU 'dualic', FAC 'factual', FUT 'future', MASC 'masculine', NP 'noun phrase', OBJ 'object', pl 'plural', PP 'past perfective', PST 'past', PUNC 'punctual', REFL 'reflexive', REP 'repetetive', RR 'reflexive / reciprocal', SEMITRANS 'semitransitive', sg 'singular', SJ 'stem-joiner', SUBJ 'subject'

REFERENCES

BAKER, MARK. 1996. *The Polysynthesis Parameter.* New York & Oxford: Oxford University Press.

COMRIE, BERNARD. 1981. *The languages of the Soviet Union.* Cambridge: Cambridge University Press.

FORTESCUE, MICHAEL. 1983. A Comparative Manual of Affixes for the Inuit Dialects of Greenland, Canada, and Alaska. (Meddelelser om Grømland, Man & Society 4.) Copenhagen: Nyt Nordisk Forlag.

FORTESCUE, MICHAEL. 1994. Morphology, polysynthetic. In R.E. Asher & J.M.Y. Simpson (eds.) *The Encyclopaedia of Language and Linguistics*, Vol. 5. Pp. 2600-2602. Pergamon Press & Aberdeen University Press.

HEATH, JEFFREY. 1975. Some functional relationships in grammar. *Language* 51.1:89-104.

HUMBOLDT, WILHELM VON. 1836. Über die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluß auf die geistige Entwickelung des Menschengeschlechts. Berlin: Königliche Akademie der Wissenschaften / F. Dümmler.

JELINEK, ELOISE. 1984. Empty categories, case and configurationality. *Natural Language and Linguistic Theory* 2:39-76.

5 See REID (forthcoming) for a remarkable account of an Australian language that has developed fully-fledged polysynthetic characteristics in the last eighty years.

JELINEK, ELOISE. 1995. Quantification in Straits Salish. In Emmon Bach, Eloise Jelinek, Angelika Kratzer & Barbara Hall Partee (eds.), *Quantification in Natural Languages.* Dordrecht: Kluwer Academic. Pp. 487-540.

KLEINSCHMIDT, SAMUEL 1851. *Grammatik der Grönlandischen Sprache.* Berlin: de Gruyter .

LAUNEY, MICHEL. 1994. *Une grammaire omniprédicative. Essai sur la morphosyntaxe du nahuatl classique.* Paris: CNRS.

MCKAY, GRAHAM. 1975. Rembarnga, a language of Central Arnhem Land. Unpublished Ph.D. thesis, Australian National University.

MITHUN, MARIANNE. 1984. How to avoid subordination. *BLS* 16:73-85.

MITHUN, MARIANNE & GREV CORBETT. 1999. The effect of noun incorporation on argument structure. In Lurella Mereu (ed.) *Boundaries of morphology and syntax.* Amsterdam: John Benjamins. Pp. 49-71.

NETTLE, DANIEL. 1999. *Linguistic Diversity.* Oxford University Press.

PAWLEY, ANDREW. 1993. A language that defies description by ordinary means. In William A. Foley (ed.) *The Role of Theory in Language Description.* Berlin: Mouton de Gruyter.

PAWLEY, ANDREW & FRANCES H. SYDER. 1983. Two puzzles for linguistic theory: nativelike selection and nativelike fluency. In Jack C. Richards & Richard W. Schmidt (eds.) *Language and Communication.* London: Longmans. Pp. 191-225.

REID, NICHOLAS. Forthcoming. Phrasal verb to polysynthetic verb. Recorded morphosyntactic change in Ngan'gityemerri. In Nicholas Evans (ed.) *Studies in comparative non-Pama-Nyungan.*

SCHULTZE-BERNDT, EVA. 2000. *Simple and complex verbs in Jaminjung. A study of event categorisation in an Australian language.* Nijmegen: Proefschrift.

The true status of grammatical object affixes: evidence from Bininj Gun-wok[1] [2]

NICHOLAS EVANS

1. Introduction

One trend in recent work on polysynthetic and other head-marking languages has been to see them as identical, at some level of structure, to more familiar European languages, except for a requirement that arguments be encoded as bound material on the verb (see e.g. Baker 1996):

> 'All Mohawk verbs are inflected to show the person, number and gender features of both their subject and their object..... These inflections are obligatory, fixed in position, and in a one-to-one correspondence with the arguments of the verb... *Thus, it is natural to say that these inflections count as pronouns, and provide the true subject and object of the verb* [italics mine]. Full NPs,[3] when they appear, have the status of some kind of adjunct or modifier.' (BAKER 1996:11)

---

1 I thank the Humboldt Foundation for financial support during the time of writing this paper, the Australian Research Council (Grants 'Non-Pama-Nyungan Languages of Northern Australia', 'Polysemy and Semantic Change in Australian Languages' and 'Reference Grammar of Mayali: cross-dialect materials'), the Australian National Parks and Wildlife Service and the Gagudju Association for supporting my fieldwork on Bininj Gun-wok since 1986, Murray Garde for checking out further data and examples, Michael Fortescue, Murray Garde, Nikolaus Himmelmann, Dagmar Jung, Myung-Hee Kim, Hans-Juergen Sasse and Jane Simpson for comments on various versions of this paper, and most importantly, the many speakers of Mayali, Kunwinjku and Kune who have taught me about their language, in particular †Toby Gangele, Eddie Hardy, †David Kanari, Goldie Blyth, Mick Kubarkku, †David Karlbuma, Tom Wood and Victor Rostron.

2 The current paper is a slightly revised version of EVANS (1999). I have taken advantage of this opportunity to correct a number of mistakes in that paper, and to incorporate some comments on an important recent paper by SIERWIERSKA (1999), which claims not to find grammatical object agreement in a sample of 272 languages. See Footnote 20 for my response to this claim.

3 By citing this passage I do not wish to imply that Baker is unaware of the problems of definiteness and referentiality to be discussed in this paper, since many of the examples of indefinite objects (e.g. in Aztec and Mohawk) are mentioned in his 1995 book. Rather, his theoretical framework

The purpose of the present paper is to contest this view, by exploring one non-trivial consequence of requiring object arguments to be represented by bound affixes: the fact that, being obligatory, they will no longer be able to encode such contrasts as referential vs non-referential, definite vs indefinite and so on. As a result, bound object affixes in at least some polysynthetic languages pattern more like subject agreement morphology in European languages than like free pronouns, in that they specify person and number information while remaining non-committal about reference and discourse status. A corollary of this is that, in order to obtain a full referring expression, external material needs to be integrated more tightly than can be captured simply by treating it as an adjunct. This is more compatible with 'dual-structure', unificational or construction-grammar approaches drawing information about arguments from both the verb morphology and external material.

Let me make clear from the outset that I do not disagree with Baker's basic claim that the polypersonal inflections on the verb are capable of saturating the argument structure of the verb. Rather, it is the specific equation of these inflections with pronouns, and the corollary that external material is therefore always some kind of adjunct, that I wish to contest. The degree to which our positions differ turns, of course, on the exact meaning of the metalinguistic term 'pronoun' in the above quote, and in general this term is used in a dangerously imprecise way in much contemporary linguistic discussion. On my reading of his book, he intends 'pronoun' here in its prototypical sense of 'personal pronoun', i.e. as 'anaphorically standing for nouns' (cf his quote 'the pronouns him, her' on p. 252), and it is when taken in this sense that I find the above quote problematic.[4] But it is not impossible that Baker intended a vaguer sense of the term 'pronoun' – something like 'participant index' or 'argument place-holder' – in which case the phenomena to be described below are compatible with his claim.

Whatever the sense he had in mind, it is crucial to understand the semantic role of argument affixes in polysynthetic languages, and it is to this goal that the present paper is dedicated. In this article the term 'pronoun', usually specified as 'free personal pronoun' unless the context makes this unnecessary, will be used in the more restrictive sense of an anaphoric expression. For want of a better term, however, I will follow the established practice of using the term 'pronominal prefix' or 'pronominal clitic', as the case may be, to refer to participant affixes in the form of bound argument affixes.

Rather than stressing the parallel between bound pronominal affixes and free personal pronouns, an alternative comparison is with subject agreement phenomena on verbs in less exotic languages. Now it is well-known that in most European languages, such as English, German, Italian or Russian, use of the 'third person singular' verb form is compatible with

---

leads him to attribute the differences to syntactic structure rather than focussing directly on the meanings of the relevant affixes. I return to this point in §6.

4 This interpretation of the above passage also correlates with his view that polysynthetic languages lack determiners, since if he intends 'anaphorically standing for nouns' by 'pronoun' the 'pronominal affixes' would then be specific enough to accomplish determination by themselves.

any possible referential status of the subject, ranging through anaphoric personal pronouns ('s/he comes'), definite NPs ('the old postman comes every morning'), indefinite expressions of various types ('a tall/different man comes here every day', 'whoever comes late gets locked out of the anatomy lecture'), negative expressions ('noone civilized rings so early') and repeated tokens in 'again' constructions ('a new president lies again every term'). BOEDER (this issue, exs 6-9) gives examples of indefinite interpretations of third person singular agreement on verbs in Classical Greek and Georgian. In other words, the system of verb agreement for subject is confined to indexing person and number, and is non-committal with respect to anaphoric or referential status.

This contrasts with the behaviour of free personal pronouns, which in the languages just mentioned are strongly identified with referential and anaphoric contexts. This is not to say there are no exceptions to this restriction. To use English examples, personal pronouns can give an indefinite reading in the following circumstances (see COOPER (1979), HAUSSER (1979), EVANS (1980) and REINHARDT (1983)):

- when combined with following relative clauses ('he who hesitates is lost')
- when functioning as object pronouns in donkey sentences ('every man who owns a donkey beats it'),
- when used as 'lazy pronouns' ('I feed my dog_i bones but my neighbour feeds it_j scraps')
- with the generic readings of 'they' and 'you' ('they always try to get you to pay more than you want to')
- in cases where reference in an apodosis is suspended from conditions in a protasis ('if you find a dodo, bring it home as a pet for my daughter').

The point is, however, that these are specific syntactic and semantic contexts that are unusual enough to warrant special treatment by logicians and formal semanticists, rather than being systematic possibilities by which pronouns can be referentially non-committal in the way comparable to person and number agreement on the verb. It is impossible to say 'she scolds them' with the meaning 'she scolds people', to say 'he will marry her again' when 'her' refers to another token of the same type (i.e. with the meaning 'he will marry another woman'), to say 'she is looking for him' with the meaning 'she is looking for a man', or to say 'he speared him again' to mean 'he speared someone again'. Yet, as we shall see, object prefixes in at least some polysynthetic languages can be used in all these cases. This illustrates the comparability of obligatory argument affixes in polysynthetic languages to verbal agreement rather than free personal pronouns, and the fallacy of the 'incorporation treatment' equating them with personal pronouns which happen to be encoded on the verb.

The reasons for this difference follow from long-recognized general structuralist principles: that small sets of choices can encode fewer semantic distinctions, and that obligatoriness deprives morphemes of expressive possibilities. Free personal pronouns, belonging to a complex set of alternatives that includes the whole range of NP possibilities (from phrases with definite or indefinite determiners, other quantifiers, bare nouns,

negatives and so on) will clearly be more specific semantically than obligatory agreement morphology with a small set of choices which in the third person is usually limited to number and gender, and this forces the agreement morphology to accept a wider range of functions.[5]

As a result, the fact that polysynthetic languages have obligatory agreement morphology for more than one argument will impact on the degree to which differences in referential and anaphoric status are highlighted in basic constructions in the language. This effect will be heightened when agreement morphology alone is often the sole means of representing arguments. SASSE (1995:205), discussing Cayuga, accurately describes the consequences of this for the average clause in that polysynthetic language: 'reference is not made directly by naming the individual, but by establishing a pronominal index which points to a relatively unspecified individual'. In this paper I will look in more detail at how such indexes encode object information in the Australian language Bininj Gun-wok.[6]

The structure of the remainder of the paper is as follows. In §2 I briefly sketch the typology of Bininj Gun-wok, including the behaviour of incorporated nominals, which parallels that of pronominal prefixes in allowing both referential and non-referential interpretations. In §3 I summarize those morphological features of the argument prefix system relevant to our discussion. In §4 I illustrate the use of object prefixes with generic and indefinite objects, and at the interaction of this referential underspecification with an adverbial prefix encoding repetition of events. In §5 I look at the various solutions polysynthetic languages can employ in their treatment of object reference, briefly discussing the parallel issues that arise as pronominal object clitics increase in obligatoriness. Finally, in §6 I set the current article in the context of work on dual representation systems, unification of morphological information from noun and verb, and on the meaning of constructions, and outline some directions needing future research.

---

5 Hans-Juergen Sasse (p.c.) points out that it may be premature to dismiss the existence of languages whose free personal pronouns pattern like argument affixes in polysynthetic languages, since there is no a priori reason to rule them out. To my knowledge, however, this remains an unattested possibility among the languages of the world.

6 I use Bininj Gun-wok as a cover term for several dialects which will be drawn upon in this paper, including Gun-djeihmi (Dj), Kunwinjku (W), Kuninjku (I) and Kune (E). The term Mayali is also used, sometimes for Gun-djeihmi dialect and several other closely-related dialects, and sometimes for the dialect chain as a whole. For a pan-dialectal grammar see EVANS (forthcoming). Note that the Gun-djeihmi orthography differs slightly from that in the other dialects. In addition to my own field-notes, examples are drawn from the following sources: texts in BERNDT & BERNDT (1951), CARROLL (1995) and NGANJMIRA (1997), and example sentences in GARDE (n.d.).

2. Typological background

Bininj Gun-wok is a highly polysynthetic language with around twelve prefix slots on the verb and two suffix slots. This is shown by the morphological template in (1). Note the following abbreviations: dir = directional ('hither' or 'thither'), imm = immediate ('right now/then/(t)here', misc = miscellaneous[7], ben = benefactive applicative, gin = generic incorporated nominal (see below), bpin = body part incorporated nominal, num = numerospatial prefix (e.g. 'many in a bunch'), com = comitative applicative, refl/recip = reflexive/reciprocal, TAM = tense/aspect/mood. For discussion and exemplification see EVANS (1995a,b, 1997, forthcoming).

(1)	-12	-11/-10	-9	-8	(-7)*	-6	-5	-4	-3	-2	-1	0	1	2
	TENSE	Subj(/obj)	dir	imm.	misc	ben.	misc	gin.	bpin.	num	com	ROOT	refl/ recip	TAM SUF.

Sample verbs illustrating some of the morphological possibilities are (2) and (3):

(2) *Birri-yawoyh-djarrk-mirnde-moname-rr-inj.*
W 3plP-again-together-many-assemble-RR-PP
'They assembled together as a group.' [Carroll 1976:62]

(3) *Aban-yawoih-warrgah-marne-ganj-ginje-ng.*
Dj 1/3pl[8]-again-wrong-BEN-meat-cook-PP
'I cooked the wrong meat for them again.'

As (3) illustrates, one source of morphological complexity in the verb is the optional incorporation of certain nominal roots inside the verb. A simpler example (4a-b) illustrates both the optionality of the process, and the fact that the incorporated form omits any noun class prefix that may be present in the external form (noun classes are shown here by roman numerals; there is a rough formal and semantic correlation with a system of gender prefixes on modifiers, such that I normally governs masculine, II feminine, III vegetable and IV neuter). The factors contributing to the more precise interpretation of incorporated nominals

7 This encompasses a wide range of adverbial type prefixes such as 'at night', 'in the wrong place', 'surreptitiously'; some of these precede and some follow the next slot, with some variation among speakers as to their exact placement.

8 In glossing subject/object combinations, I separate the subject person/number features from the object features by a slash, do not notate singular overtly, and use 'h' to mean 'higher object' and 'l' to mean 'lower object' (see §3). Thus 1/3pl means 'first person singular subject upon third person plural object', and 3/3l means 'third person singular subject upon third person singular lower animate object'.

as definite, indefinite or generic will be mentioned below. Note that only about forty nominals, almost all inanimate and generic, may incorporate in this way; other nominals must remain external. The exception is body part nouns, all of which may incorporate (see EVANS 1995b).

(4a)	*Ba-ginje-ng*	*gun-ganj.*	(4b)	*Ba-ganj-ginje-ng*	
Dj	3/3l-cook-PP	IV-meat		3/3l-meat-cook-PP	
	'(S)he cooked (the) meat.'			'(S)he cooked (the) meat.'	

The above examples show that, as is typical for polysynthetic languages, verbs may form complete sentences by themselves. Indeed, by incorporating one verb root inside another (the incorporated verb must then bear a special suffix, here glossed IVF for 'incorporating verb form'), a single verb can in fact represent a complex sentence:

(5)	*Ga-yau[-ganj-ngu-nihmi]-re.*
Dj	3-child-meat-eat-IVF-goNP
	'The child goes along eating meat.'

Subject and object prefixes are comparable to agreement suffixes in well-known European languages in being referentially open. Examples illustrating this point for objects will be given later, but the phenonemon as regards subjects can be illustrated by (6a-e), where different types of external nominal expression (or none at all) occur, providing the referential specificity needed to enrich the interpretation of the argument prefix beyond mere 'third person singular'. Note in passing that 'singular' will not be shown in glosses, but that non-singular categories (dual and plural)[9] will.

(6a)	*ga-wokdi*		(6b)	*nungga*	*ga-wokdi*
Dj	3-speakNP		Dj	he	3-speakNP
	'(s)he/it is speaking'			'he is speaking'	

(6c)	*nungan*	*ga-wokdi*
Dj	heEMPH	3-speakNP
	'HE is speaking.'	

(6d)	*bininj*	*ga-wokdi*
Dj	man	3-speak-NP
	'the/a man is speaking'	

9 In more detailed treatments elsewhere I use a different glossing system for number, based on an analysis as a minimal/augmented rather than singular/plural system, but I avoid this here so as to simplify the treatment.

(6e)	*na-mekke*	*bininj*	*ga-wokdi*
Dj	MA-DEM	man	3-speak-NP
	'the/that man is speaking'		

More generally, referring expressions need to be built up through the integration of external and internal material. Consider the following two sentences from a text commenting on a small flock of pigeons walking down towards a river to drink; for expository purposes I have bolded those portions of the sentence which are unified to give the referring expression at issue.

(7)	***na-bene maih***	*a-na-ng*	***ga-m-golu-rr-en***	*gaddum-be*	*djohboi*
Dj	MA-thatbird	1/3-see-PP	3-hither-descend-RR-NP	up-ABL	poor.thing
	'I've seen **those birds** coming down (to the waterhole) from higher up, dear little things.'				

..... [3 lines of text intervene] ...

na-mege	***maih***	***ngarrgu***	***gabarri-bódjare***	*gukku*
MA-that	bird	our	3pl-thirstyNP	water
'**Those birds of ours**, they're thirsty for water.'				

In both lines the full semantic specification needed to reach the interpretation 'those birds' and 'those birds of ours' can only be obtained by combining information from external and internal material. The first three words could either be translated as 'I saw that bird' or 'I saw those birds'; the plural interpretation is then forced by the next verb, which uses a reflexive/reciprocal suffix with a collective interpretation, even though the object prefix (as is normal with lower animates) does not encode their number. In the subsequent line the speaker, somewhat unusually, goes on to use the plural form of the prefix to indicate plurality, as a way of personifying the birds (and the theme of the whole text is the way the birds care for their offspring by taking water back in their crops).

Supplying the number of nominal expressions is only one example of how information from the verb is unified with external nominal material. Another common situation involves the integration of incorporated nominals with external modifiers, as in (8-10); note that (as with the interpretation of third person prefixes), the choice of external modifiers is an important factor in determining the definiteness and referentiality of incorporated nominals.

(8)	*Al-wakadj*	*ka-yaw-karrm-e*	***al-daluk***.
Dj	[name]	3/3l-child-have-NP	FE-female
	'Al-wakadj has **a female child**.'		

(9) *Kamak* *kan-**bolk**-bukka-n* ***ke.***
Dj good 2/1-country-show-NP your
'It is good that you will show me **your country**.'

(10) *Nga-**murrng**-bimbom* ***na-mekke.***
W 1/3-bone-paintPP MA-DEM
'I painted those bones.' (Carroll 1995:353)

In fact, the interpretation of referentiality and definiteness in incorporated nouns is as complex as with pronominal prefixes. Certain determiners favour definite interpretations (e.g. possessive pronouns, as in (9), and some demonstratives, as in (10)), but this is not the only possible interpretation, so that (9), for example, could mean 'show me (some) country of yours'. The verb lexeme is also important: for example, incorporation of generic nominals into stance verbs (typically in conjunction with an adjacent external nominal furnishing more specific information) always has a thetic or existential interpretation, as in (11):

(11) *Gonhdah* *ga-**rrulk**-di* ***an-bernbern.***
Dj here 3-tree-standNP III-ghost.gum
'There's a ghost gum tree here.'

In connected text there is a strong tendency for inanimate protagonists introduced as free nominals to progress to incorporated status, correlating (approximately) with definiteness or pronominalization in English:

(12) *Ngaye* *gorrogo* *an-bang* *nga-gurrm-i,*
Dj 1sg before III-dangerous 1/3-put down-PI

gun-gurlah *a-ma-ngi.*
IV-skin 1/3-get-PI

gun-gurlah *a-ga-ni* *djamun-djahdjam.*
IV-skin 1/3-take-PI dangerous-place

*a-**gurlah**-wo-ni,* *gun-warde* *an-wo-ni.*
1/3-skin-give-PI IV-money 3/1-give-PI

'In the old days I used to put down baits to get (dingo) **hides**.
I would take **the hides** to the police station. I would give **them/the hides** to him and he would give me money.'

Finally, it is also possible to give incorporated nominals a secondary predicate interpretation. This is particularly likely when the argument to which they are linked is elsewhere specified as first or second person (13b), or when incorporated into intransitive verbs of representation, such as 'speak, sound' in (14) and *kurrmerren* 'put oneself (into the landscape)' in (15). In this case the incorporated nominal is never modified.

(13a) *Bi-yau-bawo-ng.*
Dj 3/3hP-child-leave-PP
'(S)he abandoned/left the child.'

(13b) *Ngan-yau-bawo-ng.*
3/1-child-leave-PP
'(S)he abandoned/left me as a child.'

(14) *Nga-bekka-n* *ga-rurrk-wokdi.*
Dj 1/3-hear-NP 3-cavity-speakNP
'I hear that it (the tree) sounds hollow.' (In looking for a tree suitable to make into a didgeridoo). (Lit. 'I hear it, it sounds as a cavity' or 'I hear it, the cavity sounds')

(15) *Daddubbe* *ba-bim-gurrme-rre-ni.*
Dj [name] 3-image-put-RR-PI
'Daddubbe would put himself into the landscape as a rock painting.'

The above discussion should indicate the complex way in which construction type, modifier type, verb semantics and discourse progression combine to determine the way in which incorporated nominals are referentialized. We shall encounter similar complexities when we turn to object pronominal prefixes in §4.

3. The form of subject/object prefixes: summary remarks

Verbal prefixes for subject and object form a large paradigm, which will not be treated exhaustively here – see CARROLL 1975, ETHERINGTON & ETHERINGTON 1994 and EVANS forthcoming for fuller details. Briefly, there is variation in the degree to which subject and objects can be segmented, with portmanteaux most common when both subject and object are speech act participant (first or second person). Where a distinct object prefix is analysable it usually contains a final *-n* segment, e.g. Dj *nga-ba-n-bani* [1-3nsg-OBJ-dual] '1/3du'. As in many Australian languages, there are two competing ordering principles: subject before object, and first/second person before third person, with the second principle more highly ranked. A supplementary principle is that certain object number markers will be placed in final position, even at the cost of detachment from the morpheme whose number they are modifying, e.g. *nga-n-di* [1-OBJ-3nsgSUBJ] '3>1pl'.

Figure 1 illustrates a subset of the pronominal prefix paradigm, and includes relevant dialect variants; 'factored' morphemes for subject and object are shown in the relevant part

of the margins. Forms with third person subjects indicate a two-way tense contrast, essentially non-past *ga-/ka-* (the difference is merely orthographic) vs past *ø-*, as in Dj *gabarri-* '3pl/3NP' vs *barri-* '3pl/3P'; for brevity NP will not be glossed overtly in pronominal prefixes. In Dj the expected *ø-* prefix for '3/3P' is replaced by the form *ba-* .

OBJ / SUBJ	3sg. lower animate [also: intrans] *ø-*	3sg.higher animate *ø-* , *bi-*	3pl *bVn-*
1sg *nga-*	*nga-*	*nga-*	*ngaban-* (Dj) *ngaben-* (W) *ngabin-* (E)
2sg *yi-*	*yi-*	*yi-*	*yiban-* (Dj) *yiben-* (W) *yibin-* (E)
2pl *ngurri-*	*ngurri-*	*ngurri-*	*ngurrban-* (Dj) *ngurrben-* (W) *ngurrbin-* (E)
3pl *(g/ka)bVrri-*, *(g/ka)bV-..-di*	*(ga)barri-* (Dj) *(ka)birri-* (W, E)	*(ga)barri-* (Dj) *(ka)birri-* (W, E)	*(ga)bandi-* (Dj) *(ka)bindi-* (W, E)
3sg *(g/ka)-*	*ga-* (pr; Dj), *ka-* (pr; W, E) *ba-* (pst: Dj), *ø-* (pst: W, E)	*(ga)bi-* (Dj) *(ka)bi-* (W, E)	*(ga)ban-* (Dj) *(ka)ben-* (W) *(ka)bin-* (E)

Figure 1: a subset of the subject/object prefix paradigm. Shaded areas indicate zero marking of object.

As Figure 1 shows, 3sg objects are marked by zero whenever the subject is other than third person singular. If the subject is third person singular, there is a choice between the *(ga-/ba-)/(ka-/ø-)* and *(ga)bi-/(ka)bi-* sets which I shall refer to, in summary fashion, as a choice between *ø-* and *bi-* . Note that, because the zero-marked object forms are indistinguishable from intransitive forms, so that on formal grounds *nga-ngun* could mean 'I (am) eat(ing)' or 'I (am) eat(ing) him/her/it', the *bi-* forms are usually the only indicator (in the case of singular or lower animate subjects) that the verb is transitive and that an object argument is present.

A number of semantic factors contribute to the choice between zero and *bi-* forms. In many situations the choice resembles a direct vs inverse system, so that '(s)he > it' takes the zero form (16) while 'it > him/her' takes the *bi-* form (17). For this reason the zero form will be glossed as '3/3l' for 'third singular subject upon third singular lower object' and the zero form as '3/3h' for 'third singular subject upon third singular higher object'.

(16a)	*Bininj*	*ø-ngune-ng*	*duruk.*	(16b)	*Bininj*	*ba-ngune-ng*	*duruk.*
W, E	man	3/3lP-eat-PP	dog	Dj	man	3/3lP-eat-PP	dog
	'The man ate the dog.'				'The man ate the dog.'		

(17)	*Bininj*	*bi-baye-ng*	*duruk*
Dj, W, E	man	3/3hP-bite-PP	dog
	'The dog bit the man.'		

Where both subject and object are below the human threshold, choice between zero and *bi-* is determined by relative position on a hierarchy[10] comparable to that found in Navajo. Thus a snake biting a dog would be *bi*, while a dog biting a snake would be *ø-*. Likewise, fire frightening a dog or buffalo would require *bi-* . Where both are above the human threshold (e.g. a man and a woman) the *bi-* form is always used, except that when babies are objects there is variation between *bi* and *ø-* forms. When both are below the threshold, but equivalent in animacy, e.g. two dogs fighting each other, or two kangaroos, the choice between *ø-* and *bi-* will depend partly on empathy, with*bi-* reflecting greater empathy or personification. But the choice here also varies across dialects: the easternmost dialect, Kune, is more likely to employ it in these circumstances.

In ditransitive and benefactive constructions, in which the object pronominal slot expresses the indirect object, the *bi-* prefix represents third person singular indirect objects (18). Since indirect objects are virtually always human, the use of the form can be explained by the animacy rather than the grammatical role of the argument, but in any case it means that third person singular indirect objects will always take the *bi-* form.

(18a)	*Gabi- ganj-wo-n.*		(18b)	*Gabi-marne-ganj-ginje.*	
Dj	3/3h-meat-give-NP			3/3h-meat-cookNP	
	'(S)he gives him/her the meat.'			'(S)he	cooks him/her the meat.'

Finally, for some speakers the choice between zero and *bi* forms (with human objects) is influenced by the referentiality of the object if the verb is lower in discourse transitivity,

10 In the related language Rembarrnga, from the Eastern Gunwinyguan group, there is a similar hierarchy, but 'higher' combinations use the form *barr-* , which is homophonous with the third person plural form.

such as 'look for'; in such cases, *bi* forms are confined to referential objects. We return to, and exemplify, this point, in §4.2.

COMPARATIVE REMARKS ON BI-.

Since non-zero markers of third person singular object are typologically uncommon, it is worth making a brief excursus on the origins of the *bi-* formative, even though we cannot give a definitive account at this stage of research on the comparative morphology of the Gunwinyguan and related languages. (Note that Bininj Gun-wok belongs to the Central subgroup of the Gunwinyguan family).

Dalabon, the language closest to Bininj Gun-wok, has a cognate form *bv-* (where v represents a high central vowel), though combined in the opposite order to give *bvkah-* instead of *kabi-*. (In Dalabon the *ka-* third singular prefix is generalized to all tenses, and the *h-* is an additional formative marking (approximately) realisness). In Dalabon there is a choice between *kah-* and *bvkah-* that basically parallels the choice between *ka-* and *kabi-*, though as in the Kune dialect of Bininj Gun-wok *kabi-* is used for a wider range of objects.

In Jawoyn, a Western Gunwinyguan language which borders the Gun-djeihmi dialect there is a benefactive prefix *bi-* which follows the subject/object prefixes, and adds a benefactive object. Compare Gun-djeihmi *gabi-mang*, which means 'he gets him/her' (e.g. he marries her) because of its simple 3/3h meaning, and Jawoyn *ga-bi-mang* (MERLAN n.d.) which means '(s)he gets it for him/her', as in *giyowk gabi-mang* 'he hunts/gets meat for her'. Here *ga-* means '3/3' and *bi-* adds a benefactive object. The Jawoyn facts suggest the possibility that *bi-* in Bininj Gun-wok and Dalabon results from the generalization of an old benefactive marker to other human third person singular arguments, working outwards from high-animacy objects (much like Spanish *a* or Hindi *ke*, but in this case using an applicative in a prefixal complex rather than an adposition); this could be linked to the development of a new benefactive prefix *marne-/marnanj-* in the Central Gunwinyguan languages.

An alternative explanation would link *bi-* to prefixes marking a third person human category in a number of languages inside and outside Gunwinyguan (and perhaps related to the *bV-* formative implied by the third person dual forms *bani-/ bene-/bini-* in Bininj Gun-wok, and the corresponding third person plural forms *barri- / birri-*). In the Western Gunwinyguan language Ngalakan, for example, third person non-singular objects are marked by a prefix *bi-* or *bu-* , depending on the vocalism of the preceding prefix: cf *ngugu-* [1sg/3sgNEUT], *ngumu-* [1sg/3sgVEG], *ngubu-* [1sg/3pl], *yigi-* [1inclusive.dual/3sgNEUT], *yimi-* [1inclusive.dual/3sgVEG], *yibi-* [1inclusive/3pl] (Merlan 1983:88). Recall (f/n 9) that in Rembarrnga a form identical with the plural is used with human or higher-animate objects, suggesting extension of *bi-* from '3plOBJ' to 3higherOBJ' would not be implausible.

Finally, it should be mentioned that the word for 'man, person' is *bi* in a number of Gunwinyguan languages (Dalabon, Rembarrnga) while in others words for 'man, person' have this as their first element (Bininj Gun-wok *bininj*, Ngalakan *bigurr*). This suggests a third etymology for the *bi-* prefix, namely that it may have arisen by reanalysing an

incorporated form *bi-* from 'man, person' to 'third person higher object'. However, the restrictions against incorporating the word for 'man' in all Gunwinyguan languages make this unlikely. What is more likely is that, if *bi* turns out to be reconstructed as a word for 'man, person' at a much earlier stage, it could have served as a grammaticalization source for both a 'human' class prefix and a third person pronominal prefix to which non-singular number prefixes were added (with 'people' – *bi*-plural – being grammaticized to 'they'). In defence of attributing *bi* 'person' to an early ancestral stage, note that in Iwaidja, only distantly related, 'man' is *arrgbi*, whose form is compatible with *arrg-* being a prefix 'our, we' to *bi* 'person', i.e. 'our person, one of us').

Deciding between these rival explanations is difficult in our current state of knowledge (and various composite explanations are in any case imaginable), but all suggest that the *bi-* forms have a long association with either indirect or human objects, or both.

4. 'Non-pronoun' uses of third person object prefixes

We now examine in more detail a number of cases in which third person object prefixes are used in situations where a personal pronoun would not be used in a language like English or German. (As will become clear, Bininj Gun-wok, as well, disallows the use of free personal pronouns in such cases). Because of the difficulties of proving that zero prefixes are not simply analytic artefacts – in other words, of disproving the alternative analysis that *zero isn't there* (cf MITHUN 1986) in those parts of the paradigm where third person singular object forms are identical to the intransitive forms – I will confine myself to cases where the object is represented by a non-zero form: either a third plural object form or a *bi-*form.

In §4.1 I consider the expression of generic objects; in §4.2 I look at unspecified and other forms of non-referential objects; and in §4.3 a number of examples of various types are brought together, in a single stretch of text. In §4.4 I look at how the rules for the expression of non-referential objects can account for what, from an English point of view, appear to be anomalous uses of the object pronominal prefixes with *yawoyh-* 'again'.

4.1 Generic objects

In English, as in other European languages, third person plural personal pronouns, as in (19a), do not allow a generic interpretation; to obtain this a bare plural must be used instead of a pronoun, as in (19b):

(19a) She scolds **them**.
(19b) She scolds **people**.

In Bininj Gun-wok, however, both can be encoded with a bound plural object prefix:

(20) *al-ege* *daluk* ***gaban**-du-ng*
Dj FEM-DEM woman 3/3pl-scold-NP
a. 'That woman scolds people.'
b. 'That woman is scolding them.'

To favour a generic reading, the verb may be followed directly by a bare noun, as in (21) and (22). As mentioned in §2, number is not normally marked on nouns themselves, so the noun may be formally indistinguishable from a singular, as in (21), but in some cases reduplicated forms specific to plurals may be used (22). With eligible nouns incorporation is also possible, as in (23).

(21) *munguyh* ***kaben**-yawa-n* ***daluk*** *minj* *kabi-marnedjare*
I always 3/3pl-look.for-NP woman not 3/3h-loveNP

daluk *bininj* *na-mekke.* *Kabirri-warnyak daluk*
woman man MA-DEM 3pl/3-not.want woman
'He's always **looking for women**, but there is no woman who loves that man. Women don't want him.' (GARDE 1996)

(22) *Kulalung* *ka-rohrok* *yiman* *Namorrorddo.* *Nungka*
W [name] 3-same like [name] he

kaben**-widnan* ***wurdwurd.
3/3pl-hateNP children

Bu *kaben-bekka-n* *kabirrih-nalkbu-n* *munguyh.*
SUB 3/3pl-hear-NP 3pl-IMM-cry-NP always
'Kulanung is a spirit like a Namorrorddo. He hates **hearing babies** cry and cry all the time.' (NGANJMIRA 1997:232)

(23) ***Kaben-yaw**-wo-n* *ngal-kunburriyaymi* *yiman* *ngalyod*
I 3/3pl-baby-give-NP II-mermaid.spirit like II-rainbow.serpent

la *ngal-berd-djenj* *ngal-buyika.*
CONJ II-tail-fish FE-other
'The mermaid spirit can **have babies** just like the rainbow serpent but the woman with a tail like a fish (another variety) is different.' (GARDE 1997)

To force a definite referential reading, on the other hand a free personal pronoun can be used (24). Referential readings are also forced by the use of nominal groups containing an anaphoric demonstrative (25). Note that this counters the claim made in BAKER (1995:253) that the 'lack of semantically significant determiners seems to be a property of the polysynthetic languages as a class', since demonstratives clearly function as determiners in Bininj Gun-wok.

(24)	*al-ege*	*daluk*	*gaban-du-ng*	*bedda*
Dj	FEM-DEM	woman	3/3pl-scold-NP	them
	'That woman is scolding them.' (cf 20)			

(25)	*ngal-mekke*	*daluk*	*kabin-du-ng*	*na-mekke*	*ngong*
E	FEM-DEM	woman	3/3pl-scold-NP	MA-DEM	mob
	'That woman is scolding that mob.' (cf 20)				

A rarer alternative to the use of plural object prefixes on the verb is the use of the ø-form of the third person singular, followed by a bare noun. An example is:

(26)	*Namorrorddo*	*na-bang*	*ka-djare*	*ka-bu-n*	*bininj*
I	[name.of.dangerous.spirit]MA-dangerous		3-wantNP	3/3l-kill-NP	person
	'Namorrorddo is dangerous, he kills humans.'				

4.2 Indefinite and unspecified objects

A second situation where English and other European languages may not use personal pronouns is with unspecified objects, e.g. indefinite objects of irrealis verbs (27b), implicit indefinite pronouns translateable as 'someone' (28b), objects which are indefinite (29b) or whose exact identity is unimportant (30b), or indefinite 'intentional objects' (QUINE 1960), as in (31b):

(27a)	She should have married him, the Welshman.
(27b)	*She should have married him, a/any Welshman.
(28a)	She says to him (, to her husband), 'I've seen some honey.'
(28b)	* She says to him (, to someone), 'I've seen some honey.'
(29a)	She married him, the white man.
(29b)	*She married him, a/some white man.
(30a)	The messenger will bring them, all the people.

(30b) *The messenger will bring them, many people (whose identity is unknown or unimportant)
(31a) He is looking for her, for his wife.
(31b) * He is looking for her, for a wife.

For the first four of these situations the *bi-* form of the third person pronominal prefix is always used in Bininj Gun-wok, as illustrated by (32 - 35) respectively:

(32) ***Bi**-ma-yi* *Na-burlanj* *kun-mak*
W 3/3hP-marry-IRR [marriage.class] IV-good
a. 'She should have married Naburlanj; that fellow whose section is Naburlanj.'
b. 'She should have married a (=some) Na-burlanj man'.[11]

(33) *Bad* *ka-re* *an-kung* *ka-na-n,* *ka-rrurnde-ng*
Dnj but 3-goNP III-honey 3/3l-seeNP 3-return-NP

***kabi**-marne-yime* *"An-kung* *nga-na-ng.* *Come on* *ngarr-e,* *ngarr-ni*
3/3h-BEN-sayNP III-honey 1/3-see-PP 12-goNP 12-sitNP

'iiiiiiiii' *ka-yime* *ka-m-borled-borledme* *wanjh* *maitbi* *ngarri-dolkka-ni.*
[noise] 3-sayNP 3-hither-ITER-turn.roundNP then maybe 1pl-get.up-PI

'[Talking about the None honey wasp spirit:) But he goes and sees some honey, comes back **and says to someone**: 'I've seen some honey. Come one, let's go, he goes "iiiiiiiii" and buzzes around and around; then maybe we get up (to look).'
[Lofty Nadjimerek Bardayal per MG]

(34) *Balanda* *bi-mey*
E:D European 3/3hP-marryPP
a. 'She married the white man.'
b. 'She married a white man.'

(35) *Kakkawarr* *kaben-ma-ng* *birri-wern* *bininj.*
I messenger 3/3pl-bring-NP 3pl-many person
a. 'The messenger will bring the many people.'
b. 'The messenger will bring many people.'

11 Here *Naburlanj* refers to a marriage category (containing roughly one eighth of the male populations) from which the subject should have taken her husband.

With verbs taking a psychological object, such as *yawan* 'look for, seek', there is some disagreement among speakers about whether the *bi-* form can be used with indefinite objects. Thus for some speakers (36a) is ambiguous, allowing the two interpretations shown, while for other speakers only a definite/possessed reading is possible with the *bi*-form, and the indefinite reading requires the ø-form, as shown in (36b). Once the object is modified by a demonstrative, as in (37), all speakers choose the *bi*-form.

(36a) *Bininj* ***gabi****-yawa-n.*
Dj man 3/3h-seek-NP
a. 'She is looking for her man.' (all speakers)
b. 'She is looking for a man (any man).' (some speakers)

(36b) *Bininj* ***ga****-yawa-n.*
Dj man 3(/3l)-seek-NP
'She is looking for a man (any man).' (some speakers)

(37) ***gabi****-yawa-n* *na-mekke* *bininj* *(ngarre).*
Dj 3/3h-seek-NP MA-DEM man her
'She is looking for that man (of hers).' (all speakers)

Given the fact that 'have' constructions take non-referential objects (cf VAN GEENHOVEN 1998 on Eskimo), it is worth pointing out here that the verb *karrme*, which can mean either 'hold, grasp' or 'have', exhibits a split here: when used with the 'hold' meaning it takes the *bi-* form if singular and the appropriate dual or plural form if non-singular (38), but when used with the 'have' meaning it takes the ø-form regardless of the object (39). This holds for all speakers.

(38) a. *Kabi-karrme* *daluk.*
I 3/3h-holdNP woman
'(S)he is holding the/a woman.'

b. *Kabin-karrme* *daluk* *danjbik.*
3/3pl-holdNP woman three
'(S)he is holding (the) three women.'

(39) a. *Ka-karrme* *daluk.*
I 3/3l-haveNP woman
'He has a wive / wives.'

b. *Ka-karrme* *daluk* *danjbik.*
3/3l-haveNP woman three
'He has three wives.'

It appears, therefore, that for verbs which lexically require non-referentiality in their objects, such as 'have', the *bi-* construction is never used (alternatively, one could attribute this to their low discourse transitivity); and for the psychological-object verb 'look for' some speakers choose between ø- and *bi*-forms according to whether the object is referential. As mentioned in §3, the ø-forms are in fact identical with the intransitive prefix forms, so that in such cases one could argue that the intransitive prefix set is being used.

For other transitive verbs, however, the *bi-* form is used regardless of the referentiality, definiteness or givenness of the object.

Before leaving this section, mention should be made of a construction which is translated into an English with a singular subject and an indefinite object, but rendered in Bininj Gun-wok as an intransitivized reciprocal and a dual subject in a 'set-subset' construction, as in (40).

(40) *Ani-ma-rre-n* *Al-mardgu.*
Dj 1du-marry-RR-NP II-Mardgu (matrimoiety)
'I have to marry a woman of *Al-mardgu* matrimoiety.'
(Lit. 'We two marry, me and (some) Al-mardgu woman.)

Such constructions run counter to the claim sometimes made that first person subjects are always referential, because first person non-singular exclusives are made up of a first person referent (which is referential) plus one or more third person referents (which need not be referential).

4.3 A textual example

To illustrate how interpretation depends on the interplay between pronominal affixes, external nominals and the unfolding discourse context, consider the following textual excerpt from the *Kunwinjku Spirit* corpus (NGANJMIRA 1997:4-6), told by Alex Nganjmira and translated by a group of people at the Kunwinjku Language Centre. The translation has been slightly adapted to make it more literal; in the couple of places where the alteration is relevant to the issue of referentiality, the original translation follows mine in square brackets. Nominal expressions we will be discussing below are in bold.

(41) Background: Yingarna the Rainbow Snake, ancestress of people today, came across
(W) under the seas, and reached the island of Waminari, where:

a. *Ngaleng* ***ben-yawme-y*** ***birri-wern***
she 3/3plP-bear.child-PP 3pl-many

birri-kuk-bu-buyika ***bininj*** ***dja*** ***mayh***
3pl-.body-REDUP-different person and animal
'She gave birth to **many strange people and animals**,

b. *Birri-warre-ni* *birri-kodj-buyika-ni* *dja* *birri-burrk-buyika-ni.*
3pl-bad-PI 3pl-head-different-PI CONJ 3pl-body-different-PI
'They were ill-formed, with strange heads and bodies.

c. *Minj* ***ben****-marnedjare-niwirrinj,* *wanjh* *ø-yime-ng:*
NEG 3/3plP-like-IRR then 3P-say-PP
'She didn't like **them**, and said:

d. *"Nga****ben****-yawoyh-kuk-ngu-n,* *kaluk* *nga****ben****-djordm-ih-we*
1/3pl-again-body-eat-NP later 1/3pl-grow-IVF-throwNP
kore *ku-njam."*
LOC LOC-stomach
'I'll swallow **them** again, and later I'll grow **some more** in my belly.

e. *Kaluk* *yerrih* *na-kudji man-djewk* ***ben****-yawoyh-yawme-y.*
then after MA-one III-rain 3/3plP-again-bear.child-PP
'The next year she gave birth (**to children**) again,

f. *Kaluk* *bininj* *duninjh* *na-wu* *birri-djordm-inj* *wurdyaw* *duninjh*
then person real MA-REL 3plP-grow-PP child real
'To normal people who grew up as proper children.'

[Line omitted explaining how these children grew up to be the various tribes of the region]

g. *Yingarna* *ø-wam* *ø-durnd-i* *kore* *ngalengarre*
[name] 3P-goPP 3P-returnPP LOC 3femOBL
kun-red *Waminari*
IV-place [name]
'Yingarna went back to her place at Waminari.

h. *Wanjh* ***ben****-berd-djobke-ng* *wurdwurd* *ku-mekke* *kun-red,*
then 3/3plP-penis-cut-PP children LOC-DEM IV-place
dja *birri-dowe-ng.*
CONJ 3plP-die-PP
'Then she circumcised **some children** at that place, but they died.

i. *Yerre* *ngaleng* *ø-wam* *Marrkolidjban* *dja* ***ben****-berd-djobke-ng*
later she 3P-goPP [place] CONJ 3/3plP-penis-cut-PP
wurdwurd *ku-mekke.*
children LOC-DEM
'Later she went to Marrkolidjban and circumcised **some boys** there (to see if they would be O.K.),

j.	*Kamak*	*kun-ekke,*	*minj*	*ngad*	*karr**ben**-berd-djobke*	*kondah* .....
	good	IV-DEM	NEG	we	12pl/3pl-penis-cutNP	here

and they were O.K. (But) we don't circumcise (**people**) in this area.'

A wide range of translations for the plural object marker *ben* plus its accompanying external material are appropriate in this passage.

The only straightforward examples of *ANAPHORIC USE* are in line c, and the first occurrence in line d; in both these cases there is no external material. The second occurrence in line d receives an anaphoric translation in the *Kunwinjku Spirit* translation ('grew them again in her belly') which is essentially a *LAZY PRONOUN* use pushing the boundaries of what counts as an identical entity in tracked discourse, given that a whole new set of children are involved.

Three of the remaining occurrences involve *NEW MENTIONS*, namely the examples in line a, h and i. In each of these cases a nominal group follows the verb. In the latter two cases this takes the form of a plural nominal with a spatial demonstrative (*wurdwurd kumekke* '(some) children at that place'), and in the former case there follows a lengthy nominal group of five words.

In line (e) the translation is problematic; the original translation gives an English intransitive verb as the equivalent ('give birth' rather than 'give birth to people'), but equally valid would be a translation with a bare plural (e.g. 'gave birth to babies'). In either case, nominal material making the nature of the object clear is supplied in the next line.

Finally, line (j) exemplifies a *GENERIC-OBJECT* use, which could be translated either with a bare plural generic noun, as I have done ('circumcise people'), or again with an intransitive English verb, which is what appears in the published translation. In this example no bare nominal appears to stress the generic-object interpretation, which is clear from the contrast of one (circumcising) area vs another (non-circumcising) one.

The above text is typical in representing how far constructional means, in the form of conventional interactions between verbal-prefix and external nominal material, determine the interpretation of such pronominal prefixes. Prefixes not supplemented with external material may equally well be anaphoric or generic, while in the case of new mentions external nominals typically immediately follow the verb. The unfolding discourse context is as important in determining the interpretation as the presence and nature of immediate external material. Nonetheless, where precision is desired, generic interpretations can be forced by adding bare nouns, and specific anaphoric interpretations by adding free pronouns and anaphoric demonstratives.

4.4 Interaction with 'again'

Bininj Gun-wok has a verbal prefix, *yawoyh-* (written *yawoih-* in Dj), which is roughly translatable as 'again' (see EVANS 1995a for more detailed discussion); in eastern dialects it is common to use a free adverb *rawoyhno / yawoyhno*[12] as an alternative. As with its equivalents in many languages, interesting semantic questions arise from the use of 'again' as a diagnostic of what counts as the same event. Essentially there are three types of situation where *yawoyh-* will be used:

(a) exact replays, where both the activity and all the actants remain constant under repetition (42). In English, as in other European languages, object pronouns are acceptable in such circumstances (43, 44).

(42) *He fell over again.*
(43) *She wore it to the semantics class again.*
(44) *She passed him again on the stairs.*

(b) final-state replays, where the state is repeated but the change-of-state transition may be happening for the first time. There is token-identity, although the actant may shift from subject to object relation between the two events. An example is:

(45a) *The jungle grew over the ruins again.*

This is ambiguous between a reading where the jungle is growing over the ruins for the second time, and one where buildings were built in what was originally jungle, and now reverts to being jungle again. Interestingly, while pronominalization of the object is quite natural on the first reading, it sounds slightly awkward on the second reading.

(45b) *The archaeologists cleared away the jungle from the ruins, but eventually it grew (back) over them again.*
(45c) *?The ancient Cambodians cleared away the jungle to build their temple complexes, but eventually it grew back over them again.*

(c) replays with token replacement, where one or more of the actants are replaced with another token of the same type. In English the varied token cannot be pronominalized, and 'one' (46) or 'someone else' (47) must be used instead.

(46) *I caught a fish yesterday, and I caught one (*/it) again today.*

---

12 The suffix *-no* is found on a number of manner adverbials, e.g. *molkno* 'secretly; stealthily', and the replacement of morpheme-initial *y* by *r* is a feature of eastern dialects.

(47) *I saw someone in the shed yesterday, and saw someone else (*him/them) there again today.*

It is type (c), replays with token replacement, that concerns us here, since in some circumstances pronominal prefixes are acceptable in Bininj Gun-wok where personal pronouns would be unacceptable. (48) is from a text about a cannibal who had already speared various people, and now goes out to spear a new token of the same type; the *bi* object form is used here:

(48) In a text about a cannibal who had already speared various people:
W *wanjh* ***bi-****yawoyh-yam-i* *na-buyika*
then 3/3hP-again-spear-PI MA-other
'then he would (go again and) spear another' (CARROLL 1995:363)

In (49), a text about how the early missionaries oversaw a program that systematically paired up young men and women in monogamous marriages to break down the old system of gerontocratic polygamy, both subject and object involve token replacements, yet the *bi* form is still used in line (d).

(49) W:OM45
a. 'Well, at last that white man said:

b. *'ngaben-wono-wo-n , kabirri-ma-rr-en, yawurrinj dja yawkyawk.*
1/3pl-REDUP-give-NP 3pl-marry-RR-NP young.man and young.woman
'I want to give them in marriage, those young boys and girls.'

c. *wanjh ngokkogen bene-bad-ma-rr-en, bininj daluk.*
well at last 3duP-now-marry-RR-NP man woman
'So two of them were married, a man and a woman.'

d. *ngarri-wam, dird-buyika ngal-buyika daluk* ***bi-****yawoyh-me-y*
1pl-goPP moon-other FEM-other girl 3/3hP-again-marry-PP
na-wu bininj.
MA-REL man
'We went on, and the next month another girl was married to another man.'

e. 'He kept on doing that, that white man,
f. 'giving in marriage whoever had grown up at the same time as each other .
g. 'until at last he had finished (giving) them all.'

Not only is pronominalization impossible here in the English translation of (d) – * *the next month she was married to him again* – but it becomes very awkward to use the adverb 'again' – *? a(nother) man married a(nother) woman again* – without either placing it in a separate intonation contour – *and again, a(nother) married a(nother) woman* - or using a passive construction – *another woman was married to another man again.*

Although (49) shows that the *bi* form is compatible with indefinite objects representing token replacements, it is not the case that *bi* is universally compatible as an object marker with all non-definite token replacements.

First consider (50) and (51), which illustrate the ambiguity of the prefixed word *nga-yawoyh-mang* [I/her-again-get/marry] between repeated-token and and token-replacement readings. This can be disambiguated through the choice of external material: noun plus modifiers of identity and specification for the repeated-token reading and bare-noun (optionally with modifier meaning 'another, a different') for the token-replacement reading.

(50)	*Gare*	*nga-yawoih-ma-ng*	*daluk*	*(al-buyiga).*
Dj	maybe	1(/3)-again-marry-NP	woman	FE-other
	'Maybe I'll marry (some woman/*her) again.'			

(51)	*Gare*	*nga-yawoih-ma-ng*	*ngal-gare*	*ngal-ekke*	*daluk.*
Dj	maybe	1/3-again-marry-NP	FE-old	FE-DEM	woman
	'Maybe I'll marry her/my old wife again.'				

In this case the form of the prefix is identical whether the verb is intransitive, or transitive with a third person singular object. It is possible to argue that the *nga-* prefix in (50) is merely an intransitive form, and that the expression *mang*, or *mang daluk*, is here an intransitive one (like English 'get married', or intransitive 'marry' as in 'She's in no hurry to marry'). In (51) the presence of a determined nominal group makes it clear that the verb is transitive.

Now consider what happens once the combination involves a third person singular subject, so that overt object prefixation becomes possible through the use of the *bi*-form. In cases of exact repetition, comparable to (51), the *bi*-prefix is acceptable (52a). However, where there is token-replacement, as in (52b), only the ø-form is acceptable.

(52a)	*Gare*	*gabi-yawoih-ma-ng*	*ngal-gare*	*ngal-ekke*	*daluk*
Dj	maybe	3/3h-again-marry-NP	FE-old	FE-DEM	woman
	'Maybe he will re-marry his old wife', 'Maybe he will marry her again'.				

(52b)	*Gare*	*ga-yawoih-ma-ng*	*daluk*
Dj	maybe	3(/3)-again-marry-NP	woman
	'Maybe he will marry (someone) again.'		

These examples, as with the cases of 'look for' and 'have' discussed in §4.2, illustrate that although the conditions on pronominal object prefixes in Bininj Gun-wok are much broader than those on free personal pronouns in English, there are some limits. In this case, once one is dealing with an object that is merely intentional, as in (52b), use of overt object prefixes becomes unacceptable. Nonetheless, the availability of the *bi* form for indefinite repeated tokens, as in (49) and (50), show that its scope is much wider than that of free personal pronouns.

5. Cross-linguistic parallels

We have now established that, in Bininj Gun-wok at least, argument prefixes can be used to represent objects in a range of circumstances where free personal pronouns would be inappropriate, such as generic objects, indefinite and certain non-referential objects, and replacement tokens of the same type. We turn next to the question of whether this is a language-specific phenomenon, or is more generally characteristic of languages with multiple argument representation on the verb (or, more generally, on other clausal heads such as second-position auxiliaries). Although many grammars remain silent on this issue, we shall see that it is in fact quite a common phenonemon in head-marking languages.

In the Western Gunwinyguan language Warray, for example (HARVEY 1986:142-3), the third person plural object form can be used with an indefinite reading (53); Harvey also gives examples of the relevant third person plural forms being used for indefinite subjects.

(53) *Pat-put-na-y*
1sg.non.complete-3pl-see-REALIS
'I can see somebody (coming).'

In Mangarayi, closely related to the Gunwinyguan family, the rules for marking objects are rather complex, and number neutralization on the verbal argument prefixes and/or external numbers is responsive to definiteness (see MERLAN 1982:93-6). Nonetheless, in at least some cases indefinite objects are marked by the third person plural object prefix, whether meaning 'some (people)' (54) or 'any (people)' (55).

(54) *Ngayangayag* *wuyanba-bu-ni-wa*
some 3pl/3pl-kill-PCOMPL-SUF
'They killed some (people).' (MERLAN 1982:96)

(55) *Ja-wuyan-wu-n* *gurdarrba* *ø-malam*
3-3sg/3pl-give-PRES anywhere M.ACC-people
'He gives things to anybody, he's generous in all directions.' (MERLAN 1982:118)

The most detailed discussion for an Australian language of the use of bound object markers in indefinite and non-referential contexts, is the discussion of Warlpiri by BITTNER and HALE (1995).[13] They demonstrate that, in combinations of free nominals with object clitics, both indefinite and definite interpretations are regularly available. The only exceptions are when the external nominal is a personal pronoun, demonstrative or name, in which case the interpretation is always definite (p. 89), but here, of course, the definiteness is attributable to the external material rather than the pronominal clitics. A crucial difference from the Bininj Gun-wok data is that when no external material is present the Warlpiri clitics are always interpreted as free pronouns, with definite referential interpretations. Various treatments of this system, e.g. SIMPSON (1991), therefore treat the object clitics as polysemous between full-pronoun and agreement-marker interpretations.

Although they are discussing clitics in second position, rather than prefixes to the verb, the same issues of obligatoriness arise, since the auxiliary (including suffixes for subject and object) must be present in all finite clauses with the exception of certain verbless predication structures. And although in cases of singular objects the fact that the form is zero (e.g. 56) might suggest an alternative analysis in which the object form is simply not there, their discussion includes non-zero examples in the form of non-singular object suffixes (e.g. 57), and a non-zero indirect object suffix for third person singular (58).

(56) *Kurdu* *ka-rna-ø* *nya-nyi*
child PRES-1sg-3sg see-NP
(a) 'I see a child.'
(b) 'I see the child.'
(c) 'I see him/her, who is a child.' (B&H ex. 18)

(57) *Panu* *ka-rna-jana* *nya-nyi*
many PRES-1sg-3pl see-NP
(a) 'I see a large group (of them).'
(b) 'I see the large group (of them).'
(c) 'I see them, who are a large group.' (B & H ex. 19)

(58) *Ngarrka-ngku* *ka-ø-ø-rla* *karnta-ku* *ngapa*
W man-ERG PRES-3sg-3sg-3sgIO woman-DAT waterABS

yi-nyi *mata-ku*
give-NP tired-ERG
(a) 'A/the tired man is giving some/the water to a/the woman.'
(b) 'The man, who is tired, is giving some/the water to a/the woman.' (B&H, ex. 6)

13 See also SIMPSON (1991:153) and HALE (1983:6) for briefer remarks.

As these examples show, both definite and indefinite readings are available regardless of whether the pronominal clitic is zero (as in (56)) or non-zero (e.g. 3 plural object *-jana* in (57) and 3 sg indirect object *-rla* in (58)). Although BITTNER & HALE's main concern is with the interpretation of the whole referring expression (i.e. the pronominal clitics plus the co-referring nouns and adjectives), the point of interest here is that the pronominal clitics are not restricted to occurring in definite contexts.

It is clear, then, that the pattern in Bininj Gun-wok is far from unique, and that several other Australian languages with obligatory pronominal affixes for object – whether to the verb (Warray, Mangarayi), or to a second-position 'auxiliary' (Warlpiri) – allow a range of indefinite or generic interpretations.

Instead of seeing this semantic latitude as a corollary of obligatory pronominal affixation to the head in these languages, could this instead be related to some special features of discourse structure and referential practice in these languages? One could argue, for example, that given the 'broadcast model' of communication found in Australian Aboriginal communities (WALSH 1991), with the freedom it gives interlocutors to listen or not as they wish, speakers in such communities do not expect to monitor the set of intersubjective givens that determines decisions about definiteness and other aspects of reference. On this argument, the phenomenon outlined would be derived from pragmatic norms rather than typological features of these languages.

I believe there are two reasons to reject this view.

Firstly, if it were true it should apply equally well to free personal pronouns, but these, as we saw above, are basically comparable to free personal pronouns in English, being incompatible with the indefinite, non-referential, generic or token-replacement uses discussed above for pronominal affixes.[14]

Secondly, the phenomenon is not confined to Australian languages but is attested in typologically similar languages from other parts of the world. We now briefly consider three such languages – Lummi, Mohawk and Georgian – in which, like Bininj Gun-wok, pronominal affixes or obligatory clitics representing objects can receive either a definite or an indefinite interpretation.

For the Salishan language Lummi, JELINEK (1993:12-13), mentions that absolutive third person clitics to the verb may receive either a definite or an indefinite reading, even though the corresponding third person ergative suffixes must be definite:[15]

---

14 This is not to say that they are functionally equivalent. In particular, their optionality means they play a role in signalling discourse organization, for example by re-mentioning protagonists after paragraph boundaries.

15 JELINEK goes on to argue that 'on the definite reading, the adjunct is a topic; on the indefinite reading, it is predicative, corresponding to an adjoined relative'. This argument is irrelevant to the current issue and need not detain us here.

(59) *leŋ-t-s-ø* *cə* *ŋənə*
Lu see-TR-3ERG=3ABS DET child
'He_i saw him_j, the/a $child_j$.[* the/a child saw him]

In Mohawk, objects are marked in the pronominal prefix slot even when they are non-specific indefinites, as in the first verb of (60):

(60) *ne ónʌ* *érhar* *ʌ-hó-kʌ-'*
Mo when dog FUT-3MAsg/3MAsg-see-PUNC

ʌ-ho-tewékwʌ-'
FUT-3Masg/3MAsg-pet-PUNC
'Whenever he sees a dog he pets it.' (BAKER 1995:125)

And in Georgian (BOEDER, this issue, ex. 9), third person prefixes to the verb can be used for indefinites, at least in potentialis contexts:

(61) *romel-i* *x-e-e-ʒieb-n* *ṗ ov-i-s*
who-NOM 3IO-E:VERSION-seek-FREQ.PST find-FREQ.AOR-3sgSUBJ
'Whoever seeks something, finds it' ('he who seeketh findeth')
(Mathew 7,8; Old Georgian)

Although I lack the space to survey the topic thoroughly here – and many descriptions simply fail to provide sufficient detail to evaluate the exact range of the phenomenon – examples such as the above make it clear that Bininj Gun-wok is far from being unique in this regard, and that at least the indefinite interpretations of pronominal object affixes are found in a number of other head-marking languages.

This is not to say that such interpretations are an inevitable consequence of using object pronominal affixes, since a number of alternative strategies are available.

One is to reduce the transitivity of the verb in situations where the object is not definite, e.g. through the use of a middle or antipassive; examples are the Eskimo antipassive and (in some circumstances) the Cayuga semi-reflexive (e.g. *-kany-* 'bite someone' and its semi-reflexive form *-atę-kany-* 'bite (dispositionally), be someone who bites'; SASSE 1997). In Bininj Gun-wok, the avoidance of the *bi-* form with the non-referential objects of 'have' and (for some speakers) 'look for' (§4.2) could be analysed as the use of the intransitive prefix set, and hence of an intransitivizing strategy.

A second strategy is to have a special object form dedicated to non-specific objects; an example of this is the Aztec non-specific object prefix *te-/tla-* (ANDREWS 1975, LAUNEY 1981).

A third strategy is to have a conventionalized interpretation of some third person affix. One example is the use of the third person plural for indefinite objects in Warray (see

above). A second is the use of the third person feminine for unspecified human objects in Cayuga. Here one says *e-yá:kę-hs* [3sg.fem-go.out-HABIT] is ambiguous between 'she goes out' and 'one goes out', *khe-tęą:-s* [1sg/3sgfem-pity-HABIT] is ambiguous between 'I pity her' and 'I pity someone', and *shako-yé:na-hs* [3sg.masc/3pl.fem-grab-HABIT] is ambiguous between 'he grabs them (fem.)' and 'he grabs people'. (The last example has been lexicalized as the term for 'policeman').

A fourth strategy is to exploit the choice between incorporation and the use of pronominal affixes, e.g. by reserving nominal incorporation for indefinite or generic nouns. The result of this is that the although the argument must be marked on the verb in either case, the choice between representation by a lexical and a pronominal root comes to parallel the choice between noun and pronoun in an analytic language like English or Chinese. Again, (so-called) noun incorporation in Eskimo[16] can have this function, but it would be inaccurate to see this as always being a function of noun-incorporation, and indeed in many languages, such as Bininj Gun-wok, noun incorporation is more likely to signal givenness than indefiniteness. But note that in Bininj Gun-wok, unlike in Eskimo or Aztec, there is no incompatibility with representing an argument by both incorporated noun and pronominal affix.

Most polysynthetic languages appear to employ a mix of the above strategies with a tolerance for at least some non-definite uses of pronominal affixes. Thus in Aztec the regular object prefix *k-* rather than the non-specific prefix *te-/tla-* is used for objects of negatives such as 'I don't see anyone' (62a, from LAUNEY 1981:47) and indefinite pronouns like 'anything', as in 'did you see anything' (62b, from LAUNEY 1981:249), as well as for indefinite nominal expressions (63a,b from LAUNEY 1981:66):

(62a) *ay-āc* *ni-qu-itta*
not-someone 1sg-3sg-see
'I don't see anyone.'

(62b) *cuix* *i-tl-a'* *ō-ti-qu-itta-c*
Q ø-something-AFF PST-2sg-3sg-see-PERF
'Did you see anything?'

(63a) *ni-qu-itta* *cal-li*
1sg-3sg-see house-ABS
'I see (some) houses.'

16 Although the term 'incorporation' is often applied to Eskimo, strictly speaking it is not incorporation proper, but rather the highly productive formation of denominal verbs from nominal roots. We avoid going into this question here.

(63b) *ni-qu-itta* *cē* *cuahui-tl*
1sg-3sg-see one tree-ABS
'I see a tree.'

And in Greenlandic Eskimo indefinite objects in subordinate clauses are treated in a number of ways, including incorporation and an intransitive case frame with only the subject marked on the verb (64), and a transitive case frame with both subject and object marked on the verb by third person obviative pronominal affixes (65):

(64) *Nukappiaraq* *balloni-si-gaannga-mi*
boyABS balloon-get-whenITER-3sgPROX

qaartuur-tar-p-a-a.
break-quant.scope-INDIC-[+tr]-3sg/3sg
'When a boy gets a balloon, he breaks it.' (BITTNER 1995:64)

(65) *Piniartu-p* *puisi* *pisara-annga-gu*
hunter-ERG sealABS catch-when.iter-3sgOBV/3sgOBV

nuliata *amirlanir-tigut* *pilat-tar-p-a-a.*
wife3sgERG most-QUANTIF flense-quant.scope-IND-[+tr]-3sg/3sg
'When a hunter catches a seal, his wife usually flenses it.' (BITTNER 1995:68)

Such cases show that polysynthetic languages are not totally lacking in ways of distinguishing definiteness and referentiality levels in objects, since one or more of the above strategies may be used to show such distinctions. Rather, it seems to be the case that the semantic threshold at which representation is made by pronominal affix is lower than it is for free pronouns in polysynthetic or more analytic languages alike.

A promising topic for future research on polysynthetic languages – which will depend on having much more precise descriptions of their use of pronominal affixes– will be the question of whether there are systematic patternings in the threshold and range of situations under which pronominal affixes are used.

At the same time, it is important to relate these semantic restrictions to the point reached on the grammaticalization path from free pronouns to syntactically optional pronominal clitics to obligatory affixes.[17] Optional pronominal clitics in many languages, unlike

17 It is important to distinguish the dimension of obligatoriness from that of clitic vs affixal status, since these need not correlate: Warlpiri pronominal clitics are obligatory in finite clauses, and in this respect subject to the same semantic restrictions placed on closed classes of obligatory pronominal affixes, whereas Romance clitics, e.g. in Italian, are not obligatory, at least partially

pronominal affixes, are typically incompatible with indefinite interpretations. In Italian, for example, clitic pronouns are essentially definite. CINQUE (1990:75-6) discusses alternate readings of *qualcosa* 'something' and *qualcuno* 'someone'; only when these have referential readings, as in (66), is the pronominal clitic is obligatory:

(66) A: *Li* *conosci,* *quelli?*
3pl.MA know:2sg those:MA
'Do you know them, those people?'

B: *Sì,* *qualcuno,* **(l')* *ho* *già* *conosciuto.*
yes someone 3MA have:1sg already know:PST.PART
'Yes, I've already met one of them.'

Although one would expect there to be a correlation between increasing obligatoriness, on the one hand, and a lowering of the threshold for use of pronominal clitic-cum-affixes on the other, the existence of alternate strategies such as those outlined above undoubtedly complicates the relationship. There are also questions of direction of causation: does morphosyntactic obligatoriness lead to semantic reanalysis, or must there be an increased generality of use in the clitic stage use (such as extension to generic and indefinite uses) before the language can become truly polysynthetic rather than simply using reduced clitics?

In fact, the second alternative seems more likely on both theoretical and empirical grounds. The theoretical argument is essentially definitional: if argument affixes cannot be used for non-definites, then argument affixation cannot be obligatory, so the language will not meet the criteria for being considered polysynthetic.[18] The empirical argument is based on the existence of languages which extend pronominal clitics to generic and (some) indefinite uses, as in Modern Greek and Albanian.[19] In modern Greek, where generic Nps are definite, they are copied with a pronominal clitic. Thus one says 'I like-them the flowers' for 'I like flowers' (67), and 'He's always after-them the women' for 'He's always after women' (68). Albanian takes the use of object clitics further, to cover indefinite objects. The pronominal clitic is obligatory with definite direct objects, as in (69), and with indefinite direct objects it is obligatory when the focus is on some constituent other than the object itself (70). However, it does not occur with 'conflated' non-referential objects, i.e.

---

form part of a contrast set with free pronouns - e.g. in the case of stressed postverbal pronouns such as *ho visto LUI* 'I saw HIM' - and hence will contract a different set of semantic oppositions.

18 Though of course if the relevant affixes are represented by zeroes one can argue that they are absent, rather than present as zeroes in such cases; this leaves many polysynthetic languages in the position where they do not furnish decisive evidence either way.

19 I thank Hans-Juergen Sasse for bringing these languages to my attention, and for supplying the relevant examples.

with the nominal part of complex predicates with light verbs, as with 'take marking' for 'take notes' in (71).

(67) *ta* *aɣapao* *ta* *luluðia*
3.NE.pl.OBJ love:1sg the.NE.pl flower:pl
'I like flowers.'

(68) *panda* *tis* *kiniɣai* *tis* *yinekes*
always 3.FE.pl.OBJ chase3sg the.FE.pl woman:pl
'He's always chasing women.'

(69) *E* *hapi* *derën.*
3sg.FE.OBJ open.3MA.sg.pst door:ACC(FE.sg.)
'He opened the door.'

(70) *Atëherë po* *e* *dredh dhe* *unë një* *cigare.*
then PROGR 3sg.FE.OBJ roll also I a cigarette(FE.sg.acc)
'Then I'll roll a cigarette too.'

(71) *Marr* *shënim*
make1sg marking
'I take a note.'

The Modern Greek and Albanian examples just given suggest that the pronominal encoding of generic and indefinite objects is possible in argument-clitic languages as well as polysynthetic languages. This in turn suggests that the extension of object marking to certain types of non-definite object may in fact precede the integration of such markers as obligatory elements in a polysynthetic word. Note, though, that in both Modern Greek and Albanian there are still restrictions on the use of these clitics: in Modern Greek the construction in (67) and (68) can only be used when the focus is on the verb, and in Albanian the use of clitic pronouns for indefinite objects is not possible when the focus is on the object itself.

Ultimately, this issue can only be decided by constructing a delicate implicational hierarchy showing encoding strategies for objects of different referential status in a reasonable sample of both pronominal-clitic and argument-affix languages, and by careful diachronic studies where this is possible. Such studies will be crucial in understanding the exact sequence of steps by which a language becomes fully polysynthetic.[20]

20 A recent article by SIEWIERSKA (1999), which appeared in the same year as the original version of this paper, concludes that what she terms 'grammatical agreement' of objects is 'cross-linguistically very rare' (and in fact her sample of 272 languages failed to contain any examples,

6. Conclusion

In its treatment of object pronominal affixes Bininj Gun-wok is representative of a large set of polysynthetic languages: the verb plus its argument affixes is semantically more specified than a verb in a non-polysynthetic language (e.g. English), but less specified than an English verb plus the relevant free personal pronouns. A form like *kabimang*, therefore, lies, in its semantics, somewhere between 'marries' and '(s)he marries him/her', a fact that the interlinear gloss '3sg marries 3sg' captures to the right degree of semantic specificity. Accounts of polysynthetic structure which simply see polysynthetic languages as verbs whose arguments are directly represented on the verb are therefore inaccurate, and will be driven to forced and unnatural accounts for a whole range of construction types in which object affixes do not correspond to free personal pronouns.[21] Baker (1995), for example, seeks to account for some of these differences by appealing to the interaction of movement rules with different phrase structures, despite the notorious problems in finding clear evidence for phrase structure in polysynthetic languages – see e.g. Heath 1986 on the impossibility of finding evidence for NP or VP units in Nunggubuyu. At the same time, accounts like Baker's do not deal directly with the different semantics of the affixes themselves.

---

under her definition). However, in comparison with the definitions I employ in this paper, she only counts, as 'grammatical agreement', cases where the affix 'must necessarily be accompanied by overt nominal or pronominal agreements'; if there is no constraint on whether overt nominal or pronominal agreement can occur or not she considers this a case of 'ambiguous agreement'.

To my mind this is a methodological mistake, since it conflates two logically independent (though of course correlated) parameters – (a) loss of referentiality on the part of the affix, and (b) obligatory presence of the nominal argument with which it agrees. I would argue that (a) is the key fact about affixal agreement and the semantics of the relevant affixes, whereas (b) concerns issues of how acceptable it is to omit overt arguments, and is logically independent of (a) (as witness the many languages, such as Japanese or Lao, which lack agreement but permit the omission of overt arguments). Sierwierska herself admits that the two factors are potentially independent (pp. 227-8), and then goes on to make the following statement, which coincides with the observations I make in this paper: 'grammatical agreement, i.e. ambiguous or grammatical agreement in my typology, should not be dependent on the informational status of arguments' (p. 230).

However, because of her over-restrictive definition of grammatical agreement, it is difficult to interpret exactly what she means when she says there are no examples of grammatical agreement with objects, since it is unclear whether they have been excluded by both (a) and (b), just (a), or just (b). If only (a) is taken as the relevant criterion, as I have argued should be the case, a number of the languages discussed in this paper provide examples showing these cells are not empty.

21 AUSTIN & BRESNAN (1994:234-7) make similar points with respect to the problems JELINEK's (1984) 'pronominal-argument' analysis of Lummi has in accounting for the availability of indefinite interpretations of third person object clitics.

What is needed are formalisms capable of representing the appropriate level of semantic specificity contributed by the pronominal affix, then (where necessary) unifying this with relevant information contributed by external nominals. It seems likely that no polysynthetic language has, within its system of pronominal affixes, a set of contrasts large enough to make all the distinctions in definiteness, indefiniteness, genericity and referentiality which can be made, in free nominal expressions, by various combinations of articles, demonstratives, free pronouns, and adjectives representing discourse status. While some languages, such as Aztec with its 'unspecified object' affix, introduce some distinctions into the verbal morphology, I know of no language able to convey more than two values on this semantic dimension by its agreeent morphology alone.[22] Instead, polysynthetic languages have a complex set of strategies that involve, on the one hand, the unification of material from external nominals in construction-specific ways, and on the other, the signalling of changes in transitivity. Dual-structure approaches such as LFG, which are able to directly represent the projection of agreement information by the verb at the level of functional structure, then merge it where relevant with further information contributed by external nominals, are thus better placed to handle the different division of semantic labour that polysynthetic languages make between argument affix and free nominal and pronominal material – see NORDLINGER 1998 for an analysis in this framework.

From the point of view of descriptive and typological work on polysynthetic languages, there is a need for much more fine-grained analysis of the range of indefinite and non-referential uses available to argument affixes on verbs. Even Bininj Gun-wok, which as we have seen allows object prefixes for a range of indefinite, generic and non-referential objects, draws the line at using them for non-referential psychological objects and objects of the verb 'have', and resorts to the intransitive prefix set here. I have suggested that, in general, polysynthetic languages use argument prefixes for a wider range of situations than warrant the use of free pronouns, but that as one moves from definites to indefinites towards non-referential objects every polysynthetic language will draw the line at some point and use an alternative strategy, such as intransitivization. Once we have better data from a range of polysynthetic languages it will be interesting to see whether these cut-offs can be arranged in an implicational hierarchy.

In general, discussions of polysynthetic typology tend to stress the self-sufficiency of the exotic 'verb-sentence'. However, because closed systems are involved in the relevant verbal morphology, much of the fine-tuning information used in discourse management still comes from nominal material. We will only reach a full understanding of how polysynthetic languages function when we start paying closer attention to the way in which nominal and verbal material get integrated in particular constructions. And, paradoxically, the more unusual the behaviour of the verbal object affixes is, the more we are sent back to look at the very normal-looking contribution of external nominal and pronominal material.

---

22 Though it is possible that some of the contrasting third person pronominal affixes in Athabaskan languages could be treated as giving more than two choices here.

ABBREVIATIONS

Note the following abbreviations used in glosses. In glossing subject/object combinations, the person and number of the subject is given first, then a '/', then the person and number of the object, e.g. 1/3pl '1st person singular subject acting on 3rd person plural object'. If no indication of number is given overtly for a pronominal prefix, e.g. 1 in this case, this is to be interpreted as singular, but for non-human arguments this does not exclude non-singular interpretations. Persons are glossed as 1 'first person', 2 'second person', 3 'third person' and 12 'first person inclusive'. Other abbrevations are ABL 'ablative', ABS 'absolutive', ACC 'accusative', AFF 'affirmative', AOR 'aorist', CONJ 'conjunction', DAT 'dative', DEM 'demonstrative', DET 'determiner', EMPH 'emphatic', ERG 'ergative', fem 'feminine', FREQ 'frequentative', FUT 'future', h 'higher object, i.e. human, or higher animate acted upon by lower animate or inanimate' (see §3 for details), IMM 'immediate', indic 'indicative', IRR 'irrealis', IVF 'incorporating verb form', l 'lower object, i.e. lower in animacy than humans', ITER 'iterative', LOC 'locative', M.ACC 'masculine accusative', NEG 'negative', NP 'non-past', NOM 'nominative', IO 'indirect object', OBJ 'object', OBL 'oblique', OBV 'obviative', P 'past', PCOMPL 'past completive', PERF 'perfective', PI 'past imperfective', pl 'plural', PROGR 'progressive', PP 'past perfective', PRES 'present', PROX 'proximal', PST.PART 'past participle', PUNC 'punctual', Q 'question', QUANTIF 'quantifier', REDUP 'reduplication', REL 'relative demonstrative' (which introduces relative clauses, but is also more generally used to introduce new mentions which the hearer is expected to be able to identify easily), RR 'reflexive/reciprocal', sg 'singular', SUBJ 'subject', SUF 'suffix', TR 'transitive'. Roman numerals I, II, III and IV refer to four noun classes, generally congruent in semantic content to the genders glossed by MA (masculine), FE (feminine), VE (vegetable) and NE (neuter) respectively.

REFERENCES

ANDREWS, J. RICHARD. 1975. *Introduction to Classical Nahuatl.* Austin: University of Texas Press.

AUSTIN, PETER & JOAN BRESNAN. 1994. Non-configurationality in Australian Aboriginal languages. *NLLT* 14:215-68.

BACH, EMMON, ELOISE JELINEK, ANGELIKA KRATZER & BARBARA H. PARTEE, eds. *Quantification in Natural Languages*. Dordrecht: Kluwer Academic.

BAKER, MARK. 1996. *The Polysynthesis Parameter.* Oxford University Press.

BERNDT BERNDT, C.H. & BERNDT, R.M. 1951. An Oenpelli monologue: culture contact. *Oceania* 22.24-49.

BITTNER, MARIA. 1995. Quantification in Eskimo: a challenge for compositional semantics. In Bach et al, pp. 59-80.

BITTNER, MARIA & KEN HALE. 1995. Remarks on definiteness in Warlpiri. In Bach et al, pp. 81-106.

BOEDER, WINFRIED. (this vol.) A slot-filling constraint in the Georgian verb and its syntactic corollary.

BRESNAN, JOAN AND SAM A. MCHOMBO. 1987. Topic, pronoun and agreement in Chichewa. *Language* 63:741-782.

CARROLL, PETER. 1976. *Kunwinjku (Gunwinggu): a language of Western Arnhem Land.* ANU: unpublished Master's thesis.

CARROLL, PETER. 1995. *The old people tolds us. Verbal art in Western Arnhem Land.* Unpublished Ph.D. Thesis, University of Queensland.

CINQUE, GUGLIELMO. 1990. *Types of A-bar dependencies.* Cambridge, Mass.: MIT Press.

COOPER, ROBIN. 1979. The interpretation of pronouns. In F. Heny and H. Schelle, eds. *Syntax and Semantics*, vol. 10: *Selections from the Third Groningen Round Table*, pp. 61-92. New York: Academic Press.

ETHERINGTON, S. & N. 1994. *Kunwinjku Kunwok. A short introduction to Kunwinjku language and society.* Kunbarlanja: Kunwinjku Language Centre.

EVANS, GARETH. 1980. Pronouns. *Linguistic Inquiry* 11:337-62.

EVANS, NICHOLAS. 1995a. A-quantifiers and scope in Mayali. In Emmon Bach, Eloise Jelinek, Angelika Kratzer and Barbara Hall Partee, eds, *Quantification in natural language.* Dordrecht: Kluwer. Pp 207-270.

EVANS, NICHOLAS. 1995b. The syntax and semantics of body parts in Gun-djeihmi. In H. Chappell & W. McGregor, eds., *The grammar of inalienability.* Berlin: Mouton de Gruyter. Pp 65-109.

EVANS, NICHOLAS. 1997. Role or cast? Noun incorporation and complex predicates in Mayali. In Alex Alsina, Joan Bresnan & Peter Sells, eds., *Complex Predicates.* Stanford: CSLI. Pp. 397-430.

EVANS, NICHOLAS. 1999. Why argument affixes in polysynthetic languages are not pronouns: evidence from Bininj Gun-wok. *Sprachtypologie und Universalienforschung* 3/4:255-81.

EVANS, NICHOLAS. Forthcoming. *A pan-dialectal grammar of Bininj Gun-wok: Mayali, Kunwinjku and Kune.* Canberra: Pacific Linguistics.

GARDE, MURRAY. 1996. *'Saying nothing': the language of joking relationships in Aboriginal Australia.* Grad. Dip. Thesis: Discipline of Anthropology, Northern Territory University.

GARDE, MURRAY. 1997. *Preliminary dictionary of Eastern Kunwinjku (Kuninjku).* Machine-readable file deposited at ASEDA File, Australian Institute of Aboriginal and Torres Strait Islander Studies.

HALE, KEN. 1983. Warlpiri and the grammar of non-configurational languages. *NLLT* 1:5-74.

HARVEY, MARK. 1986. *Ngoni Waray Amungal-yang: the Waray language from Adelaide River.* Unpublished MA Thesis, Australian National University.

HAUSSER, R. 1979. How do pronouns denote?. In F. Heny and H. Schelle, eds. *Syntax and Semantics*, vol. 10: *Selections from the Third Groningen Round Table*, pp. 93-139. New York: Academic Press.

JELINEK, ELOISE. 1984. Empty categories and non-configurational languages. *NLLT* 2:39-76.

JELINEK, ELOISE. 1993. Definiteness and second position clitics in Straits Salish. Paper Delivered to Workshop on Second Position Clitics, Linguistic Institute, Ohio State University, July 1993.

LAUNEY, MICHEL. 1981. *Introduction à la langue et à la littérature aztèques.* Paris: L'Harmattan.

LAUNEY, MICHEL. This volume. Compound nouns vs incorporation in Classical Nahuatl.

MCKAY, GRAHAM. 1975. *Rembarrnga, a language of Central Arnhem Land.* Unpublished Ph. D. Thesis, ANU.

MERLAN, FRANCESCA. 1982. *Mangarayi.* North-Holland: Lingua Descriptive Series.

MERLAN, FRANCESCA. 1983. *Ngalakan grammar, texts and vocabulary.* Canberra: Pacific Linguistics.

MERLAN, FRANCESCA. n.d. *Jawoyn dictionary.* Typescript.

MITHUN, MARIANNE. 1986. When zero isn't there. *BLS* 12:195-211.

NGANJMIRA, NAWAKADJ. 1997. *Kunwinjku spirit.* Melbourne: Miegunyah Press.

NORDLINGER, RACHEL. 1998. *Constructive case. Evidence from Australian languages.* Stanford: CSLI.

QUINE, WILLARD V.O. 1960. *Word and object.* Cambridge, Mass: MIT Press.

REINHART, TANYA. 1983. *Anaphora and semantic interpretation.* London: Croon Helm.

SASSE, HANS-JÜRGEN. 1991. Predication and sentence constitution in universal perspective. In D. Zaefferer (ed.) *Semantic Universals and Universal Semantics*. Berlin and New York: Foris.

SASSE, HANS-JÜRGEN. 1995. The evolution of person prefixes in the Iroquoian languages and its functional background. In Michael Dürr, Egon Renner & Wolfgang Olischinski, eds.: *Language and Culture in Native North America. Studies in Honor of Heinz-Jürgen Pinnow. Lincom Studies in Native American Linguistics 2*. München, Newcastle: Lincom. Pp. 199-210.

SASSE, HANS-JÜRGEN. 1997. *Cayuga: a grammatical sketch.* MS: Institut für Sprachwissenschaft, Cologne.

SIEWIERSKA, ANNA. 1999. From anaphoric pronoun to gramamtical agreement marker: why objects don't mark it. *Folia Linguistica* 23.2:225-251.

SIMPSON, JANE. 1991. *Warlpiri Morpho-syntax. A lexicalist approach.* Dordrecht: Kluwer.

VAN GEENHOVEN, VEERLE. 1998. *Semantic incorporation and Indefinite Descriptions: Semantic and Syntactic Aspects of Noun Incorporation in West Greenlandic*. Stanford: CSLI Publications.

WALSH, MICHAEL. 1991. Conversational styles and intercultural communication: an example from Northern Australia. *Australian Journal of Communication* 18.1:1-12.

How referential is agreement? The interpretation of polysynthetic dis-agreement morphology in Ngalakgan[1]

BRETT J. BAKER

With respect to argument affixes in polysynthetic languages, authors (e.g. JELINEK 1984, M. BAKER 1996, SIMPSON 1991) have generally taken one of two positions. Either these affixes should be regarded as agreement markers, or as pronominal arguments ('anaphors'). In this paper I argue that there is a three-way division in the morpho-syntactic and referential behaviour of argument prefixes in Ngalakgan: bound anaphoric pronouns, agreement affixes, and a third category which cannot be properly characterised either as an agreement marker nor as an anaphor. I call this category 'pronominal generic affix'. Referentially, the generic affixes have affinities with incorporated generic nouns, and need not agree with a coreferential argument. In the case of Ngalakgan, the question of whether argument affixes constitute 'agreement' markers or 'anaphors' is unanswerable. We must instead ask to what extent such affixes ever constitute a referentially homogenous class.

1. Introduction

A defining characteristic of polysynthetic languages is the presence of verb affixes indexing (usually) up to two arguments of a transitive verb, and sometimes three arguments of a trivalent verb. Controversy has been long-standing over whether such affixes constitute arguments of the verb, or whether they merely agree with the nominal arguments, which may be covert (e.g. JELINEK 1984, SIMPSON 1991, M. BAKER 1996, EVANS (THIS VOLUME)). In such discussions, it is usually the case that affixes are regarded as an undifferentiated group, with

1 Many thanks to Roy *Golokgurndu,* Splinter *Gerrepbere* and Doreen *Nyulpbu* for their tireless instruction in all things Ngalakgan; and to Eva Endrey-Walder, Mark Harvey, Andrew Ingram, Chris Manning, and Jane Simpson, for help with material herein: but they should not be held accountable for the views expressed. Except where noted, all material comes from my own fieldwork. Examples taken from MERLAN (1983) are notated as [M:*n*] where 'n' is the example number in that work.

little or no attention paid to the differences between, for example, those indexing first and second persons as against third persons.[2]

My purpose in this paper is to show that affixes which constitute a class morphologically nevertheless must be examined for their morpho-syntactic and referential characteristics individually. Upon such an examination of the argument affixes of Ngalakgan, we find that they do indeed have distinct characteristics; these are examined in §2. Argument affixes differ in the degree to which they are obligatory in a clause (§2.1). They also differ in their ability to occur with co-referential external nominals (§2.1), as well as under the scope of 'indefiniteness'- or 'non-specificity'-introducing operators such as question words and negation (§2.2). In the case of the argument prefixes indexing referents from the two inanimate genders or noun classes (vegetable and neuter), argument prefixes differ in the degree to which they must strictly agree with a co-referential nominal (§2.3). Since the 'dis-agreement' phenomenon in Ngalakgan is interesting for theories of the function of bound agreement morphology, I examine it in some detail in §3. I conclude that the inanimate argument prefixes in Ngalakgan serve to delimit the construal of a noun independently of the nominal expression itself.

Such evidence points to a three-way division of argument affixes in Ngalakgan, and hence forces a reformulation of the question of whether argument affixes constitute arguments or agreement. Rather, we must ask '*Which* affixes constitute arguments, and which constitute agreement?[3] And are there any unifying factors between languages?' I suggest in conclusion that there are some discernible 'universal' tendencies of reference in pronominal paradigms.

In what follows I initially describe the characteristics and distribution of the argument affixes of Ngalakgan (§1.1), before moving on to examine their behaviour in 'reference-delimiting' constructions (§2).

1.1 The paradigm of argument prefixes in Ngalakgan

Ngalakgan is an Australian language spoken by a handful of old people in the middle Roper river area of southern Arnhem Land.[4] It belongs to the Gunwinyguan (GN) family-level subgroup (ALPHER, EVANS, and HARVEY, forthcoming), perhaps the largest and best described of the non-Pama-Nyungan or 'Northern' language families, many of whose members, including Ngalakgan, are polysynthetic.

---

2 Although BRESNAN and MCHOMBO (1987) differentiate between subject and object pronominal affixes in Chichewa.

3 This is the way the question is framed in Generative linguistics; cf. for example Nordlinger (1998:59).

4 The language was first described by MERLAN (1983).

Ngalakgan verbs take prefixation for a maximum of two core arguments which I will term 'subject' and 'object'.[5] Table (1) sets out the person/number/noun class categories of Ngalakgan verb agreement, using the monovalent prefixes to intransitive verbs. Note that third person singular *animate* subjects are represented by Ø in the table.[6]

Table (1): Intransitive prefixes

1st singular	*ngu-*	1st exclusive plural	*yirri-*
1st inclusive dual	*yi-*	1st inclusive plural	*ngurru-*
2nd singular	*nginy-*	2nd plural	*nurru-*
3rd non-plural	*Ø-*	3rd plural	*burru-*
NEUT	*gu-* ~ *Ø*		
VEG	*mu-* ~ *Ø*		

I will provisionally refer to these prefixes as 'argument prefixes'; this is to be interpreted as 'prefixes which index arguments'. Until the conclusion, I will remain non-committal as to whether these morphemes are pronominal arguments or agreement markers.

Transitive prefixes in some cases are clearly segmentable into two component prefixes from Table (1), as in *yirr-mi-* (1pS-VEG), *ngurr-bu-* (12pS-3p-), particularly those where one argument is a SAP (Speech Act Participant: indexing 1st or 2nd person) and the other is 3rd person.[7] The order of elements in prefixes is determined by a referential hierarchy such that SAP > 3pl > 3nonpl (where '>' means 'occurs before'). SAP objects are distinguished from SAP subjects with object allomorphs ending in an apical nasal *-N*: ***yirrirn****-bi-na+n* (1pO-3p-see+PR) 'they see us (excl.)' vs ***yirr****-bi-nan* (1pS-3p-see+PR) 'we (excl.) see them'.

---

5 In the case of trivalent verbs such as 'give', the object prefix is co-referent with what would be the 'dative' object in English, for example 'us' in 'he gave us some food'; see example (43) below.

6 The standard orthography for Ngalakgan is as follows: short and mostly voiced stops are written *b, d, rd, j, g* syllable-initially, and *p, t, rt, tj, k* syllable-finally and represent bilabial, apico-alveolar, apico-postalveolar, lamino-postalveolar and velar articulations respectively. Accordingly *pb, td, rtd, tjj, kg* represent geminate (long, tense, and voiceless) versions of the same stops. Glottal stop is written *h*. Nasals corresponding to the stops are *m, n, rn, ny, ng*, laterals *l, rl*, tap *rr* and labio-velar, retroflex and palatal approximants *w, r, y*. The digraph *nk* represents an alveolar nasal followed by a velar stop /nk/, whereas the homorganic cluster /ŋk/ is written *ngg*. Clusters are simplified orthographically: *nj* and *ntj* (by analogy with *nk*) represent the homorganic and heterorganic clusters /ɲc/, /nc/ respectively. Postalveolar clusters (only /ɳʈ/ and /ʈʈ/ are possible) are written *rnd, rtd*. The contrast between alveolar and post-alveolar apicals is neutralised in initial position in word-level morphemes (those separated by #, '-' or '='). Apicals in these positions are predictably postalveolar following a vowel (with or without an intervening glottal stop) and alveolar otherwise, subject to assimilation. These morpheme-initial apicals are written as alveolars *n, d, l*. Loanwords from Kriol or English are underlined in the examples *gu-di* 'tea'.

7 A list of abbreviations can be found at the end of the chapter.

I regard such transitive prefixes as synchronically analyzable into subject and object morphs partly because of the behaviour of speakers. In artificially slow speech, speakers divide words into their constituent morphemes. An example is given here.[8]

(1) *yirrirn-bi-bak-wotj+ma*[9] slow speech: [yírɨɳ]^[bí:]^[bák]^[wɔ́cma]
1pO-3p-APPL-steal+[get]PR (normal speech: [yɪrɨɳbibakwɔ́cma]
'they always steal from us' [2/7/96:2A]

Note that speakers do not break words up into syllables or feet; rather, the word is divided into prosodic words according to the word-level morpheme breaks (see B. BAKER 1999). There are numerous similar examples from the elicitation sessions. It is for this reason (among others) that I treat complex prefixes such as *yirrirn-bi-* as two separate morphemes (albeit subject to allomorphy and ordering rules) rather than as portmanteau prefixes.

Prefixes indexing two SAP arguments are unanalyzable synchronically (and speakers do not treat them as complex in artificially slow speech). These prefixes will be represented as portmanteaux with a gloss indicating subject on the left and object on the right, separated by a slash: *yini-* (2/1[pl]) '2nd person subject, 1st person object, one of which is plural'.[10]

As with intransitive prefixes, the transitive prefix indexing 3rd person singular *animate* subject or object is Ø-. In the simple Present positive, Present negative and Future positive tenses, a verb where the only argument(s) are 3rd person (animate or inanimate) non-plural must take a prefix *gu-* (MERLAN 1983).[11] Inanimate arguments count as 3rd person non-plural for this rule and must co-occur with it. The *gu-* prefix is distinct from the homophonous prefix indexing Neuter arguments: *gu-gohje gu-ngolonggoh* ***gu-gu****-janga+n*

---

8 The symbol '^' represents a brief pause.

9 The plus symbol links morphemes at the 'root'-level (which is lexicalised), the hyphen links those at the 'word'-level, which is productive; see BAKER (1999). The example given, *wotj+ma* 'steal', here is a root-level compound verb, where the second root is identifiable as the verb root 'to get'. Because the meanings of such lexicalised forms are not always compositional, I enclose the meaning of the second root in brackets, indicating that such compounds always inflect as a member of the paradigm of the second root, but do not always reflect the meaning of that root.

10 Indeterminacy of plural association is characteristic of portmanteau morphs, as opposed to those which are analysable at some level of abstractness. For instance, the prefix *yirrirn-* which realises a 1st person exclusive plural object, is analysable into *yi-* '1st person exclusive' *rri-* 'plural' and *-rn* 'object'. But various aspects of this analysis make it unfeasible from a synchronic point of view. For instance, *yi-* is the 1st person *inclusive* prefix, and the object marker alternates between *-rn* and *-n* in a way which is not controlled by synchronically productive processes. Hence I regard such prefixes as synchronically frozen forms.

11 Prefixes with similar distributions are found in other Non-Pama-Nyungan languages, e.g. Bininj Gun-wok (and other Gunwinyguan languages) has a prefix *ga-* which is similar conditioned by NonPast tense and 3rd person (singular or plural) (EVANS forthcoming). An example of the use of NP *gu-* with a masculine subject is (40), below.

(NEUT-there NEUT-river.redgum NP-NEUT-stand+PR) 'a River Redgum is there'. In a sense then, the prefix *gu-* indicates that the subject (and object) is 3rd person non-plural, just in those tenses/moods.

Apart from the Ø prefix for non-plural animates, and the plural argument prefix, there are two further prefixes indexing 3rd person arguments on verbs: the two *noun class* (NC) argument prefixes. The nouns of Ngalakgan are distributed without exception into four noun classes. Certain classes of nouns - terms for kinship categories such as *ge* 'child of a man' and age-grade categories like *bolo* 'old person' - can take MASC or FEM prefixes depending on the biological sex of the intended referent: *nu-ge* 'man's son', *ju-ge* 'man's daughter', *nu-bolo* 'old man', *ju-bolo* 'old woman'. Aside from some well-defined exceptions such as these, all nouns are assigned lexically to one and only one noun class. The prefixes marking these classes, and a broad semantic characterisation of the membership is given in Table (2).

NC prefixes to nominals and nominal modifiers (demonstrative, adjective) show a three-way alternation: zero, 'short', and 'long' (the terms are from MERLAN 1983).[12] The overt forms are shown in Table 2. The long form prefixes are glossed DEF for 'definite' in this paper (the referential characteristics of the distinction are discussed elsewhere: B. BAKER 2000).

Table (2): Noun class forms and domains

	gloss	**indefinite** (short)	**definite** (long)	**domain**
'animate' classes	MASC	*nu-*	*nu-gu-*	male humans; animals of any sex
	FEM	*ju-*	*ju-gu-*	female humans and female animals
'inanimate' classes	VEG	*mu-*	*mun-gu-*	(edible) flora
	NEUT	*gu-*	*gun-gu-*	other inanimate (body parts, topography, weather/celestial, manufactured objects)

Table 2 gives a rough characterisation of the semantic domains which are covered by the four classes.[13] The membership of the two animate classes is overwhelmingly predictable,

12 Nominal modifiers and predicates do not always strictly 'agree' with their heads or arguments (respectively). NC argument prefixes on the verb may be realised with a different NC to their coreferential arguments, a process described in §2.3. This non-congruence of agreement I will call 'disagreement'. As mentioned briefly in §3, demonstratives may optionally neutralise agreement for non-feminine referents to masculine (see MERLAN 1983:75). Like head nouns, modifiers and predicates may also take zero prefixation in agreement.

13 The terms 'Vegetable' and 'Neuter' are not strictly accurate, since flora species (edible or not) appear in both classes, while the Vegetable class also include many non-vegetable referents such as body parts (*mu-ralh* 'hair'), body products (*mu-jele* 'urine', *mu-gurratj* 'blood'), implements (*mu-jet* 'stone oven'), natural mineral products (*mu-bim* 'white ochre', *mu-gapbarndah* 'white

while that of the two inanimate classes must be lexically specified. Since virtually all terms for animals (including insects, fish, and reptiles) and for humans are in the MASC and FEM classes, these classes can be confidently assigned to a macro-category 'animate', in distinction to another macro-category which does not contain such terms.[14]

NC argument prefixes have some morphophonemic differences from the other argument prefixes. Only the NC prefixes have the same form for all three argument roles transitive subject, transitive object, and intransitive subject. All other person/number categories minimally distinguish transitive subject and object roles, aside from the zero morph indexing 3 non-plural as a kind of default. (In the latter case, of course, the question is inapplicable.) Some examples showing the ambivalence of the NEUT and VEG prefixes are presented in (2). (2a) shows the VEG prefix *mu-* in object function, (2b) in intransitive subject function, and (2c) in A function.

(2) a. *gohye* *mu-may* *ngu-**mu**-goh+na+n*
here VEG-food 1sS-VEG-have+[see+]PR
'I've got some vegetable food here.' [2/9/97:1A][15]

b. *bo-wi* *mun-gu-dubal* *gu-**mu**-yerrert*
river-LAT VEG-DEF-*Nauclea.orientalis*[16] NP-VEG-grow
langa riba im gro thet dubal, dubal tri[17]
'Leichhardt trees grow by the river.' [2/9/97:1B]

---

clay'), terms for weather (*mu-makgurr* 'cold weather') and others. Nevertheless, as EVANS, BROWN & CORBETT (1998) point out, 'Vegetable' and 'Neuter' are more descriptive than the accepted terms 'class I', 'class II', 'class III', etc.

14 The only exceptions to this characterisation are the following. All terms for native bee species and their products are in VEG class. There is also a class of generic nouns which can be assigned to either NEUT or MASC class: *gu-/nu-gony* '(any) macropod', *gu-/nu-janggu* 'meat', *gu-/nu-ray* 'game', *gu-/nu-gerrnge* 'live body', *gu-/nu-jeny* '(any) fish'. Apart from this, there are two terms which have truly anomalous class marking: *gu-jarla* 'crayfish', and *gu-maramara* 'maggot' (the latter example is from MERLAN 1983).

15 Examples taken from fieldnotes are glossed with a date, tape number and side (A or B). Examples taken from texts are glossed with the text abbreviation (KD, DP, DD, Txt 1/6/96) and (for large texts) the line number.

16 The long NC prefixes which I have glossed as 'definite' are used (inter alia) for references to 'kinds' in characterising statements like this one. See §2.2.1.

17 For some of the examples, I have included the speaker's translation in Kriol. These translations are mainly included for the benefit of those who know the language, and for those people these translations give an insight into the speaker's intuitions about Ngalakgan constructions.

c. *mu-malba-yih* *mu-ngorh-mi+ny* *ngun-**mu**-bohbo*
VEG-ironwood-ERG VEG-fall-AUX+PP 1sO-VEG-hit.PP
'It was an ironwood tree (*Erythrophleum chlorostachys*)
that fell down and hit me.' [2/9/97:1B]

2. The morpho-syntactic characteristics of argument prefixes

In this paper, I argue that the argument prefixes do not constitute a homogenous class in morpho-syntactic or referential terms (though they are homogenous morpho-phonologically and prosodically). The following sections examine the characteristics of argument prefixes with respect to the following parameters:

(3) a. Obligatoriness
b. Referentiality
c. Agreement

'Obligatoriness' defines the degree to which argument prefixes are optional in any given clause, and conversely, the degree to which a co-referential independent nominal is permitted in the same clause. 'Referentiality' defines the degree to which argument prefixes are referential: Are their referents required to be definite, specific entities? Or can they be indefinite to various degrees? 'Agreement' defines the ability for argument prefixes to carry morpho-syntactic features which are distinct from, or 'clashing' with, those of a co-referential nominal expression.

The way that the space of prefixal reference is carved up by these parameters is schematised in Figure (1):

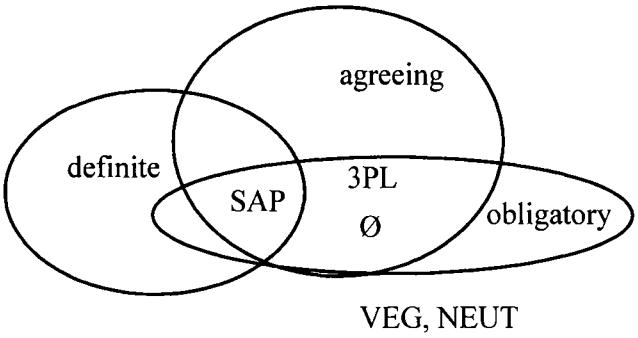


Figure (1): The intersection of three morpho-syntactic parameters with the space of prefix categories.

I address each of these parameters in turn, discussing the extent to which they are satisfied by the categories of argument prefixes proposed here: SAP, 3pl, and NC.

2.1 Obligatoriness

The argument prefixes differ in terms of obligatoriness. SAP argument prefixes are obligatory, as is the 3pl prefix when referring to humans; NC prefixes are optional. In (4), for example, there are two expressions of a first person subject, (a) with an argument prefix, and (b) with an independent pronoun. But (4b) is not a possible expression of the sense of (4a): an argument prefix is obligatory for the expression of SAP argument roles (word order does not determine the interpretation of argument-predicate relations in Ngalakgan).

(4) a. *manapburn* ***ngu-Ø-mang+iny***[18]
echidna 1sS-3p-get+PC
'I got an echidna.'

b. **ngaykgah-yih,* *manapburn* ***Ø-mang+iny***
I-ERG echidna get+PC

c. *ngaykgah-yih,* *manapburn* ***ngu-mang+iny***
I-ERG echidna 1sS-get+PC
'It was me that got an echidna.'

3rd person plural prefixes are also obligatory, at least for human referents which are plural. In this function of marking plurality for human arguments, the 3pl prefix cannot be omitted. In example (5), the sense of (a) cannot be expressed as in (b). Without a co-referential 3pl prefix, external nouns referring to humans and higher animates are interpreted as singular entities.

(5) a. *matjji* *marrawul* ***burru-nanga+niny,*** *nu-gu-mirpbarra*
indeed hungry 3pS-sit+PC MASC-DEF-child(ren)
'Because the **children** were hungry.'

b. *matjji* *marrawul* ***Ø-nanga+niny,*** *nu-gu-mirpbarra*
indeed hungry 3sS-sit+PC MASC-DEF-child(ren)
'Because the **child** was hungry.'

18 I omit zero morphs and their glosses, for the most part. Since zero occupies a paradigmatic gap in the agreement prefixes to verbs, the reader can assume that the lack of an overt (subject or object) argument prefix to a verb is to be interpreted as 3rd person singular. This is one reason why example (4b) is infelicitous.

The reason for the obligatoriness of the 3pl prefix is probably because of the danger of ambiguity: for most referents, the *only* available number morphology is provided by the third person plural prefix to verbs. Only kinterms can regularly take a plural suffix *-pbulu*.

Plurality of some higher animate referents also triggers the plural prefix. The category of higher animates includes minimally humans, but also dogs, kangaroos, and other large species including birds, as in (6).

(6) *wah-wangginy* ***burru****-wulup* *nu-gu-jerrk*
ITER-one 3pS-swim MASC-DEF-bird
'The birds are bathing one by one.' [12/7/96]

Inanimates, and some lower animates, do not typically trigger the plural agreement prefix (MERLAN 1983:89). Nevertheless, when explicitly quantified, as in example (7), lower animates such as terms for species of fish can still trigger plural agreement. Word reduplication as in (7), a strategy which is used infrequently, is the only means of explicitly indicating a count number for non-human animates apart from the two numeral quantifiers *wangginy* 'one' and *yapbanh* 'two', and the dual suffix *-pbirrah*.

(7) ***morlopborl=morlopborl=morlopborl*** *ngu-**bu**-meh+me*
catfish=catfish=catfish 1sS-3p-RED+get.PP

*ngu-**bu**-wakgiri+wo* *jajabarngh-gah*
1sS-3p-return.with+[give]PP afternoon-LOC
ketfij ketfij ketfij ai bin bringimbek aftanundaim
'I caught three catfish and brought them back in the afternoon.' [27/6/96:1A]

We could therefore regard the prefix I have termed '3rd person plural' instead as a prefix indexing '3rd person higher animate plural'. Or else we might hypothesise that Ngalakgan speakers regard inanimates and lower animates to be lacking in individuation, and that instead these kinds of referents are regarded collectively, rather like the way that we use Linnaean names (*'Acacia difficilis'*, *'Emydura australis'*) in English. I regard both of these interpretations as correct, though the reasons are not discussed in detail here (see B. BAKER 2000). In any case, the fact is that inanimates, and some lower animates (such as shellfish and insects), lack any kind of explicit number distinction morphology.

Normally, NC argument prefixes are overt in co-reference to an inanimate argument.

(8) a. ***gu****-ngolonggoh* *gu-gohje* *gu-**gu**-janga+n*
NEUT-river.redgum NEUT-there NP-NEUT-stand+PR
'A River Redgum is/some River Redgums are [standing] there.'

b. ***mu****-darda* *ngurr-**mu**-ma+nginy*
VEG-honey 12pS-VEG-grab+PC
'We collected some wild honey.'

But in contrast to the SAP prefixes, and the 3rd person plural prefix, NC prefixes are optional. For any verb having an inanimate argument the prefix indexing that argument may or may not occur, regardless of its argument role. In (9), the logical intransitive subject is *darh* '(any) tree' which is NEUT class. The predicate *ngorh* 'fall' is zero-marked for this subject.[19]

(9) *mu-yanah-ba* ***gun-gu-darh,*** ***Ø-ngorh-mi+ny***
VEG-what-INT NEUT-DEF-tree 3-fall-AUX+PP
'What kind of tree was it that fell?' [2/9/97:1B]

In (10), VEG class *mu-yawok* 'cheeky yam' (*Dioscorea bulbifera*) is the object of the transitive verb *jap+ga+*'soak in water', but there is no NC argument prefix expressing this object (this sentence is an extract from a short text where the reference to 'wild yam' was clearly to the same referent, hence the definite translation).

(10) *matjji* ***mu-yawok*** *ngu-Ø-jap+ga+na+gan*
because VEG-yam.sp 1sS-3-soak+[take+]FUT+PURP
wailjem, blanga jokim la woda
'...so I can soak the wild yam [in water]' [3/9/97:2A]

In (11), the VEG class noun *mu-rok* 'pandanus' is the ergative-marked subject of transitive *gulh-mi+* 'jab', but there is no VEG class prefix indexing this subject.[20]

(11) *yirrirn-Ø-ngarriny-gulh-gulh-mer+iny* ***mu-rok-yih***
1pO-hand-ITER-jab-AUX+PC VEG-pandanus-ERG
'The pandanus was jabbing our hands.' [DP: 7]

Again, the question of argument role (S, A or O) is irrelevant: inanimate prefixes are optional in any role.

I include here some brief remarks on independent pronouns. Independent pronouns are highly marked pragmatically. They cannot occur in the same intonational phrase (IP) as a verb inflected with a co-referential SAP prefix.[21] This means that the only expression of a first or second person argument possible *within* the IP containing the verb is that inhering in

19 In this example there is also an apparent feature 'clash' between the VEG class prefix on the question word *yanah* and the NEUT class noun *darh* that it questions. This phenomenon is discussed in §2.3 ff. I provide discussion in §3 for regarding this as clash rather than, say, neutralisation of noun class features.

20 See example (36) for an instance of overt agreement for NC agent.

21 This appears to be due to a restriction against *in situ* focus and topic in Ngalakgan. Items under the scope of focus or topic operators (which are null) are associated both with a distinctive intonation contour as well as a prosodic separation from the IP containing the matrix predicate, typically realised by potential for pause.

the argument prefixes. Therefore, I interpret independent pronouns as adjuncts. SAP pronominal prefixes fill the argument positions of the verb.[22]

In the case of 3rd person prefixes (including the inanimate prefixes), co-referential nouns *can* occur within the same IP, as in e.g. (6). On *prima facie* grounds then, 3rd person prefixes appear to act as agreement markers when there is a co-referential nominal in the same IP.

In summary, SAP prefixes are obligatory, as is the 3rd person plural prefix in reference to humans. The prefixes for the two inanimate noun classes by contrast are optional. Since the 3rd person animate singular prefix is zero, the question of obligatoriness makes no sense. In the following section, I show that all categories of argument prefixes can be co-referential to generic or 'kind' references, but only third person prefixes (3pl, NC, Ø) can have indefinite and non-specific readings, a possibility which is not open to SAP prefixes nor to anaphoric pronouns in English.

2.2 Interaction of argument prefixes with definiteness

EVANS (this volume) observes that in Bininj Gun-wok (BGW), a language closely related to Ngalakgan, 3rd person argument prefixes can agree with referents which are indefinite, nonspecific, or nonreferential in various ways. This is Evans' major evidence that argument prefixes are agreement markers, rather than anaphoric pronouns. The 3pl prefix and the NC prefixes in Ngalakgan can likewise both index an indefinite, nonspecific or nonreferential argument. A brief exemplification is presented here.

2.2.1 Generic and habitual readings of argument prefixes

All three types of argument prefix - SAP, 3pl, and NC - are found in habitual and/or generic contexts. Example (12) shows the use of NC prefixes indexed to inanimate 'kind' referents.

22 Independent pronouns are so rare in natural speech that at this point I cannot tell if 3rd person pronouns are permitted in the same IP with a co-referential verb. The non-feminine, non-plural form *ninjarnih* occurs just twice (in adjoining turns) in the data, but it is difficult to interpret its function in these two cases. The 3rd plural form *burrkgah* and the feminine form *jinjah* occur not at all. It is possible that all independent pronouns must occur under a distinct intonation contour. Given their function - to put contrastive focus on an argument - this requirement is unsurprising.

(12) a. *...warnwarn,* *gurninyarra*
Terminalia grandiflora, *Ficus racemosa*

b. *gu-baramulk*
NEUT-*Cucumis melo*

c. *gu-baramulk-yinji,* *gurnmarnh* *gu-**mu**-yerrert,*
NEUT-*Cucumis melo*-too, maybe NP-VEG-grow(PR)[23],

*gu-**mu**-yongo+n*
NP-VEG-lie+PR
thet baramulk tri im gro...la riba
'*Terminalia*, cluster fig, and wild cucumber grow [there], they are [there] (sc. 'by the river' from previous context)...'

d. *gu-wularra,* *mu-burrunburrun*
NEUT-?*Cucumis* sp. VEG-*Denhamia obscura*

e. *ngurrurndutj-yih gu-**mu**-ngu+n* *gurnmarnh* *gu-wularra*
emu(MASC)-ERG NP-VEG-eat+PR maybe NEUT-?*Cucumis* sp.
'Emus eat *Cucumis* and *Denhamia obscura.*'[24] [2/9/97:1B]

Apart from such 'kind' references indexing natural species, there are also examples in Ngalakgan of prefixes indexing human 'kind' references. In these cases, the 3pl prefix indexes the subject of a habitual verb.

(13) *bigurr* ***burru**-bordop-bordop-miny-gah*
person 3pS-DUR-cross-AUX.PP-LOC
thei bin olweis krosim, olpipul yuno? enibodi, thei bin olweis krosim that krik.[25]
'[that's] where people used to cross' [KD:171]

---

23 In Ngalakgan, the largest and only fully productive class of verbs takes a meaningless auxiliary *mi+* inflected for past tenses and reflexive/reciprocal: *yerrert-mi+ny* 'grow'-AUX+PP, *yerrert-mer+iny* ~ *-miy+iny* 'grow'-AUX+PC. The bare stem form, as in (12c) is interpreted as the simple present.

24 Note that there are two instances here (12c, e) which show 'disagreement' between a noun and co-occurring verb for NC, see §2.3 ff.

25 In this case the speaker has used Kriol *olweis* (< Eng 'always') together with past tense *bin*, to translate the habitual, indefinite interpretation of the Ngalakgan.

A similar example is presented in (14), again the 3pl prefix indexes a human 'kind' reference.[26]

(14) *mu-ngatjju* ***burr-bert*** *gu-ngoy-kgah*
VEG-cabbage.palm 3pS-roast NEUT-fire-LOC
'They (people in general) roast cabbage palm in the fire'; 'What people do with cabbage palm is to roast it in the fire.' [7/9/97]

There are also uses of the SAP prefixes in habitual sentences that are similar to the kind references exemplified here, as in (15).

(15) *gu-dirtda* *ngurr-yehye-gen* *je-nggorre-kgah*
NEUT-nosepeg 12pS-put-REL nose-12pDAT-LOC
'*Dirtda* is what [or 'the thing that'] we put in our nose.' [30/5/96:1A]

I return to the issue of which pronoun categories can be used in such sentences in §2.2.5.

2.2.2 'Presentational' clauses

I have no clear examples of a presentational construction with a human referent as subject. Non-human, but animate, arguments in presentational contexts show *zero* verb agreement. In (16), it was clear from context that the reference here is to more than one fish, yet the verb carries only the NP prefix *gu-*, which is obligatory for *non*-plural subjects.[27]

(16) *jeny* *yeh-yerreh* *bo-kgah* *gu-yongo+n,*
fish(MASC/NEUT) RED-down river-LOC NP-lie+PR

gurlagarl *gu-yongo+n* *yerreh* *bo-kgah*
big NP-lie+PR down river-LOC
'There are fish down in the river, big ones are down in the river.'
[2/9/97:1B]

---

26 I have no examples of kind references to humans in object function. However, MERLAN (1983:89, 164) cites the following example, where both object and subject are referentially unspecific:
bigurr-yih yirrirn-bi-nan-ji
man-ERG 1pO-3p-see-FNEG
'Men can't look at us.'

27 Given the optionality of overt agreement in coreference to nonhumans, we cannot be sure that zero agreement in this case (and others in this section) is due to referentiality factors (definiteness and specificity). Without more research I cannot fully determine the motivation for zero agreement at this point.

We can contrast example (16) with example (6), where a nonhuman, animate but definite referent triggers plural agreement.

Therefore, there is a contrast in the use of the 3pl prefix for non-humans according to whether they are definite (and usually they must be explicitly quantified), or indefinite. The indefinite use of 3pl is possible only for human referents (see §2.2.5 for the use of 3pl prefixes in other indefinite contexts).

Inanimate arguments are frequently the subjects of presentational predicates. (17) is an example.

(17) *gohje-bugi* *gu-ngolonggoh* *gu-ngolkgo*
there-just NEUT-*E. camaldulensis* NEUT-big

***gu**-janga+n*
NP-NEUT-*gu*-stand+PR
'A big river redgum stands there.' (KD: 169)

In this case the appearance of an overt NC argument prefix is typical, indicating that the subject prefix slot can also index presentational/indefinite subjects.

2.2.3 NC prefixes indexing arguments of intensional verbs

It is well known that 'intensional' verbs such as 'look for', 'want', 'ask for' tend to take indefinite objects, which are ambiguous between non-specific and specific readings (see e.g. HEIM AND KRATZER 1998:301ff; LYONS 1999:167-8). All the examples I have of these constructions in Ngalakgan take an indefinite and non-specific object. In the examples below, the verb *leh-* 'search for' is used with reference to wild honey (VEG class) in (18) and firewood (NEUT class) in (19). In both cases, this verb shows an overt NC prefix for the indefinite object:

(18) *mu-darda* *ngu-**mu**-leh+ya*
VEG-wild.honey 1sS-VEG-seek+FUT
ai gana wokabat jugabeg
'I'm going to look for wild honey.' [13/3/95:2]

(19) *alakgo* *ngu-**gu**-ma+nga,* *ngu-**gu**-leh+ya*
later 1sS-NEUT-get+FUT 1sS-NEUT-seek+FUT
'Later I'll go get some, I'll look for some...[firewood]' [13/3/95:2]

Another verb *baya+* 'seek and (potentially) attain' has much the same interpretation as *leh-* when referring to inanimates. It shows the same characteristics; in (20) taking an argument prefix for the VEG class object *mu-rok* 'pandanus':

(20) *matjji jarndiyah ngurr-marninyh+ya+gan*
indeed mat (MASC) 12pS-make+FUT+PURP
'Because/if we want to make pandanus mats...'

*mu-rok ngurr-**mu**-baya+n, ngurr-**mu**-ma+nga,*
VEG-pandanus 1sS-VEG-seek.attain+PR, 12pS-VEG-get-FUT
laik gejim pendemis bla meigim jet mets
'[then] we have to go and look for pandanus, we have to get some...'

In (21), an intensional context is set up using the modal future inflection of the transitive verb *na+* 'to see/find'.[28] The object is *manapburn* 'echidna'. Even though echidnas, like macropods, are usually individuated and can take a plural argument prefix, in this case the reading is indefinite and non-specific; there is zero agreement for this object.

(21) *ngurr-rorrongh+nga mirh-wi, gurnmarnh*
12pS-peep.at+FUT cave-LAT, maybe

*ngurr-**Ø**-na+na manapburn, ngurr-gatjja-ma+nga-gan,*
12pS-see+FUT echidna 12pS-nothing-get+FUT-REL

gurnmarnh marrawul ngurru-yongo+na
maybe hungry 12pS-sleep+FUT
if nathing wi nomo faindim enijing biif, wal wi silip anggri
'We'll look around in the caves, maybe we'll find *Tachyglossus aculeatus* [echidna]. If we find nothing then we might have to sleep hungry.'
[2/9/97:1B]

Hence, we find a distinction between NC argument prefixes and the 3pl prefix when coreferential to an animate but non-human argument under the scope of some indefiniteness-introducing operator or predicate. NC prefixes are permitted under such conditions, but the 3pl prefix, as in the presentational examples in §2.2.2, is disfavoured (and I have no examples where a 3pl prefix is coreferential to an indefinite and non-specific nonhuman argument).

SAPs *can* be the objects of such verbs, but they can never have an indefinite interpretation in this case. SAP affixes can only have definite interpretations. This restriction is a consequence of the meaning of SAP morphemes: SAPs are inherently anaphoric.

---

28 The default interpretation of the 'future' inflection of verbs in Ngalakgan is modal: 'want to, might, have to, should' are equivalents of the translations given by speakers (in Kriol: *garra, labda, maiti, gana,* and *andi*). In (21), the modal interpretation is reinforced by the uncertainty operator *gurnmarnh* 'maybe', as well as the speaker's translation in Kriol using 'if'.

2.2.4 *Argument prefixes under the scope of interrogative, negative, and indefinite operators*

NC argument prefixes are commonly co-referential to nominals under the scope of a question word (22).[29] In (22), the NC argument prefix *mu-* for VEG class is co-referential to a WH-word *mu-yanah* 'what' which is also marked for VEG class.

(22) a. *mu-yanah*$_i$ *ju-**mu**$_i$-jong+iny*
VEG-what 2sS-VEG-chop+PC
'What kind of [VEG class object, sc. 'tree'] did you cut down?' [2/9/97:2A]

b. *mu-yanah*$_i$ *ngurr-**mu**$_i$-nguna*
VEG-what 12pS-VEG-eat+FUT
'What are we going to eat?' [11/9/97]

I take it that the questioned referent in such examples (from the speaker's point of view at least) can only be characterised as indefinite and non-specific. That being the case, NC prefixes can therefore be co-referential to such entities.

In English, the determiners 'any' and 'no' also derive indefinite, non-specific readings of noun phrases (LYONS 1999:37). Ngalakgan has equivalent constructions, which include a negative suffix on the verb, and the use of the 'definite' prefix form on the quantified expression, as in (23).[30] In this example, there is a NEUT argument prefix co-referential with the expression under the scope of the 'any' construction (here: *yang* 'language'),

(23) *ju-**gu**-goh+nani+kgorro,* *gun-gu-yang*
2sS-NEUT-have+[see+]PRNEG NEUT-DEF-language
yu nomo gadim eni langgus
'You don't have any [traditional] language.' [2/9/97:1A]

I have no examples of NC prefixes co-referential to a nominal under the scope of a polar interrogative operator. All examples are like those in (24), where the verb shows zero agreement for the questioned object.

(24) a. *mu-may* *ju-goh+na+n?*
VEG-veg.food 2sS-have+[see+]PR
'Have you got any vegetable food (VEG)?'

---

29 Unfortunately, I have no examples of this construction with a 3pl argument.

30 The reason for the definite prefix in this construction is discussed elsewhere (B. BAKER 2000). I argue there that in these 'any/no' constructions, as well as when modified by the indefinite/interrogative pronouns, nouns are interpreted as partitive 'kind' references, i.e. *gun-gu-yang* is interpreted something like '[any] *kind of* language' in this context.

b. *gu-birn* *ju-goh+na+n?*
NEUT-money 2sS-have+[see+]PR
'Have you got any money (NEUT)?'

The 3pl prefix can be used in the same way to refer to indefinitely quantified human referents.

(25) a. *bigurr-tjji* ***burru****-bolh-yongoni+kgorro*
person-PRIV 3pS-track-lie+PRNEG
'No one has been here.' (Lit. '[it] is without people, their tracks are not lying [here].') [2/9/97:1A]

b. *rerre-kgah=rerre-kgah* *warlamanh* ***burru****-nang+an*
DIST=camp-LOC everybody 3pS-sit+PR
'Everybody sits in their own camp.' [3/9/97]

c. *gorreh-gorreh* ***burru****-man-nang+an*
DIST-alone 3pS-group-sit+PR
'Every group keeps to itself.' [3/9/97]

d. *nu-gu-malaborno* ***burru****-mah,* *nu-gu-malaborno*
MASC-DEF-some 3pl-good MASC-DEF-some

burru*-borde-bordewk*
3pl-INTENS-bad
'Some [children] are good, some are very bad.' [4/7/96]

The examples in section §2.2 have shown that both NC prefixes, and the 3pl prefix when in reference to human arguments, can index arguments which are indefinite, nonspecific, or nonreferential in various contexts.

2.2.5 Discussion of indefiniteness

In English, it is possible for pronouns to be co-referential to kind references whether singular, as in (26), or plural, if the kind reference acts as an antecedent to the pronoun. It is also possible for SAP pronouns in English to refer to kinds, as in (26c), and the translation of (15), above.

(26) a. *The echidna* lives in a hole, *it* has quills and (*it*) lays eggs.
b. *Acacia* is native to Gondwanaland. In Australia, *it* is commonly called 'wattle'.
c. When *we* see an echidna, *we* always track *it*.

Since anaphoric pronouns in English can refer to 'kinds', the fact that argument prefixes in Ngalakgan can also be referential to kinds does not mean that argument prefixes in Ngalakgan are not anaphors. In order to find contexts in which Ngalakgan and English contrast in their use of pronominal elements we must look to non-referential contexts, as in §2.2.

The use of both the NC prefixes and the 3pl prefix in indefinitely quantified conditions contrasts with English, as Evans (this volume) points out. In English, the equivalents of the examples in (25), using anaphoric pronouns as in (27), sound ungrammatical or marked.[31] Hence Evans' argument against the interpretation of argument prefixes in Bininj Gun-wok as always being pronominal anaphors (cf. RIZZI 1986).[32]

(27) a. * 'Everybody, they sit by themselves.'
b. ? 'Some children, they are good.'
c. * 'You don't have it, any traditional language.'

Evans argues that argument prefixes are more like agreement markers than pronominal arguments when co-referent to an external nominal expression, and that in that case, it is nominal expressions, together with information from the argument prefixes, which instantiate the arguments of the clause.[33] I have already noted the restriction against external nominal expressions of SAP arguments in Ngalakgan. In the following section, I show that NC prefixes need not agree with their co-referent arguments, and that, indeed, in some cases they are required *not* to agree.

---

31 However, it is not clear that the reason for the strangeness of the examples in (27) is because of an incompatibility between anaphoric pronouns and indefinite quantification. The strangeness could instead be because of an incompatibility between, on the one hand, indefinite pronouns (e.g. 'everybody') and expressions, and, on the other, NPs in a dislocated, hence topical function. The function of such dislocations in English is to reiterate topics which are, by definition, referential. Contrast the strangeness of the examples in (27) with the seemingly natural translations of the examples in (25), which also contain co-referential indefinite pronouns and anaphors such as 'their' and 'itself'.

32 RIZZI (1986:395) shows that indefinite pronouns are incompatible with a resumptive (anaphoric) pronoun, as shown in the contrast between the following examples (= 8 and 9 in RIZZI 1986):

a. Non conosco nessuno in questa citta'
'I don't know anybody in this city'

b. ***Nessuno, lo** conosco in questa citta'
'Nobody, I know him in this city'

c. NESSUNO, conosco in questa citta'
'Nobody, I know in this city' (with contrastive emphasis on *nessuno*)

Rizzi accounts for this contrast with the following principle:
'A pronoun cannot be locally bound by a (non lexically-restricted) quantifier.'

33 EVANS (this volume) does not rule out the possibility that argument prefixes in Bininj Gun-wok function as arguments at least some of the time (e.g. in the absence of co-referential nouns).

2.3 Disagreement

In (28a), speaker A asks speaker B for an identification of a specific spear from a known set. This example contains two instances of noun class *disagreement*. In (28a), the questioned generic noun *nu-gu-barrakgarlh* '(any kind of) spear or species used for spear shafts' is marked as MASC class, but the argument prefix co-referential with this argument is marked as NEUT class. The specific noun *mu-jukgul* '*Acacia holosericea*' is assigned to VEG class, but in (28c), the verb takes a NEUT class prefix co-referential with this argument:

(28) a. *Yanah-ba* ***nu-gu-****barrakgarlh* *ju-**gu**-mehme?*
what-INT MASC-DEF-spear 2sS-NEUT-get.PP
A: 'Which spear did you get?'

b. *jukgul,* *mu-jukgul*
A. holosericea VEG-*A. holosericea*
B: 'Acacia, an acacia one.'

c. ***mu-****jukgul* *ju-**gu**-mehme?*
¡VEG-*A. holosericea* 2sS-NEUT-get.PP
A: '[Oh, it was] an acacia one that [spear] you got?' [11/9/97]

Note that the verb retains the NEUT marker in both cases while the class of its coreferential noun is MASC in one case and VEG in the other.[34]

One's first reaction to such examples is that they must be speaker errors. The following example shows that this is not the case. In (29), speakers were asked 'What do emus eat?' Speaker B (the junior speaker) has a plant species in mind but can't remember the name. Speaker A suggests the name to her (at line 29a) with its lexically-specified NC prefix (NEUT). This is the term that Speaker B was looking for, and she concurs (in 29b), using NEUT class agreement on the verb 'eat'. Speaker A corrects her (at line 29c), repeating the verb but replacing her NEUT class verb prefix with VEG class.[35] The way speaker B seems to interpret this is that she has placed the *noun* into the wrong class, and repeats the noun with a VEG class prefix (reflecting speaker A's prefix on the verb). Speaker A then repeats both noun (NEUT class) and verb (VEG class) together in line (e), making it quite clear to speaker B (and to us), what the correct form of the construction is in this case:

---

34 The disagreement phenomenon described here has been observed only between a verb and its argument(s). It does not apply between a noun and its modifiers (see §3 for a superficially similar process applying to nominal-modifier agreement in Ngalakgan).

35 Speaker A is senior to speaker B, and a first-language speaker of Ngalakgan. Speaker B is highly fluent in Ngalakgan but through years of under-use she has lost some proficiency. The intonation of this stretch of text indicated that Speaker A was actively correcting Speaker B, not merely interacting.

(29) a. ***gu-wularra?***
NEUT-?*Cucumis sp.*
A: '*Wularra*?'

b. *yo,* *gunhbirri-gunh* *gu-**gu**-ngu+n*
yes, that-SPEC NP-NEUT-eat+PR
B: 'Yes! it eats THOSE ONES [fruits].'

c. *gu-**mu**-ngu+n*
NP-VEG-eat+PR
A: 'It eats it/them.'

d. ***mu**-wularra*
VEG-?*Cucumis sp.*
B: 'Melons'

e. ***gu**-wularra,* *gu-**mu**-ngu+n*
NEUT-?*Cucumis sp.* NP-VEG-eat+PR
A: 'It eats them, melons.' [10/9/97:1A]

This example demonstrates that instances of disagreement in Ngalakgan are not considered to be errors here. Indeed, speaker A's grammaticality judgement indicates that, in some circumstances, verbs *must* disagree with a coreferent argument. At this stage, it is unclear exactly what those circumstances are.

Example (30) presents another instance of the same kind. Again, NEUT class marked nouns *gu-baramulk* and *gu-wularra* are co-referential with verbs marked for VEG class objects *gu-mu-ngun* 'it eats them (VEG class)'.

(30) a. *yanipbi* ***gu**-baramulk,* *gu-**mu**-ngu+n,*
whatsit NEUT-*Cucumis melo,* NP-VEG-eat+PR
'Whatsit? wild cucumber it eats...,'

b. *bo-wi* *nu-gohje* *gu-mu-jung-hgun*
river-LAT MASC-that NP-**VEG**-shade-REL
'around the river where that one is in the shade,'

c. ***gu**-baramulk,* ***gu**-wularra* *gu-**mu**-ngu+n*
NEUT-*Cucumis melo* NEUT-*Cucumis sp.* NP-VEG-eat+PR
'wild cucumber, [and] wild melon it eats.'

In §3, I consider and reject an analysis of Ngalakgan disagreement under a proposal by CORBETT (1983), whereby constituents can agree 'semantically' as well as [morpho-] 'syntactically'. In §3.1, I find that more convincing parallels to Ngalakgan can be found in other Australian languages.

3. Why Ngalakgan disagreement is not 'semantic' agreement

The noun class disagreement behaviour shown by Ngalakgan in examples such as (28) is distinct from other noun class variation or disagreement patterns reported in the literature on Australian languages and languages elsewhere. Typically in Gunwinyguan languages, contrast between classes is neutralised in modifiers and verbs, so that these constituents display a smaller class system than that of nouns. The contrast between two inanimate classes may be neutralised in favour of one of them (in e.g. Gun-djeihmi), or all non-feminine classes may be neutralised in favour of MASC agreement (in e.g. Jawoyn) (EVANS 1997b, HARVEY 1997, who call this kind of pattern 'superclassing'). Certain constructions in Ngalakgan do show this kind of neutralisation: demonstratives optionally neutralise to MASC agreement for all non-feminine head nouns (example 30b is an instance).

The superclassing phenomenon is rather different to the Ngalakgan disagreement phenomenon. Superclassing is an optional or obligatory process which reduces the total number of agreement classes. Disagreement does not reduce the number of agreeing elements: both VEG and NEUT argument prefixes can disagree with their co-referential nominal. On the contrary, disagreement increases the number of potential agreement classes to which a head class can be co-referential.

What is at issue in the disagreement examples in §2.3 is the lack of fit between the noun class of the noun, and the *function* or *use* to which it is being put in a particular circumstance. Generic terms such as 'leaf', 'wood', 'spear', 'medicine' and so on are very rarely used in Ngalakgan to directly modify species terms. In Ngalakgan, as in Australian languages in general, nouns such as *mu-jukgul* have variable construals: *mu-jukgul* can refer to the tree species *Acacia holosericea*, either as an individual or as an ecozone, or *mu-jukgul* can refer to any part of the tree with no further modification of the noun, including items manufactured from its wood such as spear shafts. Generic terms such as *barrakgarlh* 'spear (shaft)' occasionally occur in *apposition* to species terms, but are not required to appear in order for these terms to have such differing construals, and are far less common than the equivalent construction in English. Constructions like 'Acacia leaf', 'Acacia tree', 'Acacia grove', with rare exceptions (see §4), simply do not occur in Ngalakgan.

Rather than using generic nouns to indicate the construal of a specific term (as in English), in Ngalakgan the form of the NC prefix on the verb can be used to indicate the construal of *mu-jukgul* in the clause: whether, for instance, it is a '(type of) spear shaft' or a '(type of) tree' that is being spoken about. In a sense then, the NC prefix to the verb 'agrees with' the *construal* of the noun, rather than its grammatical form. In (30), *gu-baramulk* refers to *Cucumis melo* qua 'plant with edible fruit' and hence the verb takes the argument prefix which is appropriate to the domain of edible plants: VEG class *mu-* (similarly with example 29). In (28), by contrast, *mu-jukgul* in this instance refers to *Acacia holosericea* in its capacity as a plant which is a source of heavy spear shafts. The NEUT class is most clo-

sely associated with inedible plant species and (especially) the implements derived from plant species, and that is the argument prefix used in this case. *Acacia holosericea* is also used for other purposes, and we expect that different NC argument prefixes might be used in these cases.

Can this use of the NC prefixes be said to be 'agreement', in the usual linguistic sense of the term?[36] It is appropriate at this point to discuss why the Ngalakgan pattern is different from other cases of morphosyntactic feature clash discussed in the literature. CORBETT (1983) for instance, discusses a number of Slavic languages in this regard. In Czech (31), the noun *děvče* 'girl' inflects like a NEUT noun, and modifiers and predicates must be in NEUT form (e.g. *přišlo* 'came (NEUT)'). However, an anaphoric pronoun which is co-referent with this noun may be either NEUT, or FEM as in (31) (from CORBETT 1983:9).

(31) a.	*To*	*děvče*	*přišlo*	*včera*	*ale*	*já*
	that	girl(NEUT)	came(NEUT)	yesterday	but	I
	jsem	*ji*	*nenajmula*[37]			
	did	her(FEM)	not.hire			

Corbett refers to agreement of the kind we see between the noun *děvče* and the pronoun *ji* in (31) as 'semantic agreement': the pronoun is FEM because the noun *děvče* refers to a female, even though it is formally NEUT. Semantic agreement of this kind is opposed to 'syntactic agreement', requiring identity of the morpho-syntactic features relevant to agreement (as in the modifiers and predicates in 31). The reason for the disagreement in this case is fairly clear. In Czech, the noun *děvče* 'girl' always refers to a female, even though the form of the noun is NEUT.

But there is a difference between the Slavic examples discussed in CORBETT (1983) and the Ngalakgan disagreement phenomenon. In Russian, nouns such as *vrač* 'doctor', can have variable construals (and hence agreement) depending on the sex of the referent (CORBETT 1983:10). But in every case in the Slavic examples, there is only one construal of a noun which is relevant to the agreement behaviour of modifiers, pronouns and predicates. That is because the only semantically-relevant construal is inherent in the referent itself (masculinity, femininity). In Ngalakgan by contrast, the construal of a noun which is relevant to agreement is contingent on function or use (see below). Unlike the case in Slavic languages, semantic agreement in Ngalakgan is a context-dependent phenomenon, which is contingent on the particular construal of a term in discourse. This kind of behaviour has been reported for 'classifier' languages, but not for 'gender' languages.

---

36 As CORBETT (1983:1) and ANDERSON (1992:103) have noted, it is by no means straightforward to define 'agreement'.

37 CORBETT (1983) supplies no translation for this example.

The discourse-dependent nature of disagreement in Ngalakgan can lead to the situation that there are usually at least two, and often many more, construals which are relevant to agreement. The differences between Czech and Ngalakgan are presented in Table (3). At the bottom of the table we see the terms in question and the gender/noun class which is assigned to them in the lexicon: for *děvče* it is NEUT, for *jukgul* it is VEG. While *děvče* can only have the construal 'female', and hence the only possible 'semantic' agreement class for this noun is FEM, the term *jukgul* can have many possible construals, depending on its function or use in a particular discourse context, and each of these has characteristic noun class assignments.[38]

Table (3): Semantic agreement vs. disagreement

	Czech	**Ngalakgan**					
Construal	'young female'	'soap'	'medicine'	'tree'	'fish poison'	'spear'	'firewood'
Semantic Agree't	FEM	VEG	NEUT	NEUT	VEG	NEUT	NEUT
Assign't	NEUT	VEG					
Term	*děvče* 'girl'	*jukgul 'Acacia holosericea'*					

Therefore, I submit that the kind of 'semantic' agreement we see in Ngalakgan is of a different order to what we see in Czech and other Slavic languages. In the next section, I consider some more fruitful analogies to the Ngalakgan phenomenon, before proposing an analysis of the function of NC prefixes in §4.

3.1 Australian analogies: classifier systems and classifier constructions

Ngalakgan disagreement finds a more relevant analogy in other Australian languages, than the Slavic examples Corbett discusses. I briefly discuss two here: Mparntwe Arrernte

38 Some of these characteristic noun class assignments can be postulated on the basis of disagreement examples, others are extrapolated from gender assignment to nouns. For instance, the major ichthycide and soap-producing trees, as well as the terms for 'poison' (*mu-marawirrina*) are all VEG class. All terms for fire and firewood (e.g. *gu-munhmunh* 'tinder', *gu-ngurnih* 'firestick'), as well as the generic terms for tree/stick (*gu-darh, gu-jandah*), are NEUT, and loanwords referring to this domain are classed as NEUT also: *gu-jambakgu* 'tobacco', *gu-metjjitj* 'matches', *gu-beypba* 'rolling papers' (contrast *gu-got* 'writing paper, book'; lit. 'paperbark tree'). I have come across no indigenous term meaning 'medicine' so far, but the loanword is also NEUT class: *gu-merritjjin*. At least one term for an important medicine plant *gulukgulun* (poss. *Melaleuca acacioides*) has been observed to change its noun class *assignment* (as well as its agreement) from VEG to NEUT when it was discussed in a processual text about making medicine from the plant.

(Central Australia, NT: WILKINS 2000), and Ngan'gityemerri (Daly region, NT: REID 1997). I have already noted that it is characteristic of Australian languages for any noun to have a wide range of construals. While Ngalakgan makes some distinctions in the construal of a noun by using oppositions in the NC argument prefix paradigm, many Australian languages make similar distinctions by using generic nouns to specify the function or use to which the referent of a species term is being put.

WILKINS (2000) discusses 'classifier' constructions in Mparntwe Arrernte: phrases consisting of a specific and generic noun in co-reference to an entity. In (32) (WILKINS 2000:151), 'kangaroo' is referred to once as *kere aherre*, literally 'game animal kangaroo' and once simply as *kere* 'game animal'. By contrast, in example (33) (WILKINS 2000:173), the reference to kangaroo is realised as *aherre* 'kangaroo' alone.

(32) a. *Ikwere-nge* *re-therre* *perte-ke* *anteme,*
3sgDAT-ABL 3dlNOM creep.up-PC now

b. ***kere*** ***aherre*** *ikwere.*
game/meat kangaroo 3sgDAT

c. *Kele* *itwe-k-irre-mele* *arratye*
OK near-DAT-INCH-SS true

d. *atanthe-ke* *irrtyarte-le* ***kere*** *re-nhe.*
spear-PC spear-INST game/meat 3sg-ACC
'After that, the two of them now crept up on the kangaroo. So, when (they) got closer (they) speared it truly with a spear.'

(33) a. *Inarlenge* *kenhe* *kwele* *ane-me* *kwerrke* *kwele*
porcupine BUT QUOT be-NPP young.one QUOT

b. *arrate-me* *apethe* *ikwerenhe-nge,* ***aherre-arteke***
appear-NPP pouch 3sgPOSS-ABL kangaroo-SEMBL
'Now echidnas are supposed to be (like this); baby echidnas, so they say, come out of their pouches just like kangaroos do.'

Wilkins shows that the choice - whether an Arrernte speaker uses the noun alone, as in (33) or the generic+specific construction, as in (32) - depends on the nature of the activity in which the referent is embedded. If the reference is to kangaroo *qua* potential source of meat, then the generic noun *kere* 'game/meat' precedes or is used in place of *aherre*. If the reference is to the natural habits and habitat of kangaroos, or the kangaroo Dreamtime ancestor, then *aherre* is used alone. Furthermore, many other specific nouns, such as *arlkerrke* 'meat ant', can appear in construction with various of the set of generics such as *yerre* 'ant' or

awelye 'medicine', depending on their inherent nature, use, or function in any particular instance. That is, generics in Arrernte perform something like the function that NC prefixes can perform in Ngalakgan: they delimit the construal of a specific noun. Wilkins uses the term 'discourse-relevant' to describe such contextually-based construals, and I follow his usage here.

Ngan'gityemerri is a language with a system of nominal classifiers: a closed set of morphemes which occur in construction with another noun and also appear in 'agreement' with modifiers (demonstratives, adjectives, numerals).[39] There are fifteen nominal classes in Ngan'gityemerri, nine of which are realised as agreement affixes on nominal modifiers. In examples (34a) and (35a), we see the standard or 'syntactic' agreement pattern. In (34a) the species term *menem* 'billygoat plum' (*Terminalia ferdinandiana*), which is assigned to VEG class, being a food source, takes a VEG class agreement proclitic on a co-occurring modifier *biny* 'ripe'. In (35a), the possessive modifier of *tyulut* 'hookspear' takes an agreement classifier of the 'bamboo spear' (B'BOO) class, to which *tyulut* itself belongs.[40]

Apart from the regular agreement, Reid also finds 'semantic' agreement like the (b) examples in (34) and (35). In (34b), the discourse-relevant construal of 'billygoat plum' is as a tree that needs to be cut down, not as a source of food. The modifier takes a proclitic marking the 'tree' class. And in (35b), the possessive modifier of *tyulut* 'spear' also takes an agreement classifier of the 'tree' class, to which most implements belong. Reid suggests that the speaker viewed the spear in this instance less as an offensive weapon than as an implement needing repair (REID 1997:202).

(34) a. ***mi-menem*** ***mi=biny*** *werrmim-ba-ket*
VEG-billygoat.plum VEG=ripe 3pS:AUX-arm-cut
'They are picking ripe Billygoat Plums [*Terminalia ferdinandiana*].'

b. ***mi-menem*** ***yerr=kinyi*** ***yerr=syari*** *yubu-ket-Ø*
VEG-billygoat.plum TREE=this TREE=dry 2sS:AUX-cut-IMP
'Chop down this withered Billygoat Plum tree!'

(35) a. *tyulut* ***yeli=nem*** *ngarim-pawal*
hookspear B'BOO=his 1sS:AUX-throw
'I threw his hook spear ('bamboo spear' class).'

---

39 The morphology of classifiers in Ngan'gityemerri is rather complex, I refer the reader to REID (1997). Noun class can be marked on both heads and modifiers in Ngan'gityemerri, but in various ways: 'Of the fifteen noun classes, six are denoted by only freeform generics, five by only bound generics, and four employ both bound and freeform generics. Furthermore, among the noun classes denoted by bound generics, there is diversity regarding the phonological status of those generics, with some being prefixes, and others proclitics' (REID 1997:172-3).

40 Not all nouns have their class marked endocentrically. As Reid observes, the class of many nouns is only revealed through [syntactic] agreement patterns.

b. *tyulut* ***yerr***=*nem* *ngerim-wurity-ngirim*
hookspear TREE=his 1sS:AUX-fix-1sS:sit
'I'm fixing his hook spear ('bamboo spear' class).'

REID (following CORBETT 1991) regards the pattern in (34b) as resulting from a covert re-assignment of *menem* 'Billygoat Plum' to the 'tree' class. This assignment is covert, since the noun still takes a VEG class prefix, and is only realised in the class taken by co-referring modifiers. While this is not the place for a critique of Corbett's theory, it is clear that the circumstances under which disagreement is favoured in Ngan'gityemerri are similar if not identical to those in which disagreement is observed in Ngalakgan.

I propose that the NC argument prefixes of Ngalakgan are performing much the same function for the construal of nouns in Ngalakgan as generics do in Arrernte, and classifier agreement prefixes do in Ngan'gityemerri. That is, given the fact that in Ngalakgan, as in the other languages mentioned, nouns can have a range of construals, the NC prefixes serve to 'fix' or 'delimit' the construal of a noun to one or a few possibilities, which are dependent on context (including the choice of predicate). In the following section I examine the implications of this analysis for the morphology of verbs.

4. A morpho-semantic analysis of NC prefixes

Ngalakgan, like Arrernte and Ngan'gityemerri, also has a range of generic nouns such as *gurndu* 'country', *darh* 'tree', *barrakgarlh* 'bamboo spear (shaft)', *darren* 'canegrass spear (shaft)', *may* 'vegetable food', *ray* 'small game, meat', *janggu* 'meat', *weh* 'water, watery liquid', and others. As I show below, these generics can be co-referential to a specific noun, either as verb-incorporated stems, or in complex nominal constructions. In that case, one may ask whether generic nouns in Ngalakgan play the same role that they do in Arrernte and Ngan'gityemerri, and if so, what then is the function of NC prefixes, which I have suggested perform the function that generic nouns or classifiers perform in these other languages.

In Ngalakgan, generics occasionally occur in combination with a specific, as in (36), in a 'classifier' construction similar to that found in Arrernte (WILKINS 2000). In (36), note that the ergative case is marked only on the generic *gu-darh* 'tree' - the specific term *mu-malbah* 'ironwood' takes no case marking. This indicates that the two words constitute a complex nominal expression. In this kind of construction, generics perform a similar function to the generics of Arrernte or the classifiers in Ngan'gityemerri: they delimit the use, function or inherent state of the specific term. In (36) for instance, the choice of *darh* 'tree' indicates that it was the whole tree which fell down, rather than just a branch (for which the speaker would perhaps use *jandah* 'stick, tree', *wanjat* 'arm' or just *malbah* alone). (Note that here we have an instance where we would expect disagreement - trees are prototypically NEUT

class - nevertheless, the verb takes VEG class agreement, perhaps because there is no ambiguity in the interpretation of *malbah* given the presence of the generic in this case.)

(36) *ngun-mu-bohbo,* *gu-**darh**-yih* *mu-**malbah***
1sO-VEG-hit.PP NEUT-tree-ERG VEG-ironwood

mu-ngorh-miny en *ngun-mu-bohbo*
VEG-fall-AUX.PP and 1sO-VEG-hit.PP
jet wadi bin foldan la mi en kilim mi
'An ironwood tree [lit. 'tree ironwood'] fell down and hit me.' [2/9/97:1B]

A subset of lexical generics can also be incorporated into verbs, as in (37), where they perform much the same function as external generic+specific phrases (the difference between the two relates to information structure). In (37), the external noun *langgah* 'billabong' [a seasonal or permanent pool of water in sand, clay, or loam] is a type of *gurndu* 'country', which occurs here as an incorporated co-referential noun.[41] In (38), the incorporated stem *mili-* 'water' is co-referential to the external (borrowed) noun *di* 'tea'.

(37) *gun-gohje-bugi* *gu-**langgah*** *ngu-**gurndu**-ngey+bu+n* *'Yalwarra'*
NEUT-there-just NEUT-billabong 1sS-country-name+[hit+]PR[toponym]
'I call that lagoon "Yalwarra".' (Lit. 'I country-name that lagoon "Yalwarra".') [KD:5]

(38) *ju-na+n* *ju-gu-**mili**-boylimhma+gan* *gu-**di***
2sS-see+PR 2sS-NEUT-water-boil+PURP NEUT-tea
'You look, if you want to boil tea...' [31/5/96:1B]

The incorporated nouns in examples (37-8) here perform a 'classificatory' function, according to MITHUN's (1984:863) criteria: they narrow the scope of the verb, while allowing a more specific noun to function as object.

---

41 Naming country is a particular kind of activity undertaken by Dreaming beings and also by humans as an invocation of the Dreaming, hence, this compound is 'name-worthy' in MITHUN's sense (1984:856). But it is not a lexicalised compound in the same way that *weh-nyarh* (lit. water-die) 'dying of thirst' is. Unlike the latter, *gurndu-ngey-bu+* can also be expressed phrasally: *gun-gu-gurndu ngu-gu-ngey-bun* 'I call that country...' In addition, *gurndu* can be omitted from example (37) without changing its (logical) interpretation. That is, as with true incorporation in general, *gurndu* is optional here. See EVANS (1995, 1997a for similar arguments with respect to noun incorporation in Mayali (Bininj Gun-wok).

Generics can also occur alone, as in (39) and (40). In this case they do not refer to specific referents, but rather to some unknown referent, or to a 'kind' referent of a higher taxonomic level (a 'lifeform').

(39) *gu-wangginy* *gun-gu-**darh*** *gu-gu-yerrert*
NEUT-one NEUT-DEF-tree NP-NEUT-grow
'There's one tree growing there (I don't know what kind).' [3/9/97]

(40) ***jandah**-gah,* *gu-yongo+n,* *nu-gu-dugurlah*
stick-LOC NP-lie+PR MASC-DEF-ringtail.possum
dugurlah, thet posim, lib la tri, dugurlah
'The ringtail possum lives in the branches [of trees] (any kind/different kinds of tree).' [2/9/97:1B]

Such generic+specific constructions, whether the generic is external or incorporated, are not common in Ngalakgan. The examples presented here (from a corpus of around 4000 lines of elicitation and text material) represent the most unambiguous instances where generic and specific nouns are co-referential.

The situation in Ngalakgan contrasts strongly with the status of generics in Eastern Arrernte as reported in WILKINS (2000). Wilkins examined a database of 26 texts, 15 of which contained NP references to 'kangaroo' headed by a noun: generic, specific or generic+specific. There were a total of 50 occurrences of NPs of this kind, of which 17 (34%) were generic+specific constructions. Therefore, I conclude that Ngalakgan speakers make very little use of generic nouns in reference to specific entities, whether external or incorporated, as compared to other Australian languages. If Ngalakgan speakers are delimiting the reference of nouns, as I have claimed they are, then that function is performed almost exclusively by the NC argument prefixes.[42]

One might suppose that NC argument prefixes are, in fact, merely reduced or alternate forms of the incorporated generic nouns. There are several characteristics distinguishing incorporated nouns from NC argument prefixes which militate against this proposal. Firstly, NC prefixes can be distinguished on morpho-phonological grounds from incorporated generics. On several morpho-phonological criteria, NC argument morphemes are clearly *prefi-*

42 Ngalakgan therefore lends support to the claim by FOLEY (1997) that elaborated and productive noun class systems are in complementary distribution to productive noun incorporation in polysynthetic languages. It makes sense of my informal impression that the degree of incorporation in Ngalakgan is much less than that in closely related Gunwinyguan languages like Rembarrnga and Bininj Gun-wok. In these languages, there is no NC agreement in verbs (indeed, verbal agreement for noun class is rare in GN languages, with only Ngalakgan, Ngandi (HEATH 1978) and Nunggubuyu (HEATH 1984) having this feature). It appears that the capacity for verbal agreement, and especially *dis*agreement, is the crucial feature separating Ngalakgan from languages like Bininj Gun-wok and Rembarrnga.

xes, not stems. Like other prefixes, they consistently lack stress (41), and hence a following monosyllabic open verb root must have a long vowel at the surface (B. BAKER 1999). This is because final stress in Ngalakgan is only licensed when the syllable on which it falls is both a monosyllabic root which is a possible word, and is prosodically heavy: i.e. is either closed, or has a long vowel at the surface.

(41)	a.	*yirr-**mi**-ma*	[jɪrmimá:]	
		1pS-VEG-get.PR	'we get it (VEG)'	[DP]
	b.	*ngurrurn-**mu**-ne*	[ŋʊrʊɳmuɳé:]	
		12pO-VE-burn.PR	'It [sc. 'sun' (VEG)] burns us.'	[13/3/95:3A]

Prefixes contrast in this respect with compounded/incorporated stems, which *must* be stressed regardless of size (42) (B. BAKER 1999).

(42)	a.	*gu-**weh**-mah*	[guwèɛ̱máa̱ʔ]
		NP-water-good	'it is freshwater'
	b.	*yi-gi-**bo**-wa+na*	[jigibò:wána]
		12S-NEUT-river-follow+FUT	'we're following the river'

NC prefixes are therefore distinct morphologically and phonologically from incorporated generic nouns such as *may* 'food', *mili-*, *binyi-*, and *weh*, all of which mean 'water' or 'liquid', and *gurndu* 'topographic', since the latter are treated like stems.

Secondly, NC prefixes, unlike incorporated generics, cannot co-occur with two filled prefixal argument positions, and in this respect are like other argument prefixes. Like many multi-argument indexing languages, Ngalakgan permits a maximum of just two argument prefixes to a verb. In the case of trivalent verbs such as *wu-* 'give', *mutjji-* 'show, teach', and *bak-bawunh-mi-* 'leave for', which have three syntactic-semantic roles - agent (e.g. 'giver'), theme (e.g. 'gift'), and goal (e.g. 'recipient') - the two roles expressed by verb prefixes are the agent and the goal (43a). When the theme is an inanimate entity, as it typically is, there can be no argument prefix indexing this entity (43b). But it may be expressed by an incorporated (or external) noun, as in (43c, cf. also 44).

(43)	a.	*ngu-bu-wohwo*
		1sS-3p-give.PP
		'I gave them [something].' (Not: 'I gave them to him/her.')
	b.	**ngu-bu-**mu**-wohwo*
		1sS-3p-VEG-give.PP

c. *ngu-bu-**may**-wohwo*
1sS-3p-food-give.PP
'I gave them vegetable food (*may* VEG class).'

d. **ngu-bu-**mu-may**-wohwo*
1sS-3p-VEG-food-give.PP

(44) a. *jun-**gurndu**-mutjji+nga*
2s/1s-country-show+FUT
'You have to show me your country.' [22/5/98]

b. **jun-**gu**-mutjji+nga*
2s/1s-NEUT-show+FUT
'You have to show me it [sc. 'country'].' [22/5/98]

That this restriction is morpho-syntactic, not morpho-phonological, in nature is shown by examples such as (45), where a co-referent NC prefix and incorporated noun co-occur.

(45) *ngurr-**mu-birtdin**-marninyh-nya*
12pS-VEG-bread-make+FUT
'We'll make some lilyseed bread.'

Apart from the phonological and morphological criteria, NC prefixes are much more frequent than generic nouns. These patterns indicate that NC argument prefixes are not some kind of reduced incorporated generic noun. Clearly, they are argument prefixes.

However, it has been my contention throughout that the argument prefixes of Ngalakgan do not constitute a homogenous class. Before proposing a classification of these prefixes, I outline in summary the differences between them, as shown in Table (4).

Table (4): Contrasting properties of argument prefixes

	SAP	3pl (+hum)	3pl (-hum)	inanimate NC
Obligatory?	yes	yes	no	no
Stress?	no	no	no	no
Generic refs?	yes	yes	no	yes
Indefinite (∃) refs?	no	yes	no	yes
Neg.indefinite (¬∃) refs?	no	yes	no	yes
Questioned refs?	no	yes	no data	yes
Can disagree?	no	no	no	yes

On many tests, the NC prefixes perform just like the 3pl prefix when co-referential to humans, but not animals. In both cases, and in contrast to SAP prefixes and to 3pl prefixes

referring to non-humans, both definite and indefinite references are possible. The use of the 3pl prefix in reference to animals is restricted to definite, referential cases only.[43] I have argued that the argument prefixes of Ngalakgan can be divided into at least three categories according to their morphological behaviour and their referential characteristics. In concluding I suggest how these differences can be characterised informally.

5. Conclusion

There have been two major currents to the debate over the status of multiple argument-marking affixes in polysynthetic languages. On the one hand, authors such as JELINEK (1984) and M. BAKER (1988, 1996) have proposed, following earlier authors such as BOAS (1911), that such affixes constitute the 'real' arguments of the clause, and that the optional nominals represent adjuncts linked by apposition. BRESNAN and MCHOMBO's (1987) term for such pronominal arguments is 'pronominal anaphor'.[44] On the other hand, SIMPSON (1991), AUSTIN and BRESNAN (1996), NORDLINGER (1998), EVANS (this volume), and others have shown that argument affixes cannot constitute pronominal anaphors in all cases.

I have shown here that we must distinguish more carefully between the categories of person, number, and gender since Ngalakgan evinces several such categories that have differing distributions according to their inherent person/number/noun class features or else the characteristics of their co-referential nominal expressions. There are three formal categories of argument prefixes according to the criteria examined in this paper: SAP, 3rd person plural, and inanimate.

These categories cannot be characterised uniformly neither as 'agreement' markers nor as pronominal anaphors. On the one hand, SAP argument prefixes do not countenance such external co-referentiality at all, and hence must be interpreted as the arguments of the predicate to which they are affixed. That is, SAP argument prefixes are bound pronominal anaphors.

On the other hand, the argument prefixes for the two inanimate noun classes do not necessarily *agree* with their co-referential nominal expressions at all. I am assuming a definition of agreement such that two coreferential elements x and y can be said to 'agree' if the morphosyntactic features expressed by x and y (for person/number/gender) are compatible.

---

43 However, as noted, this may be an epiphenomenon due to the common tendency in languages not to show overt agreement for non-humans except where these arguments are highly foregrounded.

44 Note that this use of the term 'anaphor' differs both from the traditional grammatical use and from the more recent use in Generative grammar. Traditionally, an 'anaphoric' pronoun literally 'refers back'. In Generative grammar (e.g. CHOMSKY 1981), 'anaphor' refers exclusively to reflexive and reciprocal pronouns. In BRESNAN and MCHOMBO's use, the term refers to pronominal elements which are also referring expressions, rather than mere agreement markers. I follow their usage here.

In Ngalakgan, both nouns and verbs can express the morphosyntactic features associated with the VEG and NEUT classes (and hence we cannot regard disagreement as an example of superclassing: where NC features are neutralised in some domain). Therefore, in those instances where a noun expresses features of one inanimate NC but the verb expresses the features of another, the two sets of NC features cannot be said to be compatible. Hence, these NC prefixes cannot be considered to be 'agreement' markers either. Rather, in referential terms they behave more like generic nouns and classifiers in other languages such as Eastern Arrernte and Ngan'gityemerri, which delimit a domain of function/use to which the external noun is being put in any particular environment.

Nevertheless, NC markers, and the third plural prefix, cannot be called 'pronominal anaphors', if what we mean by this term is elements which are both arguments and definite referring expressions, such as the anaphors of English ('he, she, it', etc). That is because, as pointed out by EVANS (this volume), the prefixes for both these categories can appear in a wide range of environments where they are co-indexed with entities which are indefinite, non-specific or non-referential in various ways. English anaphoric pronouns cannot refer to such entities (Evans claims), therefore the two - argument affixes, and anaphoric pronouns - are distinct. I concur with Evans that the 3pl affix should be termed an 'agreement affix', and should be considered to be analogous to such affixes in European languages and Warlpiri. Such affixes can commonly be ambiguous between agreement and argument interpretations. BITTNER and HALE (1995), for example, show that Warlpiri 3rd person clitics may, in the absence of a co-referential noun, be interpreted as anaphoric pronouns.

That is, *neither* of the commonly accepted characterisations of argument affixes - agreement markers, nor bound pronouns - seems appropriate to the characterisation of NC prefixes in Ngalakgan. Rather, I propose the term *bound pronominal generic* for these prefixes. Like generic nouns in other Australian languages pronominal generic affixes can (optionally) serve to delimit the function/use of an external noun (though they do not always do so), and in some cases NC argument prefixes in this function will clash with the morphosyntactic features of the co-referential noun. Again, like generic nouns, they are not necessarily referential in such constructions. Indeed, it's possible that they may be more important to the interpretation of the *predicate*, rather than the referent (as MITHUN 1984 suggests is the case for incorporated nouns). They may also be anaphoric when occurring independent of any co-referential noun (hence 'pronominal'). Pronominal generic affixes are like the 3pl prefix in this regard.

Table (5) summarises the essential properties and proposed terms for the various categories of argument prefixes in Ngalakgan.

Table (5): The argument prefix categories of Ngalakgan

Affixtype	Always agrees?	Always referential?	Allows co-reference?	Term
SAP	yes	yes	no	'bound anaphoric pronominal'
3pl	yes	no	yes	'pronominal agreement affix'
NC	no	no	yes	'pronominal generic affix'

Finally, we can ask whether the three- or four-way division of the argument affixes in Ngalakgan finds any universal support. I have already noted that SAP argument affixes in Ngalakgan may never have indefinite and non-specific readings, and that such a restriction is probably lexical in nature. I suspect that the restriction on indefinite or non-specific uses of SAP argument affixes is very common, if not universal, among polysynthetic languages at least, though few grammars explicitly discuss this asymmetry. In short, we expect to find that if any argument prefixes behave more like bound pronominal anaphors, it will be those indexing SAP referents. In addition, it also seems common for the 3rd person plural pronoun in languages to take on the functions of an indefinite, non-specific, or non-referential pronoun. This is the case in English, as well as Ngalakgan and Bininj Gun-wok.

I hope to have shown a way forward out of the debate between 'argument-ist' and 'agreement-ist' positions over the function of argument-indexing affixes. This will come from a recognition of the diverse referential and morpho-syntactic functions that argument affixes can play in many languages, not just those characterised as 'polysynthetic'.

Abbreviations

Boundary symbols, in order of increasing freedom and productivity:
+ ('root-level'), - ('word-level'), = ('clitic') see B. Baker (1999).

1, 12, 2, 3	1st excl., 1st incl., 2nd, 3rd person	GN	Gunwinyguan language family
ABL	ablative	INCH	inchoative
ALL	allative	INST	instrumental
AUX	auxiliary, finite verb stem	INT	interrogative clitic
BGW	Bininj Gun-wok	INTENS	intensified
COM	comitative	ITER	iterative
DAT	dative	LAT	lative
DEF	definite	LOC	locative
DIST	distributed	MASC	masculine noun class
DUR	durative	NC	noun class (MASC, FEM, VEG, NEUT)
ERG	ergative	NEUT	neuter noun class
F/FUT	future	NOM	nominative
FEM	feminine noun class	NP	non-past
GEN	genitive	O	object
p	plural	RED	reduplication

PC	past continuous	REL	relative/subordinator
PNEG/	past/	RR	reflexive/reciprocal
PRNEG/	present/	s	singular
FNEG	future negative suffixes	S	subject (that is, Agent of a transitive or Subject of an intransitive verb)
POSS	possessive		
SAP	speech act participant (1st and 2nd person referents)	SEMBL	semblative
POT	potential	SS	same subject
PP	past punctual, present perfective	VBSR	verbaliser
		VEG	vegetable noun class
PR	present		
QUOT	quotative		

REFERENCES

ALPHER, BARRY, NICHOLAS EVANS, AND MARK HARVEY (forthcoming): Proto-Gunwinyguan verbal suffixes. To appear in Nicholas Evans ed. *Studies in comparative Non-Pama-Nyungan.* Canberra: Pacific Linguistics.

ANDERSON, STEPHEN R. (1992): *A-morphous morphology.* Cambridge Studies in Linguistics. Cambridge: Cambridge University Press.

AUSTIN, PETER AND JOAN BRESNAN (1996): Non-configurationality in Australian Aboriginal languages, in *Natural Language and Linguistic Theory* 14: 215-268.

BAKER, BRETT J. (1999): *Word structure in Ngalakgan.* Unpublished PhD thesis, University of Sydney.

BAKER, BRETT J. (2000): *The referential function of the noun class prefixes of Ngalakgan in nominal expressions.* MS, University of Sydney.

BAKER, MARK C. (1988): *Incorporation: a theory of grammatical function changing.* Chicago: University of Chicago Press.

BAKER, MARK C. (1996): *The Polysynthesis Parameter.* Oxford: OUP.

BITTNER, MARIA AND KEN HALE (1995): Remarks on definiteness in Warlpiri, in: Emmon Bach, Eloise Jelinek, Angelika Kratzer, and Barbara H. Partee eds., *Quantification in natural languages.* Dordrecht: Kluwer. Pp. 81-106.

BOAS, FRANZ (1911): *Chinook.* In *Handbook of American Indian languages, Vol. 1.* Washington: Government Printer. Bulletin of the Bureau of American Ethnology 40. Pp. 599-677.

BRESNAN, JOAN AND SAM A. MCHOMBO (1987): Topic, pronoun, and agreement in Chichewa, in: *Language* 63: 741-82.

CHOMSKY, NOAM (1981): *Lectures on Government and Binding.* Dordrecht: Foris.

CORBETT, GREVILLE (1983): *Hierarchies, targets and controllers: Agreement patterns in Slavic.* London and Canberra: Croom Helm.

CORBETT, GREVILLE (1991): *Gender.* Cambridge: Cambridge University Press.

EVANS, NICHOLAS (1995): The syntax and semantics of body parts in Gundjeihmi, in: Hilary Chappell and William McGregor eds., *The grammar of inalienability: a typological perspective on the part-whole relation.* Berlin: Mouton de Gruyter.

EVANS, NICHOLAS (1997a): Role or cast? Noun incorporation and complex predicates in Mayali. In Alex Alsina, Joan Bresnan and Peter Sells eds., *Complex predicates*. Stanford: CSLI. Pp. 397-430.

EVANS, NICHOLAS (1997b): Head classes and agreement classes in the Mayali dialect chain, in: Mark Harvey and Nicholas Reid eds., *Nominal classification in Aboriginal Australia*. Studies in Language Companion Series. Amsterdam and Philadelphia: John Benjamins. Pp. 105-146.

EVANS, NICHOLAS, DUNSTAN BROWN AND GREVILLE CORBETT (1998): Emu divorce, in: M. Catherine Gruber, Derrick Higgins, Kenneth S. Olson, and Tamra Wysocki eds., *CLS 34*. Chicago IL: Chicago Linguistics Society.

EVANS, NICHOLAS (forthcoming): *A pan-dialectal grammar of Bininj Gun-wok: Mayali, Kunwinjku and Kune*. Canberra: Pacific Linguistics.

FOLEY, WILLIAM A. (1997): Polysynthesis and complex verb formation: the case of applicatives in Yimas, in: Alex Alsina, Joan Bresnan, and Peter Sells eds., *Complex predicates*. CSLI Lecture Notes No. 64. Stanford, CA: CSLI Publications.

HARVEY, MARK (1997): Head and agreement classes: an areal perspective, in: Mark Harvey and Nicholas Reid eds., *Nominal classification in Aboriginal Australia*. Studies in Language Companion Series. Amsterdam and Philadelphia: John Benjamins. Pp. 147-164.

HEATH, JEFFREY (1978): *Ngandi grammar, texts and dictionary*. Canberra: Australian Institute of Aboriginal Studies.

HEATH, JEFFREY (1984): *Functional grammar of Nunggubuyu*. Canberra: Australian Institute of Aboriginal Studies.

HEIM, IRENE AND ANGELIKA KRATZER (1998): *Semantics in generative grammar*. Malden, MA and Oxford: Blackwells. Blackwell Textbooks in Linguistics Series.

JELINEK, ELOISE (1984): Empty categories, case, and configurationality, in: *Natural Language and Linguistic Theory* 2: 39-76.

LYONS, CHRISTOPHER (1999): *Definiteness*. Cambridge: Cambridge University Press.

MERLAN, FRANCESCA (1983): *Ngalakan grammar, texts and vocabulary*. Canberra: Pacific Linguistics.

MITHUN, MARIANNE (1984): The evolution of noun incorporation, in: *Language* 60: 847-93.

NORDLINGER, RACHEL (1998): *Constructive case: evidence from Australian languages*. Stanford, CA: CSLI.

REID, NICHOLAS (1997): Class and classifier in Ngan'gityemerri, in: Mark Harvey and Nicholas Reid eds., *Nominal classification in Aboriginal Australia*. Studies in Language Companion Series. Amsterdam and Philadelphia: John Benjamins. Pp. 165-228.

RIZZI, LUIGI (1986): On the status of subject clitics in Romance, in: Osvaldo Jaeggli and Carmen Silva-Corvalán eds., *Studies in Romance linguistics*. Dordrecht: Foris. Pp. 391-420.

SIMPSON, JANE (1991): *Warlpiri morpho-syntax: A lexicalist approach*. Dordrecht: Kluwer.

WILKINS, DAVID P. (2000): Ants, ancestors and medicine: a semantic and pragmatic account of classifier constructions in Arrernte (Central Australia), in: Gunter Senft ed., *Systems of nominal classification*. Cambridge: Cambridge University Press. Pp. 147-216.

Syntax and morphology of polysynthesis in the Georgian verb[1]

WINFRIED BOEDER

1. Georgian is one of the almost sixty autochthonous Caucasian languages, and it is the one with the highest number of speakers (3.5 to 3.8 million, depending on whom you count as speakers of Georgian) and with the oldest literary tradition. It should therefore offer us the unique opportunity to study a history of 1,500 years of polysynthesis. But as far as personal verb-marking is concerned, Georgian has been stable over this time, and its related languages, Laz, Mingrelian and Svan, seem to allow the conclusion that Georgian person-marking largely mirrors the system of the South Caucasian or Kartvelian protolanguage. Changes refer to the loss or extension of third person plural object suffixes and to third person subject suffixes, but not to the subject and object prefixes which I will discuss here. This stability of prefixation over thousands of years is no less remarkable a phenomenon than the better-known "natural" preference for suffixation in other domains of Georgian morphology.

The purpose of this paper is to provide a better understanding of verbal agreement in a polysynthetic language like Georgian. After a short outline of verbal morphology (§ 2), I will describe some constraints on the combination of subject and object markers (§ 3), which lead to a reinterpretation of the dummy head noun *tav-* 'head' which is often described as a reflexive pronoun (§ 4). The special problem of indirect object reflexives raises some questions about the nature of agreement and its status in grammar (§ 5). Finally, I will point to the tradition in which these questions are embedded (§ 6).

---

1 This is a substantially revised version of the article (Boeder 1999). I am extremely grateful for the list of errors, mistakes and improvements kindly provided by George Hewitt (School of Oriental and African Studies, London), in particular for his comment on example (36), which unfortunately I had taken to be grammatical in the former version - a correction that may have major theoretical implications. - Many thanks go to the editor Nicholas Evans for his very careful corrections and for his invaluable insistence on much-needed clarification in some places. - I would also like to thank my colleague Kevin Carpenter for carefully correcting my English.

2. In comparison with its northwestern Caucasian neighbour languages (Abkhaz, Cherkez etc.), Georgian has a rather modest form of polysynthesis in the verb. In Standard Modern Georgian,[2] most verb forms can be segmented into agglutinatively concatenated morphs whose shape is phonetically very stable and whose boundaries are not blurred. (Note, however, that some of these well-delimited segments of the verb are portmanteau morphs.).

Let us take some simple examples (in fact, (2) is the utterance of a young child (age 1,10,18)!):

(1) a) *me gada- v- ċer ma-s*
I Prev-1S-write it-Dat
'I will copy it'

b) *šen m- nax-av me*
you 2O-see-TS I
'you will see me'

(2) *did-i* *tevz-i*
big-Nom fish-Nom

mo- m- i- ṭan- a babua-m Tamaz[i]-s[3]
Prev-1O-OV-bring-3S(Aor) grandfather-Erg Tamaz-Dat
'Grandfather brought me, Tamaz, a big fish' (IMEDA3E – TUITE 1992: 63)

In the pronunciation of these and other forms, morphophonemic changes are almost absent. Word order is "free", although, for instance, the post-verbal position of the clitics *ma-s* and *me* in (1) is the most natural one, and the word order of (2) is determined by information structure. The order of verbal morphemes does not necessarily mirror clausal word order.[4] Normally, it cannot mirror this order at all, since Georgian seems to have one and only one slot for subject and object prefixes which is occupied by *m-* 'me' in (1b) and (2). On the face of it, the term "polypersonal" may even look misleading in the case of Georgian verbal person-marking.[5] But let us look more closely at the structure of Georgian verb forms.

In contrast to nominal forms, whose inflectional affixes are exclusively suffixal, verbs have both suffixes and prefixes. Disregarding interdependencies between morphemes and combinations with the copula, the following sequential positions may be distinguished:

---

2 For a concise and very clear survey of Georgian grammar, see IMEDA3E – TUITE 1992: 44-54.

3 This is the only non-adult form in (2): the child has reanalyzed the nominative *Tamaz-i* 'Tamaz-Nom' as a stem by analogy with the pattern of (adult) *Giorgi-ø* 'Giorgi-Nom', *Giorgi-s* 'Giorgi-Dat'. The correct adult form would be *Tamaz-s*.

4 As it does to some extent in Abaza; see ALLEN 1956.

5 On this disparity between valency and person marking see e.g. GAMQRELI3E 1979; 1981.

1. Preverb I: *uḳu-* 'back'
2. Preverb II, e.g. *še-* 'into, onto'
3. Preverb III: *mo-* 'hither, German *her-*'
4. Subject marker, e.g. *v-* 'I'
5. Object marker, e.g. *m-* 'me'
6. Version vowel: *a, e, i* or *u*
7. Root, e.g. *naxv-* 'see'
8. Passive/inchoative marker, *-d*
9. Thematic suffix, e.g. *-eb*
10. Participle formant: e.g. *-ul-*
11. Causative formant: *(-ev)-in*
12. Causative thematic suffix
13. "Extension marker",[6] e.g. *-d*
14. "Paradigm marker",[6] e.g. *-i*
15. portmanteau 3rd person subject number marker
16. Object Plural marker

Table 1: Morphological slots of the Georgian verb

Two additional examples[7] might give an impression of what a Georgian verb form can look like:

(3) *uḳu- še- mo- g- a- brun- a*
back(1)-PrevII(2)-PrevIII(3)-2O(5)-NV(6)-(re)turn(7)-3Sg(Aor)(14)
'(s)he made (Aor) you return here'

(4) *da- g- a- ç̣er- in- eb- d- a- t*
PrevII(2)-2O(5)-NV(6)-write(7)-Caus(11)-CausTS(12)-Ext(13)-3Sg(Past)(14)-Pl(O)(15)
's/he would make (Cond) you(Pl) write it'

---

6 "Extension (marker)" (13.) is FÄHNRICH's 1991 translation of A. Shanidze's term: *savrcobi* (ŠANIʒE 1973); ARONSON 1991 uses "imperfect/conditional marker" and includes *-n-* of some of the perfect forms (*uq̇ep-n-ia* 's/he/it has (apparently) barked') and the copula in complex verbs in this position. "Paradigm marker" (14.) is again taken from Fähnrich; ARONSON 1991, following A. Shanidze (mç̣krivis nišani), uses "screeve marker". VOGT 1971 calls the paradigm marker *–e* of the "weak aorist": "suffixe thématique", and *–i* of the "strong aorist": "voyelle d'appui"; HEWITT 1995 calls *-e* "indicative".

7 The examples are taken from DAMENIA 1982, who offers a complete survey.

I will not discuss these "slots" in detail. Suffice it to say that one important feature of Georgian verb morphology is that in some cases the presence or absence of one affix requires the presence of another.[8] For example, the causative suffix *–in-* (slot no. 11) in the last example requires the presence of a version marker (slot no. 6), which is *–a-* in the default case.

Of primary interest here are the personal prefixes of the slots 4, 5 and 6 and, to some extent, the number markers 15 and 16. These markers are displayed in table 2:

Person		Singular	Plural		
(a)	1	v / Ø___	v / Ø___t		morpho-
	2	Ø / x___	Ø / x___t		logical
	3	___ a / o / s	___n / an / en / es / nen		subject
(b)	1	m___	gv___		
	2	g___	g___t		
(c)	3	s / h / Ø___(t)			
(d)	1	[Ø-]i___		reflexive	
	2	[Ø-]i___		non-	
	3	[Ø-]i___		intensified	
(e)	1	m-i___	gv-i___	non-	
	2	g-i___	g-i___	reflexive	
	3	Ø-u___(t)			
(f)	1	m-e___	gv-e___		
	2	g-e___	g-e___t		
	3	Ø-e___(t)			
(g)	1	m-a___	gv-a___		
	2	g-a___	g-a___t		
	3	Ø-a-___(t)			

beneficiary; specified indirect object; morphological object

Table 2

8 Note that circumfixation is one form of Georgian nominal derivation, e.g. *Sa-kartvel-o* 'Prefix-Georgian-Suffix' = 'Georgia'.

The leftmost column of numbers refers to person. The items in section (a) are subject markers, the rest are object markers of different types. Note that the terms "morphological subject" and "morphological object" are defined in terms of their markers in the verb: morphological subjects are referred to by the markers in (a), morphological objects by those in (b)-(g). The whole question of a definition of subject, direct object and indirect object in a syntactic or semantic sense cannot be discussed here, and it is largely irrelevant to the issue of slot-filling and its syntactic corollary.

Let us look at some properties of the morphological system tabulated in table 2.

Firstly, consider the coding of subjects in section (a): the opposition between 1st and 2nd person on the one hand and 3rd person on the other is doubly coded by the presence vs. absence of a prefix and by the absence vs. presence of what is called a 3rd person suffix,[9] which is a portmanteau morph coding 3rd person, number and tense at the same time. This differential coding of 3rd vs. non-third person subject is of interest here because subject prefixes are often non-overt (see below sequence constraint Ia)). Non-overtness (of *x–* and *v–*, for example) has to be distinguished from absence (non-existence according to table 2).

Secondly, 1st and 2nd direct and indirect object markers are the same, but there is no 3rd person direct object marker: section (c) refers to indirect objects only. Number is a different matter: Old Georgian, but not Modern Georgian, had a 3rd person direct object plural marker *–n* :

(5) a. Modern Georgian
mi- s- c- a ma-s
Prev-3IO-give-3Sg(Aor) s/he-Dat
's/he gave(Aorist) **it** to him/her' **or** 's/he gave(Aorist) **them** to him/her'

b. Old Georgian
mi-s-c-a 's/he gave **it** to him/her' vs.
*mi- s- c- **n-** a ma-s*
Prev-3IO-give-**Pl(O)**-3Sg(Aor) s/he-Dat
's/he gave **them** to him/her'.

Plural marking in Modern Georgian cannot be discussed here. Roughly speaking, the plural *t* bracketed in (c), (e)-(g) occurs with dative subjects and with those dative indirect objects that are higher on the animacy/empathy/agentivity scale than the "morphological subject" (BOEDER 1989: 171; 181).

Thirdly, indirect objects, which do have 3rd person markers, are differentiated into subtypes by the presence or absence of a "version vowel", which occurs in four different forms:[10] *i* in sections (d) and (e), *u* in section (e), *e* in section (f) and *a* in section (g). The

9 With many verbs, this contrast is also marked by ablaut.

10 For an extensive discussion, see BOEDER 1969.

vowelless forms in sections (b) and (c) are neutral indirect object markers, while, for instance, *i* and *u* typically mark beneficiaries, and so on. These subtypes need not concern us here, except that the opposition between simple, vowelless indirect object markers and indirect object markers specified by a version vowel will be relevant for the slot-filling hierarchy discussed below.

Fourthly, person markers are often described as a kind of personal or anaphoric pronoun inside the verb (see section 5. below). But third person verb subjects need not be definite. 3rd person markers can be indefinite in conditional clauses and potentialis contexts (BOEDER 1989: 162):

(6) Modern Georgian (from a fairy tale)
tu Ø- e- codineba, ima-s Ø- e- codineba baṭ-is garemoeba-Ø
if 3IO-EV-knows, he-Dat 3IO-EV-knows goose-Gen surrounding-Nom
'if anybody knows, he will know the whereabouts of the goose'

Compare the following examples from Old Latin (7) and Old Greek (8)-(9)(Wackernagel 1926: 112):

(7) *si in ius vocat, ni it, antestamino* (XII Tables, 5th cent. BC)
if in law he.calls, if.not he.goes, *antestatio*.should.be made
'if **somebody** brings **somebody** else into court, and he does not go, he shall call witnesses'

(8) a. *θυγατρὶ ἢν διδῶ* (Cretan inscription)
daughter.Dat if he.gives
'if **somebody** gives (a dowry) to his daughter'

b. *οὐδέ κεν ἔνθα τεόν γε μένος καὶ χείρας ὄνοιτο* (Homer N 287)
and.not Particle there your Particle might and hands he.would.blame
'[If we were chosen for an ambush fight] **nobody** would blame then your force and your hands'

(9) a. Old Georgian (Matthew 7,8)
romel-i x-[11] *e- ʒieb- n pov- i- s*
who- Nom 3IO-EV-seek- Frequ(Pres) find-Frequ(Aor)-3S(Sg)
'he who seeketh findeth'

11 One of the oldest versions of the Gospel, the so-called Khanmeti text, has *x-* instead of *s-/ h-/Ø-* as a 3rd person indirect object marker.

b. Modern Georgian (proverb)
kurd-ma kurd-s mo-h-p̣ar-a da ɣmertma gaicina-o
thief-Erg thief-Dat Prev-3IO-steal-3S(Aor) and God laughed-Quot
'A thief stole **something** from a thief, and God laughed.'

In Georgian, then, person markers are not anaphoric by themselves, and in this they differ from what are commonly considered "3rd person pronouns" in Georgian. As pointed out by N. EVANS (this volume), they are rather semantically unspecific in some languages and lack the "expressive possibilities" of free pronouns. It remains to be seen which non-specific arguments can be coded by 3rd person markers (or zero, for that matter). It is clear that only non-intensified indefinites do in Georgian,[12] and that the "semantic threshold" for this type of marking is much higher than in the Australian languages described by N. Evans.[13] What seems to be required is some degree of specificity based on a context that introduces an imagined referent into the universe of discourse, but whithout implying an impossibility or arbitrariness of choice as in: [Giorgi needs a birthday present:] *We must buy **something** for him* (where Georgian **must** have *rame* 'something' = Russian *čto-nibud'*; here 3rd person object marking would be insufficient).

3. What happens if subjects and objects are combined? As I have said above, Georgian verbs are polypersonal in the sense that subjects and different objects are coded in the verb. On syntactic grounds, I will argue that the template of the Georgian verb must contain one slot for subject prefixes and one for object prefixes.

However, prefixed person markers are almost always restricted to one overt prefix, and suffixed person and number markers to one overt suffix in Modern Georgian. The following morphological and morphosyntactic rules concur to produce this effect:[14]

12 For a study of the scale of indefiniteness, its emphatic, "expressive" members ("who/whatever you want", "God knows who/what"), and their bleaching see HASPELMATH 1991.

13 For instance, "maybe he will marry again" requires an explicit direct object: *vin icis, ikneba ḳvlav šeirtos col-i* (RḲ) 'who knows, it.will.be again that.he.marries wife-Nom'.

14 There are many proposals in the "western" literature on this topic that try to capture the regularities of person and number coding in the Georgian verb and its concomitant "tavization" (see 4. below), e.g. DEETERS 1931:41-42; VOGT 1971: 84-87; BOEDER 1969: 85-86; ANDERSON 1984; 1992: 137-156; HEWITT 1995: 128-142; HARRIS 1995: 1382. One of the best is BRAITHWAITE 1973, who discusses the relevant facts in terms of "output constraints" that block concord. Similarly, ANDERSON (1984: 209) invokes the concept of "filter".

I. Sequence constraints

(a) Prefix sequence constraint
There is only one prefix slot available for **overt** subject and object markers. Prefixed subject markers are non-overt either with lexically specified verbs or in the presence of overt prefixed object markers.

The only obvious exception is the combination of 1st person subject plus 3rd person indirect object marker in literary Old and Modern Georgian (and in some modern dialects), e.g. in:

(10) *v-s-ç̣er* (spoken Georgian: v-ç̣er)
1S-3IO-write 'I write it (to) him/her'.

But the subject marker *x-* is non-overt in Modern Georgian, except with the verb "to be" and "to go" (*x-ar* 'you are', *mo-x-val* 'you will come', etc.), and *v-* is non-overt before the object marker *g-*, e.g. in:

(11) *minda g- a- ḳoc-o, m- a- ḳoc-o* (from a poem by Važa Pšavela)
I.want.it 2O-NV-kiss-Opt, 2O-NV-kiss-Opt
'I want to kiss you, [and] you to kiss me'

As we will see, these markers are "suppressed", not "disposed of" (to use ANDERSON's (1984) terminology).

Note that this constraint is not of a phonological nature: the expected forms

(11') **v- g- a- ḳoc-o, (x-) m- a- ḳoc-o*
1S-2O-NV-kiss-Opt, (2S-)1O-NV-kiss-Opt

are perfectly possible from the point of view of phonotactics,[15] but they do not occur, neither in Old Georgian nor in Modern Georgian. It is a morphological constraint, and it is superficial in the sense that from a syntactic point of view, the 1st and 2nd person markers count as being present in the verb. Therefore, they are in a sense zero allomorphs (entered as Ø in table 2),[16] and must be present at one level of representation.[17] The "suppression" of verbal

15 Cf. *v-galob* 'I sing', *xma* 'voice'; see BRAITHWAITE 1973: 66.

16 HARRIS (1981: 31) interprets the non-overtness of subject markers as deletion. I prefer a zero allomorph interpretation because it accounts for the syntactic availability of non-overt markers. ANDERSON's (1984) interpretation (in which affixes are spelt out by disjunctively ordered rules) presupposes an availability of the subject entities (features?) on a different level of morphosyntax.

markers may have a functional motivation. As AMIRI3E (1998) points out on the basis of R. Langacker's empathy hierarchy, it is the less prototypical (less expected) arguments (3rd person (agent) subjects and 1st and 2nd person (patient) objects) that are preferrably marked by an overt verbal affix.[18]

The suffix sequence constraint is different in many respects:

> (b) Suffix sequence constraint
> Object plural markers oust the subject 3rd singular marker *-s*,
> and are usually ousted by subject 3rd person plural markers.

For example:

(12) *g- xaṭ- av- en* (Old Georgian g-xaṭ-av-en-t)
2O-paint-Pres-3Pl 'they paint you(Sg/Pl)'
g- xaṭ- av- t (Old Georgian g-xaṭ-av-s-t)
2O-paint-Pres-Pl(O) 's/he paints you(Sg/Pl)'

vs. *da- g- xaṭ- a-* t
Prev-2O-paint-3S(Aor)-Pl 's/he painted you(Pl)'

In contrast to the prefix constraint, the suffix constraint did not exist in Old Georgian. The modern form is the result of morphological simplification.[19] This simplification entails no communicative problems, in spite of so-called "pro-dropping". Suffixes will not be considered for the rest of this paper.

As will become clear later in this paper, these sequence constraints have no syntactic impact, and this distinguishes them from the following constraint which restricts object coding in the verb:

> II. Prefix slot-filling constraint
> In the morphological template of the Georgian verb, there is one, and only one,
> slot for object markers.

---

17 Note that subject marking *v-* makes its first appearance after its object counterpart *m-* in language acquisition (at the end of the second or the beginning of the third year; IMEDA3E - TUITE 1991: 58-64). As is pointed out by IMEDA3E and TUITE, this order may have different reasons, e.g. similarity of *m-* '1O' with *me* 'I', frequency of occurrence of a class of indirect verbs with *m-* (ib. pp. 63-64). But one additional factor could be transparency (code simplicity: phonological stability of object marking) vs. allomorphy involving zero.

18 Similar ideas are found in ASATIANI 1994.

19 I hesitate to interpret it as a phonotactic phenomenon, although word-final *n-t#* and *s-t#* occur only rarely in Modern Georgian.

In other words, one, and only one, object can be coded in the verb, and all other objects are confined to a verb-external coding.

Now, if there is only one slot available, how is the competition problem solved if there is more than one object? The answer is that the object slot is filled according to the following hierarchy:[20]

> III. Slot-filling hierarchy
> specified indirect object > 1st/2nd person indirect object > other objects

That is, specified indirect objects take precedence over 1st and 2nd person indirect objects, which in turn take precedence over other objects. This hierarchy combines a functional role hierarchy with the person hierarchy.[21] Consider a sentence such as:

(13) *tvalebi g- i- gav- s maq̇val- sa* (from a folk-song)
eyes 2O-OV-resemble-3Sg blackberry-Dat
'your eyes are like blackberries' (lit.: 'eyes resemble you to blackberry')

In this sentence, there are two indirect objects: 'you' and 'blackberry'. But only the former is a specified indirect object, because it is marked by the version vowel *–i* in the sequence *g-i-* 'to you'. The specified indirect object 'you' takes precedence over the non-specified 3rd person indirect object 'blackberry' which cannot be marked in the verb. Notice that precedence in this example is based on which argument is specified not on person: the same precedence relation holds for 3rd person specified indirect objects (as in (lit): 'eyes resemble her to blackberry', in the sense of: 'her eyes are like blackberries'). - Similarly, a quadrivalent verb[22] as in:

---

20 Although most of the "facts" of person marking are covered in the earlier literature, BRAITHWAITE (1973: 73) seems to have been the first to introduce the idea of "hierarchical ranking based on Function (Benefee, Object, Subject)" and on the "peer" relation between 1st and 2nd person on the one hand and the non-peer relation between 1st/2nd person and 3rd person on the other. "We might, metaphorically, characterize the situation as one in which the various UAS [= underlying abstract structure] functions/persons "jockey" for the privilege of concord in the FVF [= finite verb form]. The fate of any "loser" is not uniform for all eligible "competitors" in a particular SV(B)O [= subject verb (benefee) object] syntagm, nor will the particular entity (e.g. a particular Person) that loses in one specified syntagm (e.g. a Series I [= present tense group] collocation) also lose out in another (say Series III [= perfect tense group])" (BRAITHWAITE 1973: 65-66).

21 See BRAITHWAITE 1973: 75.

22 Avtandil Arabuli (Georgian Academy of Sciences, Tbilisi) has kindly provided me with an example of even higher "valency": *mi-g-i-šver-in-e (me šen) ǯox-i mc̣q̇ems-s xbo-s-tvis / xbos-s* Prev-2O-OV-hold.out-TS (I you) stick-Nom sheperd-Dat calf-Gen-for / calf-Dat 'I made the shepherd hold out the stick at your calf'; cp. Khevsurian: *sam-s da-g-i-ɣer-eb-av qel-sa* (A.

(14) *Tamaz-i* *m- i- čm- ev- s* *švil- s* *pur-s*
Tamaz-Nom 1O-OV-eat-Caus-3Sg child-Dat bread-Dat
'Tamaz makes the child eat bread for me'; 'Tamaz feeds my child bread'[23]

does not allow any object to be marked in the verb except the specified indirect object, *m-i-* 'for me'. (Note that causees are indirect objects in Georgian, as in many other languages.)[24] In a sentence like 's/he commits me to you', the indirect object (*-g-* '(to) you') takes precedence over the direct object ('me'):

(15) *ča- g- a- bar- eb- s* *šen* *čem-s* *tav- s*
Prev-2O-NV-commit-TS-3Sg you my-Dat head-Dat
's/he commits me to you'

where the 2nd person indirect object ('you') is coded in the verb and the dative direct object 'me' is not. The direct object must appear in a syntactically different form: *čem-s tav-s* 'my head', that is the dummy head noun 'head' which we will consider below. An example with different case alignment (ergative - nominative with an aorist (optative) verb form) is:

(16) *xširad gv-a-nax-o-s* *ɣmert-ma tkven-i* *tav-i* (MD)
often 1PlO-NV-see(Caus)-Opt-3SgS God-Erg your-Nom head-Nom
'May God let us often see you'

where only the indirect object causee 'us' is coded in the verb.

Činčarauli, Xevsurulis taviseburebani p. 328,15) three-Dat Prev-2O-OV-hold.out-TS-Quot hand-Dat 'I will reach for three things'.

23 For more examples of this type see Šanize 1973 § 402; Vogt 1971: 123.

24 It will be seen that my evidence for the slot filling hierarchy is incomplete. No example is given for the combination: "specified indirect object + 1st/2nd indirect object." Quadrivalent verbs should yield sentences like (14'): 'Tamaz makes you (indirect object causee) eat bread (direct object) (for) me (specified indirect object beneficiary)'. But as far as I can see now, these combinations seem to be unacceptable, and the corresponding examples given in an earlier paper (Boeder 1989: 175; 183, note 21) were probably based on the compliance of too liberal an informant. Demotion of the beneficiary indirect object is the preferred coding strategy in this case: as in English, an adpositional phrase is used instead (*čem-tvis* 'me-for' = 'for me'), which avoids the complexity, because postpositional phrases are not eligible for intraverbal coding. However, the expected combination does occur under specific conditions: *Vaxtang-i m-i-rčevni-a šen-s tav-s* 'Vakhtang-Nom 1O-OV-prefer-3Sg your-Dat head-Dat' (lit. something like: 'for me Vakhtang is preferrable to you' = 'I prefer Vakhtang to you'), where the "specified indirect object" (*m-i-* in our example) has subject properties (for example control and number agreement). The whole issue need not be discussed here since the problem of slot competition is independent of these details.

Under certain conditions, "other objects" leaves us with a choice. In: 'he will sell me to him', both objects are eligible for the slot. If the 1st person direct object ('me') is chosen, the form will be:

(17) *mi- m- qid-i- s me ma-s*
Prev-1O-sell-TS-3Sg I he-Dat
's/he will sell me to him/her'[25]

Alternatively, we may have:

(18) *mi- Ø- qid- i- s ma-s čem-s tav-s*
Prev- 3IO- sell-TS-3Sg s/he-Dat my-Dat head-Dat
's/he will sell me to him/her'

where Ø marks the 3rd person indirect object, whereas the direct object is not coded in the verb and must appear again in a syntactically different form. It remains to be determined whether the person hierarchy normally overrides the functional hierarchy, that is, whether (17) is the preferred variant.[26] Example (18) illustrates a further point: the 3rd person object has a non-overt, phonologically empty marker. Yet zero in the sense of table 2 is not absence, since zero marking prevents other object markers from filling the same slot (here: the marker for 'me'). Note that one archaic dialect of Old Georgian had *x-* instead of zero (see (6b)), and the morphosyntactic rules of prefix person marking has not changed essentially during the last 1,500 years.

So the coding of objects in the verb has syntactic consequences: the presence or absence of a verbal marker for one object entails a specific form of other object noun phrases. What is this form?

4. As the arguments (subjects and objects) of a verb are more or less coded in the verb itself, most verb forms can function as self-sufficient clauses, and they often do: (19) (a), (b) and (c) are complete sentences. In terms of generative grammar, they result from "pro-drop" (or "stripping" in K. Braithwaite's terminology), and (a'), (b') and (c') represent the "full" form:

(19) (a) *v- čam*
1S-eat 'I eat it'

---

25 G. Hewitt has reminded me of the fact that even in the absence of *me mas*, *-m-* in (19) cannot refer to an indirect object ('to me') because the preverb *mi-* indicates a direction towords a 3rd person (roughly speaking, *mi-* corresponds to German *hin-* and contrasts with *mo-* 'her').

26 For some data and observations see ŠANIǯE 1973: 333 §400; VOGT 1971: 86-87; 122; VAMLING 1988: 316-318; BOEDER 1989: 181-182.

(a') *me v- ċam ma-s*
I 1S-eat s/he/it-Dat 'I eat it'

(b) *m-ċam*
1O-eat 'you eat me'

(b') *šen m-ċam me*
you 1O-eat me 'you eat me'

(c) *ċam-s*
eat-3S(Sg) 's/he eats him/her/it'

(c') *is ċam-s ma-s*
s/he(Nom) eat- 3S(Sg) him/her/it 's/he eats it'

(c'') *is ċam-s q̇vel-s*
s/he(Nom) eat- 3S(Sg) cheese-Dat 's/he eats cheese'

But there is a difference between the personal pronouns *me* in (a') and (b') and *šen* in (b') on the one hand, and case-marked pronouns like *is* in (c') *ma-s* in (a') and (c'), which parallel *q̇vel-s* in (c''), on the other. *me* and *šen* are verb-external counterparts of verb-internal 1st and 2nd person markers, and they are underspecified in that they do not show the subject-object contrast of the latter (*v-* in (a) vs. *m-* in (b)). They are not case-marked, whereas appositions are (BOEDER 1979: 438):

(19a'') *me šeċuxebul-i v-ċam ma-s*
I worried-Nom 1S-eat it-Dat 'I eat it, worried'

'Worried' is a case-marked apposition of the verb-internal pronoun *v-*, and *me* is its verb-external counterpart simply marking contrast, etc., as do other verb-external pronouns in "pro-drop" languages. Now the following generalizations seem to apply:

IV. Generalization on morphological case-marking
Nominals (pronouns and nouns) that cannot be coded in the verb must be coded in verb-external case-marked noun phrases.

Consistent with my interpretation, *is*, *ma-s* and *ɋvel-s* in (a'), (c') and (c'') are case-marked noun phrases, but while *me* (in (19a'') is a noun phrase,[27] it is not case-marked.

With this in mind, we are now in a position to formulate a rule for the *tav*-headed noun phrases of (15) and (18):

> V. Generalization on dummy heads
> 1st and 2nd person, reflexive and intensified pronouns cannot be case-marked. In case-marked noun phrases without head nouns, they become possessive determiners, and a dummy head noun *tav-* 'head' is supplied (to which generalization IV applies).

So we get for instance *čem-s tav-s* (see (18)). The singular possessive pronouns are *čem-* 'my', *šen-* 'your', and *tavis-* elsewhere (for non-1st, non-2nd reflexive and intensified pronouns).

There are some exceptions to V which cannot be discussed here. For instance, although they are not coded in the verb, some, mostly non-argumental, noun phrases occur without *tav-*: complements in identificational sentences (as in: *me rom šen viɋo* I if you were 'if I were you'[28]) and forms like *čem-it* 'my-Instr' = 'by myself', *tavis-it* 'by him/her/itself' where an underlying head noun seems likely;[29] for the 2nd person vocative see below.[30]
IV and V jointly "trigger" the use of the *tav*-headed noun phrases in examples such as:

(20) *rac mo- g- i- va dav- ita- o,*
whatever Prev-2O-OV-will.come dispute-Instr-Quot,
ɋvela ***šeni tav- ita-*** *o* (proverb)
all **your head-Instr**-Quot
'... it all happens to you **by/because of yourself**, it is said'

where 'by yourself' has no counterpart in the verb *mogiva* 'it happens to you'.

---

27 Coordination requires a noun phrase status of *me* for instance in: *me da čem-ma kmar-ma v-i-angariš-e-t* I and my-Erg husband-Erg 1S-NV-calculate-TS-PIS = 'I and my husband calculated (and nothing fell to my brother-in-law's share)' (proverb).

28 Interestingly, the 19th century writer Ilia Chavchvadze uses a possessive form: *me rom šen-i viɋo* I if your-Nom were 'if I were you'; see BOEDER 1989: 183.

29 Compare *tu me čem-it mo-g-a-çon-e-t tav-i* (Ilia Chavchavadze) if I my-Instr Prev-2O-NV-please(causative)-TS-PIS head-Nom 'if I ingratiated myself with you by myself' with Old Georgian: *me tav-it čem-it viṭɋw* (John 7,17) I head-Instr I.say.it 'I speak of myself'.

30 One important issue is raised by Old Georgian examples like: *mɣdel-t moʒɣuar-ta mo-m-c-es me šen* (John 18,35) priest-PlObl master-PlObl Prev-1O-give(aorist)-3PlS I thou 'the chief priests have delivered thee unto me', where Modern Georgian must have *šen-i tav-i* 'your-Nom head-Nom' instead of *šen*. It is not clear how far such examples are the result of translational calquing. For a detailed study of this phenomenon see ŠALVAŠVILI 1992.

Some comments on IV and V are now in order.

Firstly, pronouns are determiners; some are case-marked (*is*, *mas* in (19) (a'), (c')), others are not, for instance *me* in (19) (a'), (b'), but also *še* in (21a) and *me* in (21b):

(21) (a) *še* *mamaʒaɣl-o!*
you(Sg) son.of.a.bitch-Voc!
'you son of a bitch!'

(b) *me mamaʒaɣl-s* *da-m-a-viçq̇-d-a* *es* *çeril-i*
I son.of.a.bitch-Dat Prev-1O-NV-forget-Inchoative-3SgS this.Nom letter-Nom
'I, a son-of-a-bitch, forgot this letter!'[31]

But why is it that the non-case-marked pronouns in V are possessive determiners instead of simple determiners as in (21)? There seems to be a restriction on 1st and 2nd person, reflexive and intensified pronouns: they must occur in an argument position or have a counterpart in an argument position. As long as they occur in a verb, they are its arguments; in sentences like (19)(a'), *me* has a verb-internal counterpart, and the possessive pronoun in *čem-s tav-s* in (18) is not just a determiner, but functions as the equivalent of a genitive noun phrase in argument position. As a result, Georgian noun phrases with possessive pronouns mirror the structure of the verbal complex with a major category as its head and the pronouns as its sister categories. *tav*-headed phrases are a device providing dummy head nouns with extra argument positions.[32]

Secondly, the use of the dummy head noun is one single phenomenon, not the coincident output of two distinct operations: reflexivization and so-called "object camouflage",[33] which cover only part of the uses anyway (they do not cover examples like (20)!). And once one assumes that *tav*-headed noun phrases are not (necessarily) reflexive, their occurrence in a subject position is not particularly surprising:[34]

(22) *čem-ma tav-ma* *m- a- iʒul- a* *me*
my-Erg head-Erg 1O-NV-force-3S(Aor) I
'It was me who forced myself'

---

31 BOEDER 1985: 64-64.

32 BOEDER 1989: 176, on the basis of a suggestion I owe to L. Máracz.

33 HARRIS (1981: 51) describes "Object camouflage" as follows: "If a clause contains an indirect object, a first or second person direct object in that clause is realized as a possessive pronoun + *tavi*, where the possessive reflects the person and number of the input form". She argues that "Object Camouflage and *Tav*-Reflexivization cannot be accounted for with a single syntactic rule" (ib. 52). See also ANDERSON 1984: 209.

34 ASATIANI 1982: 89; BOEDER 1989: 169.

While the head noun *tav-* controls verbal agreement (the verb bears a 3rd, not a 1st person subject marker),[35] it is irrelevant for anaphoric relations: it is the possessive determiner, and not the head noun *tav-*, that bears the referential index.[36]

Thirdly, note that IV and V presuppose an assumption we made earlier about 1st and 2nd person verbal subject markers, namely that they are present for the purpose of syntax (that they are syntactically available), although they do not surface as phonologically non-empty segments. Syntax presupposes for instance the form (11'); if the non-overtness of *v-* in (11) (according to constraint Ia)) were syntactically relevant, the rules IV and V would yield:

(11'')**čem-ma tav-ma g-a- ḳoc-o-s*
my-Erg head-Erg 2O-NV-kiss-Opt-3S(Sg)

The distribution of these "hidden" person markers justifies the positing of zero subject marker allomorphs in the presence of 1st and 2nd person object markers. The positing of zero subject markers allows a unified account of the occurrence of the dummy head noun.

There are three instances where the joint conditions of IV and V are met.

A. The pronoun is not a verb-internally codable argument but some adverbial expression like *šeni tav-ita* 'by yourself' in (21) or the self-adressing vocative in:

(23) ***čem-o tav-o,*** *bedi ar giçeria* (from a poem by A. Çereteli)
my-Voc head-Voc, fate not it.is.written.for.you
'**O me**, you aren't granted good fortune'.

The address or call:

(24) *šen* 'you!' (not: *šen-o tav-o! 'your-Voc head-Voc')

is an exception. Alternatively, it may be considered as a simple summons, not a noun phrase.[37]

B. The argument cannot be coded in the verb because no appropriate marker is available. Firstly, verbal reflexive direct object markers simply do not exist, they are not just Ø. Therefore they must be determiners with *tav-*:

---

35 BOEDER 1969: 89, 103; 1989: 176.
36 BOEDER 1969: 104; 1989: 174.
37 See BOEDER 1985; 1989: 169.

(25) *ǯavr-it* *mo- v-i- ḳal- i* *(čem-i)* *tav-i*[38]
anxiety-Instr Prev-1S-SV-kill-TS (my-Nom) head-Nom
'I killed myself with anxiety'

(26) *tav-i* *še-m-a-q̇var-e-t* (from a toast)
head-Nom Prev-1O-NV-love(Caus)-TS-Pl
'you (polite plural) made me love you'

Secondly, since no intensified verbal person markers are available, the construction with *tav-* supplies an intensified variant; see:

(27) *čem-s* *tav-s* *v- Ø- u- ḳrep*
my-Dat head-Dat 1S-3IO-OV-pick
'(For whom are you picking the apple?) I am picking it for **me**!'

(where the *tav*-headed noun phrase has a 3rd person indirect object counterpart in the verb). Compare non-intensified:

(28) *v- i- ḳrep*
1S-SV-pick 'I am picking it for me'

C. The argument cannot be coded in the verb because of the morphosyntactic slot-filling constraint. This is the case we are primarily interested in. Consider:

(29) *a- m- i- zard- a* *šeni* *tav-i*
Prev-1O-OV-raise-Aor your-Nom head-Nom
's/he raised you for me'

38 Here and in (28), the "subjective version vowel" (SV) *-i-* is used in the verb. This is the traditional term for the morpheme *-i-* in pre-root position (no. 6 in table 1) which codes the "beneficiary" or similar roles of a reflexive indirect object. In (25), it indicates the reflexivity of the "beneficiary" possessor; literally: 'I kill (my) head to/for me'. Similarly, *-i-* indicates the beneficiary status of "for myself" in (28). Notice that the same *-i-* is used in the non-reflexive paradigm (table 2 (e)), traditionally called "objective version vowel" (OV), except in the 3rd person, where *-u-* can be considered an allomorph of *-i-* after the non-reflexive 3rd person indirect object marker *ø* (*h-, x-* in Old Georgian). In this sense, "subjective" and "objective version" (in (d) and (e) of table 2) are the same (Boeder 1969), although the formal contrast between 1st, 2nd and reflexive person (with *-i-*) on the one hand, and 3rd non-reflexive person (with *-u-*) on the other is, of course, not accidental.

The slot is filled with a 1st person indirect object marker *m-*, therefore the direct object 'you' cannot be marked in the verb and occurs as a possessive determiner of *tav-*. Similarly, in the causative construction:

(30) *g- a- ḳvl-evin-eb čem-s tav-s*
2O-NV-kill- Caus-TS my-Dat head-Dat
'I make you kill me'
(**v-g-a-ḳvl-evin-eb*)

the direct object 'me' must be a phrase with *tav-*, because the object marker position is filled with the indirect object marker *g-* 'you' for the causee. (For further examples see (15), (16), (18).) Note again that the purely morphological sequence constraint Ia) has no syntactic consequence at all; although *v-* is deleted, *g-a-ḳvl-evin-eb* behaves as if the complete form **v-g-a-ḳvl-evin-eb* were there (i.e. *čemi tavi* is not required as a subject noun phrase).

5. The square brackets of reflexive [Ø] in table 2 (d) deserve special attention. On the one hand, [Ø] behaves like Ø in (18), because no direct object can be marked in the verb: 's/he paints you for herself/himself' has to be rendered by:

(31) *[Ø]- i- xaṭ- av- s šen-s tav-s*
Refl-SV-paint-TS-3S your-Dat head-Dat

where the direct object 'you' is expressed by 'your head'. The form:

(32) *g- i- xaṭ- av- s*
2O-OV-paint-TS-3S

is perfectly correct, but it has a different meaning, namely: 's/he paints it for you'.[39]
However, although the verbs in (18) and (31) share the blocking behaviour based on rules II and III, they are different (a) morphologically and (b) syntactically.

(a) Contrary to overt object markers, reflexive [Ø] does not oust the subject marker *v-* according to rule Ia):

---

39 Forms like (31) must be distinguished from forms of lexically fixed "reflexive" verbs in which the object slot preceding a neutral version vowel *-i-* is empty and can be filled (TSCHENKÉLI 1958: 386; BOEDER 1969: 106), as in: *gamo-m-i-ɕer-s* Prev-1O-NV-write-3S(Sg) 's/he will summon me (by letter)'. As a consequence, *-m-i-* can be ambiguous: it is either an object marker specified by the "object version vowel" *-i-*('for me') or an object marker plus neutral version vowel *-i-* : *m-i-gd-eb-s* 1O-OV-throw-TS-3S(Sg) 's/he throws it for me' vs. *m-i-gd-eb-s xel-ši* 1O-NV-throw-TS-3S(Sg) hand-in 's/he achieves control of me' (see VOGT 1971: 122).

(33) *v-[Ø]-i-ċer*
1S-Refl-SV-write 'I write for me/myself'

(b) Morphological presence does not equal syntactic availability. Non-reflexive Ø in (e)-(g) of table 2 has all kinds of external adjuncts:

(34) *q̇vela-s / p̀rezident-s a-v-Ø-u-šen-e did-i saxl-i*
all-Dat / president-Dat Prev-1S-3IO-OV-build-TS big-Nom house-Nom
'I built a big house for everybody / for the president'

where 'all' or 'president' has to be linked to the indirect object marker Ø. Similarly in (35), 'children' is linked to *g-*, and 'adults' to *v-:*

(35) ***bavšv-eb-s*** *čai-s da-**g**-i-sxam-t,*
child-Pl-Dat tea-Dat Prev-**2O**-OV-pour-Pl,

upros-eb-i *ḳi ɣvino-s da-v-lev-t* (RḲ)
adult-Pl-Nom but wine-Dat Prev-**1S**-drink-Pl
'We will give you, the children, tea, but we, the adults, will drink wine.'

Reflexive zero is different: as far as I can see, there is no indication of its syntactic availability ("syntactic activity").[40] For instance, it cannot have appositions:

(36) **mq̇opad p̀rezident-s a-Ø-i-šen-e did-i saxl-i*
future president-Dat Prev-Refl-SV-build-TS(Aor) big-Nom house-Nom
'You built a big house for you(rself), the future president'

A dative form like *p̀rezident-s* cannot be an "apposition" of the reflexive indirect object marker [Ø] in the way that *Tamaz-s* is an apposition of *–m–* in (2). Instead, the apposition has to be linked to the subject:

(37) *mq̇opad p̀rezident-ma a-[Ø]-i-šen-e did-i saxl-i*
future president-Dat Prev-Refl-SV-build-TS(Aor) big-Nom house-Nom
'You, (as) the future president, built a big house for you(rself)'

Note that verb-external, *tav*-headed "reflexive" noun phrases do not preclude appositions:

---

40 See SADLER - SPENCER 1997: 224-226 for some discussion of the syntactic availability of morphological reflexives.

(38) *es* *ṗropesor-i* *tavis tav-s,* *q̇vela ḳoleg-is* *mṭer-s,*
this(Nom) professor-Nom his head-Dat, all colleague-Gen enemy-Dat,

ʒalian simṗatiur adamian-ad miičnevs (BR)
very nice person-Adv he.will.consider.it
'This professor will consider himself, the enemy of all colleagues, a very nice person.'

where the dative-marked apposition must belong to the dative-marked object, not to its coreferential nominative-marked subject.

How shall we describe this syntactic "inactivity" of verb-internal indirect object reflexives? One possibility seems to be rule ordering. First, the object "slot" is filled according to the slot-filling hierarchy (III), where the indirect object specified by the "subjective" version vowel *-i-*, the reflexive beneficiary, "wins the competition" and all other competitors have to be *tav*-headed noun phrases according to rule V. Then the reflexive is deleted. Finally, the morphological sequence constraint (Ia) simply will not apply, because there is no object marker that could oust the subject marker *v-*. However, this interpretation has some implications for Georgian grammar. Since the deletion of the reflexive makes it unavailable for syntax, it has to be considered as a valency-reducing morpholexical operation like e.g. the formation of middle verbs in English.[41] And since the deletion follows the distribution of pronouns among verbal slots on the one hand, and clause level noun phrases on the other, this distribution should also be a lexically determined matter - if morpholexical representations are the input of syntax. In other words, at least part of Georgian verb structure[42] would seem to be morpholexical, and not the result of syntactic agreement rules (copying features from independent noun phrases into verbs). In this sense, the purely morphological terms "slot" and "[ø]" turn out to be inappropriate: they have to be reinterpreted on the level of morpholexical structure. Reflexive [ø] is not a zero morpheme

41 See SADLER - SPENCER 1997: 221-223. - Nicholas Evans rightly points out to me that [ø], since it indicates reflexivity and thereby removes an argument from the need for encoding on the verb, is no longer an argument with exceptional behaviour. In this it is comparable to valency-reducing morphological devices of other languages: "In Bininj Gun-wok reflexives are marked by a suffix, whereas arguments are marked by prefixes; the suffix reduces the valency by one, giving reflexive or reciprocal readings, but is certainly not an argument itself." It is true that "[ø]" together with the "subjective version vowel" *-i-* (table 2 (d)) is a valency-reducing device, a view which is supported by its occurrence in one type of passive formation (*i-çereba* 'it is (being) written'; for this connection, see DAMENIA 1982). However, on the morphological level, [ø] itself is simply nothing; under my interpretation, the forms with [ø] contain no more of a zero morph than, say, middle verbs in English. [ø] is a notational device to represent the fact that for the purpose of "competition", its behaviour parallels the behaviour of overt nominal expressions.

42 Part of it, not all: number agreement is better described on the basis of syntax. But alignment of arguments would also seem to be morpholexical, as far as assignment of morphological subject and indirect object status is concerned (with indirect verbs and perfect verb forms).

(like the 3rd person indirect object markers in table 2 (c), (e)-(f)). It is not a morphological unit, but stands for the presence of an argument at the level where arguments enter into the competetion described above.

As we said above, *tav*-headed noun phrases do not code reflexivity by themselves, which explains their occurrence in subject position (see (22)). But could we say that the possessive pronouns of *tav*-headed noun phrases (which bear its referential index) can be reflexive? Consider:

(39) *ǫvav-sa-c tavis-i baxala mo-s-ṣon-s* (proverb)
crow$_i$-Dat-too its$_i$-Nom young(Nom) Prev-3IO$_i$-please-3SgS
'Even the crow$_i$ likes its$_i$ own young.'

where the dative subject is co-referential with the possessive pronoun *tavis-*. But although this pronoun is called "reflexive" in the grammatical tradition, it is not necessarily controlled by the subject:

(40) (a) *ma-s tavis-i barat-i Svanet-ši gavugzavne* (GT)
she$_i$-Dat her$_i$-Nom letter-Nom Svanetia-in I.have.sent.it.to.her$_i$
'I have sent her$_i$ her$_i$ letter to Svanetia'

(b) *ʒma gavisṭumre tavis-i col-švil-it*
brother$_i$ (Nom) I.saw.him$_i$.out.as.a.guest his$_i$-Instr wife-child-Instr
'I saw my brother$_i$ out with his$_i$ wife and his$_i$ child'

(c) *Nodar$_i$-i šeaḳrto tavis$_i$-ma xma-m*
Nodar$_i$-Nom it.frightened.him$_i$ his$_i$-Erg voice-Erg
'His$_i$ own voice frightened Nodar$_i$'[43]

In other words: while *tavis-* codes intra-clausal anaphora, it is not specifically controlled by the subject of the clause (in whatever sense). Although we have a clear case of subject-non-subject asymmetry on the morpholexical level (where a reflexive indirect objects is specifically coded in the verb), and although we need subjects to account for control properties (coreference across clause boundaries) and the like, pronouns are not criterial of subjecthood on the clause-level.

6. The main point made in this paper is the specificity of the Georgian type of "agreement" between verb-internal and verb-external entities. If we conceive of agreement only as a kind of feature copying from verb-external entities into the verb, one distinctive property of Georgian verb agreement is not accounted for: the form and occurrence of verb-external

43 (b) and (c) are from an as yet unidentified Georgian source.

entities partially depends on the properties of the verb. This recalls the old concept of the finite verb form as a clause *in nuce* which can be expanded by adding "appositional" subject and object noun phrases. In this sense, *Tamaz-s* in (2) is an "apposition" of the verbal marker *–m–* 'I'.[44] On an analytical level, this idea turns up as "catalysis" in LOUIS HJELMSLEV's theory (1961: 93-96), where e.g. the expanded forms in (19a'), (19b') and (19c') are the result of an "interpolation" of external entities that must be posited to extract the internal categories of the verb (VOGT 1971: 88) – a procedure that every Georgian child has to master in school. In the framework of generative grammar, on the other hand, the non-expanded forms are interpreted as "stripped" of underlying argument co-constituents (BRAITHWAITE 1973: 10 et passim) whose non-overtness has come to be known as "pro-drop".

A more sophisticated variant of the "appositional" approach was proposed in the configurationality debate (see e.g. JELINEK 1989). According to this theory, the verbal complex exhibits the fully-fledged array of its arguments, and noun phrases are linked with verbal arguments if they are compatible with them. Indeed, compatibility is an intuitively satisfying concept to characterize the agreement relations described in this paper. In particular, the assumption of the linking of compatible entities is a straightforward way to describe such things as 3rd person "appositions" of 1st person subjects or objects (as for instance in (35)) without resorting to pro-drop for non-overt extraverbal arguments. I am still in favour of this approach, and yet a better understanding of "agreement" will depend both on an account of complex data like those presented in this paper and on a theory that characterizes the relationship between morphology and syntax in a principled way. Such a theory has to clarify the question of whether morphological operations necessarily and exclusively depend on the output of syntax (ANDERSON 1984; see HARRIS 1995: 1382), or whether other interdependencies are conceivable. If the interpretation of reflexives proposed above is correct, it means that more operations of Georgian grammar are morpholexical than standard descriptions suggest.

ABBREVIATIONS

Adv	= adverbial case	O	= object
Aor	= aorist indicative	Opt	= optative, subjunctive II
Caus	= causative	OV	= objective version
Cond	= Conditional	Part	= participle formant
Dat	= dative	Perf	= perfect
DO	= direct object	Pl	= plural
Erg	= ergative	Poss	= possessive
EV	= e-version	Pres	= present

44 For more examples see BOEDER 1989: 178

Ex	= extension marker	Prev	= preverb
Frequ	= frequentative (Georgian: "xolmeobiti", Deeters: "Permansiv")	Quot	= quotative
		Refl	= reflexive
		S	= subject
Fut	= future	Sg	= singular
Instr	= instrumental	SupV	= superessive version
IO	= indirect object	SV	= subjective (reflexive) version
Gen	= genitive	TS	= thematic suffix ("paradigm marker")
Neg	= negation		
Nom	= nominative	Voc	= vocative
NV	= non-specifying (neutral) version		

Other abbreviations are initials of my informants

References

ALLEN, W. SIDNEY. 1956. Structure and system in the Abaza verbal complex. *Transactions of the Philological Society* 1956: 127-176.

AMIRIǯE, NINO (NINO AMIRIDZE). 1998. *Conceptual Archetypes and their Reflection* in *Morphosyntax*. Dissertation Bulletin. Submittence (sic) for a candidate's degree (Ivane Javakhishvili Tbilisi State University). Tbilisi.

ANDERSON, STEPHEN R. 1984. On representations in morphology: case, agreement and inversion in Georgian. *Natural Language and Linguistic Theory* 2: 157-218.

ANDERSON, STEPHEN R. 1992. *A-Morphous Morphology* (= Cambridge Studies in Linguistics 62). Cambridge: Cambridge University Press.

ARONSON, HOWARD I. 1991. Modern Georgian. In: A. C. HARRIS (ed.) 1991: 219-312.

ASATIANI, RUSUDANI. 1982. *Marṭivi ċinadadebis ṭipologiuri analizi* (tanamedrove saliṭeraṭuro enis masalaze) (Sakartvelos SSR Mecnierebata Aḳademia, Aɣmosavletmcodneobis Insṭituṭi). Tbilisi: Mecniereba.

ASATIANI, RUSUDANI. 1994. *Kartvelur enata ṭipologiis saḳitxebi* (gramatiḳul ḳaṭegoriata sisṭemaši dominanṭuri ḳaṭegoriis dadgenis tvalsazrisit) (Sakartvelos SSR Mecnierebata Aḳademia, Aɣmosavletmcodneobis Insṭituṭi). Tbilisi: Mecniereba.

BOEDER, WINFRIED. 1969. Über die Versionen des georgischen Verbs. *Folia Linguistica* 2 (1968): 82-152.

BOEDER, WINFRIED. 1979. Ergative syntax in language change: the South Caucasian languages. In: FRANS PLANK (ed.): *Ergativity.* Towards a theory of grammatical relations. London...: Academic Press, 435-480.

BOEDER, WINFRIED. 1985. Zur Grammatik des Vokativs in den Kartwelsprachen. In: URSULA PIEPER — GERHARD STICKEL (edd.): *Studia Linguistica Diachronica et Synchronica.* Werner Winter Sexagenario Anno MCMLXXXIII. Berlin — New York – Amsterdam: Mouton de Gruyter, 55-80.

BOEDER, WINFRIED. 1989. Verbal person marking, noun phrase and word order in Georgian. In: LÁSZLÓ MARÁCZ - PIETER MUYSKEN (eds.) 1989: 159-184.

BOEDER, WINFRIED. 1999. A slot filling constraint in the Georgian verb and its syntactic corollary. *STUF 52*: 241-254

BRAITHWAITE, KIM. 1973. *Case Shift and Verb Concord in Georgian.* Ph.D. dissertation. University of Texas at Austin.

DAMENIA, MERI. 1982. *Kartuli zmnuri morpemebis strukturuli modelebi* [Russian résumé: Strukturnye modeli glagol'nyx morfem v gruzinskom, pp. 263-271]. Tbilisi: Mecniereba.

DEETERS, GERHARD. 1931. *Das kharthwelische Verbum.* Vergleichende Darstellung des Verbalbaus der südkaukasischen Sprachen (= Sächsische Forschungsinstitute Leipzig. Forschungsinstitut für Indogermanistik, Sprachwissenschaftliche Abteilung. Band 1). Leipzig: Markert & Petters.

EVANS, NICHOLAS. 1999. Why argument affixes in polysynthetic languages are not pronouns: evidence from Bininj Gun-wok. *STUF* 52: 255-281.

FÄHNRICH, HEINZ. 1991. Old Georgian. In: A.C. HARRIS (ed.) 1991: 129-217.

GAMQRELIZE, TAMAZ. 1979. Zmnis 'pirianoba' da 'valentoba' / Valentnost' i personal'nost'glagola [Résumé p. 51]. In: *Saenatmecniero krebuli ezyvneba Givi Mačavarianis xsovnas dabadebis 50 clistavze / Studies in Linguistics Presented to the Memory of late prof. Givi Machavariani on the Occasion of His Fiftieth Birthday.* (Sakartvelos SSR Mecnierebata Akademia / Academy of Sciences of the GSSR). Tbilisi: Macniereba, 33-51.

GAMQRELIZE, TAMAZ. 1981. 'Personenhaltigkeit' und 'Valenz' des Verbs. *Georgica* 4: 65-70.

HARRIS, ALICE C. 1981. *Georgian Syntax.* A study in relational grammar (= Cambridge Studies in Linguistics 33). Cambridge: Cambridge University Press.

HARRIS, ALICE C. 1991. *The Indigenous Languages of the Caucasus.* Volume 1: *The Kartvelian Languages.* Edited by Alice C. Harris. Delmar, New York: Caravan.

HARRIS, ALICE C. 1995. Georgian. In: J. JACOBS et al. (eds.): *Syntax.* Ein internationales Handbuch zeitgenössischer Forschung. Berlin - New York: Mouton, 1377-1397.

HASPELMATH, MARTIN. 1991. Zur Grammatikalisierung von Indefinitpronomina. In: NORBERT BORETZKY et al. (eds.): *Sprachwandel und seine Prinzipien.* Bochum: Universitätsverlag Brockmeyer, 103-125.

HEWITT, BRIAN GEORGE. 1995. *Georgian.* A structural reference grammar (= London Oriental and African Library 2). Amsterdam - Philadelphia: Benjamins.

HJELMSLEV, LOUIS. 1961. *Prolegomena to a Theory of Language.* Translated by Francis J. Whitfield. Madison: University of Wisconsin Press

[Imedaze, Natela] NATELA IMEDADZE - KEVIN TUITE. 1992. The acquisition of Georgian. In: DAN ISAAC SLOBIN (ed.): *The Crosslinguistic Study of Language Acquisition.* Volume 3. Hillsdale/N.J. - London: Erlbaum, 39-109.

JELINEK, ELOISE. 1989. The case split and pronominal arguments in Choctaw. In: L. MARÁCZ – P. MUYSKEN (eds.) 1989: 117-141.

MARÁCZ, LÁSZLÓ - PIETER MUYSKEN (eds.). 1989. *Configurationality.* The typology of assymmetries (= Studies in Generative Gramar 34). Dordrecht/Holland - Providence/RI: Foris.

SADLER, LOUISA - ANDREW SPENCER. 1997. Morphology and argument structure. In: ANDREW SPENCER - ARNOLD M. ZWICKY (eds.): *The Handbook of Morphology* (= Blackwell Handbooks in Linguistics). Oxford: Blackwell, 206-236.

ŠALVAŠVILI, LEILA. 1985. *Čemit* rigis pormebi piris nacvalsaxelta mokmedebitis punkciit. *Tbilisis Universitetis Šromebi* 262 (Enatmecniereba 9): 188-202.

ŠALVAŠVILI, LEILA. 1989. *mit - misit* rigis pormebi Kartulši. *Kartuli Sitqvis Kulturis Sakitxebi* 9: 240-254.

ŠALVAŠVILI, LEILA. 1992. Pirveli da meore piris nacvalsaxelta moulodnel kombinaciebši gamoqenebis šemtxvevebi 3vel Kartulši. *Iberiul-Kavkasiuri Enatmecnierebis Celicdeuli* / Annual of Ibero-Caucasian Linguistics 18-19 (1991-1992): 53-64.

SANIJE, AKAKI. 1973. Kartuli enis gramatikis sapuzvlebi I: Morpologia. *Meore gamocema (= Txzulebani III; = 3veli Kartuli enis katedris šromebi 15*). Tbilisi: Tbilisis universitetis gamomcemloba.

TANDAŠVILI, MANANA. 1996. Kartuli zmnis egocentruli bunebisatvis I. Pirta strukturisatvis Kartul zmnaši. *Saenatmecniero 3iebani* 5: 60-67.

TANDAŠVILI, MANANA. 1998. Kartuli zmnis egocentruli bunebisatvis II. Zmnur tipta depiniciisatvis Kartulši / On the egocentric nature of the Georgian verb II. On the definition of verb types in Georgian [Résumé p. 67]. *Saenatmecniero 3iebani* 7: 62-67.

TSCHENKÉLI, KITA. 1958. *Einführung in die georgische Sprache. I: Theoretischer Teil.* Zürich: Amirani.

VAMLING, KARINA. 1988. A feature-based analysis of Georgian morphology. Case study: object agreement. In: FRIDRIK THORDARSON (ed.). *Studia Caucasologica I.* Proceedings of the Third Caucasian Colloquium, Oslo, July 1986 (The Institute for Comparative Research in Human Culture Oslo / Instituttet for sammenlignende kulturforskning Serie B: Skrifter LXXV). Oslo: Norwegian University Press, 309-319.

VOGT, HANS. 1971. *Grammaire de la langue géorgienne.* (= Instituttet for sammenlignende kulturforskning. Serie B: Skrifter LVII). Oslo: Universitetsforlaget.

WACKERNAGEL, JACOB. 1926. *Vorlesungen über Syntax mit besonderer Berücksichtigung von Griechisch, Lateinisch und Deutsch.* Erste Reihe. Basel: Birkhäuser.

Compound nouns vs. incorporation in Classical Nahuatl

MICHEL LAUNEY

1. On sentence-words

As it was first used (and still is nowadays in an informal and intuitive way), the term *polysynthetic* is related to the idea of "sentence-words", but this splits into two further notions. The most common one is that notions which in another language would be expressed by independent words or phrases often merge to form long compound words: this is clearly expressed by SAPIR (1921:128):

> The elaboration of the word is extreme. Concepts which we should never dream of treating in a subordinate fashion are symbolized by derivational affixes or "symbolic" changes in the radical element, while the more abstract notions, including the syntactic relations, may also be conveyed by the word.

CRYSTAL (1987:293) is a more recent expression of this:

> Words are often very big and complex, containing a mixture of agglutinating and inflectional features, as in Eskimo, Mohawk, and Australian languages. For instance, the aborigine language Tiwi expresses "I kept on eating" as *ngi-rru-unthing-apu-kani* ('I - past tense - for some time - eat - repeatedly').[1]

In this sense, incorporation is often cited as the epitome of polysynthesis, as in SPENCER (1991:38):

> The polysynthetic languages are those which, like Chukchee, permit processes such as noun incorporation, so that a single word can encode a meaning which would require a fairly elaborate sentence in many other languages

or DUBOIS (1994:370), who apparently uses the terms in a synonymous way:

> *polysynthétique* ou *incorporante*: se dit d'une langue dans laquelle les diverses parties de la phrase se soudent en formant de longs mots composés (cas de l'esquimau et de nombreuses langues indiennes).

---

1 See Osborne (1974:42) for the original data.

In a second sense, sentence-words are characteristic of languages of the head-marking type (NICHOLS 1986); in LAUNEY (1986, 1994) I suggest calling them *langues indiciantes* ("indexing languages"), where argument positions appear as affixes on verbs or noun heads. In this sense, sentence-words need not be long, provided they contain the required number of argument affixes. There has been a lot of discussion about this type – see FOX (1966) on Mayan languages (though he does not use the term *polysynthetic*), and, in a more humorous way, QUENEAU (1965) who points out the use of French clitics as very parallel to Chinook morphosyntax.

In his latest book, BAKER (1996: 17) combines the two meanings in a more theoretically restricted definition:

> First, there is the Mohawk type, in which both agreement morphemes and lexical roots count as rendering an argument visible. Languages with this property I call *polysynthetic*, thereby giving a technical sense to a word that BOAS (1911) and SAPIR (1921) use rather impressionistically in their typologies (....) However, it must be emphasized that many languages that Sapir and Boas would have called polysynthetic are not polysynthetic in my narrower and more technical sense. Some languages with quite impressive amounts of morphological complexity may not use that complexity to *systematically* represent *argument* relationships....

In Classical Nahuatl (or Aztec), very long compound words are fairly uncommon, but this language does have incorporation, as we will see later. And it is a strictly head-marking language, without case (i.e. not dependent-marking), so it fits Baker's definition pretty well. In this paper, however, I focus on another major feature of Nahuatl morphosyntax, which I will call *omnipredicativity*; this reinforces the head-marking parameter and leads to interesting conclusions about the status of NPs, and about the parallelism between incorporation and compound nouns. Before going into the latter issue, I will explain briefly what I mean by omnipredicativity (for a more detailed account, see LAUNEY 1994).

2. Omnipredicativity

At first sight, Nahuatl is a canonical head-marking language, with personal prefixes for subject, object and noun possessor:

(Subject) Singular 1 *n(i)-*, 2 *t(i)*, 3 *ø-* (zero); Plural 1 *t(i)-*, 2 *am-*, 3 *ø-*
(Object) Singular 1 *-ne:ch-*, 2 *-mitz-*, 3 *-k(i)-* ; Plural 1 *-te:ch-*, 2 *-ame:ch-*, 3 *-kim-*
(Possessor) Singular 1 *-n(o)-*, 2 *-m(o)-*, 3 *-i:-* ; Plural 1 *-t(o)-*, 2 *-am(o)-*, 3 *-i:m-*[2]

---

2 In this transcription, *ch, tl* and *tz* stand for affricates which are pronounced, respectively, [tš] (as in Spanish), [tl] and [ts]; *kw* is a labiovelar; ' stands for a glottal stop; some of the morphemes have allomorphs with or without a vowel according to the syllabic context (this is shown by vowels between brackets).

In the third person, argument marking is overt for object (1b) and possessor (2b); for the sake of consistency we posit a zero marker for subjects, as in (3b):

(1a)*Ni-mitz-itta* 'I see you'; *ni-k-itta* 'I see him/her'; *ni-kim-itta* 'I see them'
(1b)*Ni-k-itta in kone:-tl* 'I see the baby'; *ni-kim-itta in ko:-kone-'* [3] 'I see the babies'
(2a)*no-tlakwal* 'my food'; *i:-tlakwal* 'his/her food'; *i:n-tlakwal* 'their food' [4]
(2b)*i:-tlakwal in kone:-tl* 'the baby's food'; *i:n-tlakwal in ko:-kone-'* 'the babies' food'
(3a)*Ni-tza'tzi* 'I shout'; *tza'tzi* 'he/she shouts' (i.e. *ø-tza'tzi*)
(3b)*Tza'tzi* (i.e. *ø-tza'tzi*) *in kone:-tl* 'The baby shouts'

Now, let us delve into the nature of the phrase *in kone:tl*. It looks like a commonplace NP with a determiner and a noun, and the relationship with the argument marker in the head can be expressed by coindexation:

(4a)*Ni-k$_i$-itta [in kone:tl]$_i$* (= 1b)
(4b)*i:$_i$-tlakwal [in kone:tl]$_i$* (= 2b)
(4c)*ø$_i$-tza'tzi [in kone:tl]$_i$* (= 3b)

But the reality is different. We must take into account several features which combine to form the omnipredicative parameter. First, the subject pronominal prefixes shown combining with verbs in (3) are used with nominal predicates as well:

(5a)*Ni-kone:tl, ti-kone:tl, ø-kone:tl* 'I am a baby, you are a baby, he/she is a baby'
(5b)*Ti-no-kone:-w* [5] etc. 'You are my baby'

Second, word order can be reversed, producing a focus effect on the noun:

(6a)*Kone:tl in ni-k-itta* 'I see *a baby*', 'It is a baby whom I see' (cf. 1b)
(6b)*Kone:tl in i:tlakwal* 'It is *baby's* food' (cf. 2b)
(6c)*Kone:tl in tza'tzi* 'It is a baby who is shouting'

Third, there may appear NPs with first or second person markers:

(7a)*Ni-mitz-itta in ti-kone:tl* 'I see you, (who are a) baby'
(7b)*Ti-tza'tzi in ti-kone:tl* 'You shout, (you who are a) baby'

---

3 *-tl* (allomorphs *-tli* or *-li*) is a so-called absolute suffix which appears in the singular of non-possessed nouns; in the plural, there is another suffix (-' or others) sometimes combined with first syllable reduplication.

4 Or rather 'it is my food', etc., see below.

5 *w* (allomorph: zero) replaces the absolutive suffix in possessed nouns; its plural is *-wa:n*.

Fourth, a vocative form (which can appear with nouns only) must be explicitly marked (by a stressed suffix *-é*):

(8a) *Kone:tl-é!* 'O baby!'
(8b) **Kone:tl!*

These data show that nouns have the same predicative properties as verbs, but that goes much deeper than a mere lack of copula, as in, say, Russian or Arabic; such languages do not allow argument affixes in noun predicates[6], while Nahuatl does, so that (within the polysynthetic frame) noun predicates are sentence-words just like verb predicates are. It would be intuitively fruitful to express the meaning of verbs and the meaning of nouns just in the same way: if we say that *tza'tzi* means '(to) shout', then we must say that *kone:tl* means, not 'baby', but rather '(to) be a baby'. Note further that (2a-b) may have a sentence status: 'it is my/your/his/her/the baby's food'.

Not only do nouns behave like verbs when used as sentence predicates, but argument phrases may be built with verbs just in the same way as they are with nouns. The structure **in V** which appears in (6a) and (6c) may look like a syntactic monster, since it is not a noun phrase (its lexical part is indeed a verb, which unlike nouns may be inflected by tense[7]); but it is not a verb phrase either, for at least two reasons: it is specified by the same determiner as a noun, and it occupies an argument position. This point is crucial: in (6a-c) *kone:tl* is in a predicate position, in every sense of this term. It is the informative part of the sentence (call it *comment*, *focus*, or, as in the European tradition, *rheme*): I say 'It is a baby' about the person whom I see (6a), the person whose food it is (6b), or the person who shouts (6c). We will see that it is also the keystone of syntactic relations. And the phrase that follows is in subject position, referring to an entity which specifies the semantic value of the subject of this predicate: 'the person such that I see him/her (is a baby)', 'the person such that this is his/her food (is a baby)', 'the person such that he/she shouts (is a baby)'. These are undoubtedly argument phrases, but they cannot be expressed in categorial terms such as NP or VP; if need be, we could follow ABNEY (1987), and say that in both cases the syntactic head is actually the determiner, so that they are DetPhrases.

Let us now see how our semantic analysis is expressed syntactically. No overt person marker appears in (6c) but we do see an object marker in (6a) and a possessive marker in (6b). Moreover, we see in (7a-b) that overt 1st or 2nd person subject marker may appear in the argument phrase: this is admittedly much less common than in 3rd person, since the first and second persons are already fully specified referentially, but it may be the case that we need to bring some semantic specification or restriction about them ('As an N, I/you...', 'Being an N, I/you...', etc.). In all such cases, there is a coindexation between a prefix in the argument phrase and a prefix in the sentence predicate. As a consequence, the most consi-

---

6 For further arguments, see LAUNEY (1994).

7 E.g. *Kone:tl in ni-k-itta-s / ni-k-itta-ya / ni-k-itta-k* 'It is a baby whom I will see / was seeing / saw'.

stent way to analyse (4a-c) and their reverse counterparts in (6a-c) is to postulate a zero third person subject prefix on the nouns, as follows:

(9a) *ni-k$_i$-itta in ø$_i$-kone:tl* (= 1b, 4a)
(9b) *i:$_i$-tlakwal in ø$_i$-kone:tl* (= 2b, 4b)
(9c) *ø$_i$-tza'tzi in ø$_i$-kone:tl* (= 3b, 4c)
(10a) *ø$_i$-kone:tl in ni-k$_i$-itta* (= 6a)
(10b) *ø$_i$-kone:tl in i:$_i$-tlakwal* (= 6b)
(10c) *ø$_i$-kone:tl in ø$_i$-tza'tzi* (= 6c)

In other words: all argument phrases are subordinate predicates, which designate an entity to which this predicate can be applied. For instance, in (9c), *in ø-kone:tl* is 'the entity such that (**in**) it is a baby', and in (10c) *in ø-tza'tzi* 'the entity such that he/she shouts'. Put simply, you can refer to some entity as *the baby* if and only if this entity is such that you can say about it/him/her *It is a baby*; if not, you are lying, or you are mistaken, or you are making an improper use of the word. Incidentally, this explains why vocatives have to be overtly marked: they are non-predicative uses of nouns (when you call someone, you do not say anything about anybody or anything), and in Nahuatl every lexical item in the sentence is either the sentence predicate or a subordinate predicate: a word like *kone:tl* must always be interpreted as *ø-kone:tl*, 'it is a baby', or '(entity such that) it is a baby'. There are no words specialized just for predicative or for referential use: all words are basically predicative, and all of them may refer to an entity; but reference is a secondary use, derived from predication, which is primitive. Hence the suggested term *omnipredicativity* to refer to such a language.

The condition for this kind of subordination is that there exist a coindexation (representing coreference) between one argument place in the subordinate predicate and one argument place in the main sentence predicate. If I say of an entity, such that (it is already admitted that) *it is a baby* that *it shouts*, or of an entity such that (it is already admitted that) *it shouts* that *it is a baby*, it must be the same entity: otherwise the statement would be meaningless. Incidentally, there is no focalization in such a language: there is no reason to derive (10c) from (9c) or the contrary. Rather, both (9c) and (10c) are built by compounding two "simple" predicates. One of them, which is taken as the informative centre of the sentence (*rheme, focus, comment...*) gives the sentence predicate and can form a sentence by itself; the other represents "old" information and appears as subordinate when the semantic value of the argument needs to be specified.

These phenomena may lead us to revisit a common issue in typology: the *subject-prominent* vs. *topic-prominent* opposition. Surprisingly enough, little attention has been paid to the opposite point of view, namely the other member of the pairs, respectively *verb* and *comment* (*rheme, focus...*). Let us assume that there are *verb-prominent* and *rheme-prominent* (or *focus-prominent*) languages: what could the syntactic characteristics of rheme prominence be? Nahuatl offers a prototypical case of precisely this: every lexical item repre-

sents a predicative notion, which can be applied to one or two[8] arguments. These items may have specific properties, such as aspectual variability (verbs) or aspectual stability (nouns), but all of them can form a sentence, provided that they are *saturated*, i.e. that a semantic value is assigned to each argument place. That is what argument affixes stand for, and they are sufficient if for whatever reason (situational evidence or anaphora) the semantic value of the argument is supposed to be clear.

It goes without saying that I totally agree with JELINEK's (1984) suggestion that in such a language there is no *pro-drop* parameter, but rather a *NP-add* one (I hope that I am not misinterpreting Jelinek's idea by putting it that way): sentence-words are perfectly well-formed sentences and all argument phrases are adjuncts.

However, this does not necessarily mean that argument affixes *are* the arguments, because there *are* no arguments, but two different things. One is *argument positions*, whose number and semantic properties are determined by each predicate. The second is *argument denomination*, which takes the form of a phrase of some kind. In order to make a valid statement, a semantic value must be assigned to argument places, but if for whatever reason this semantic value is supposed to be known, no overt denomination is required: this is why argument affixes are always present, while argument phrases occur in some cases only, and if they do, argument affixes are not deleted.

Moreover, there is no other possible interpretation in an omnipredicative language, since the noun itself is not basically a designating item which fills a sort of argument slot, but a predicate with its own argument structure. In (9c) for instance, the subject place of *ø-tza'tzi* is not linked with the noun *kone:tl* (or the NP *in kone:tl*), but *with the subject position* of the (subordinate) predicate *ø-kone:tl*. And in (10a), the subject position of the sentence predicate *ø-kone:tl* (marked by ø- '3sg') is linked with the object position of the subordinate predicate *ni-k-itta* (marked by k- '3 sg O'). In sum: nouns are basically predicates, even though their referential use may be more frequent than is the case for verbs. This reflects a cross-linguistic norm of predication which can be formulated as "apply unstable properties to stable entities rather than the reverse".

3. Compound words in an omnipredicative language

Since there are such striking morphosyntactic similarities between verbs and nouns[9], we may try to see how far they go and search for other similarities. Let us recapitulate our first

---

8 Three in the case of ditransitive verbs, which I will not examine here.

9 Actually, there are lots of differences, but these are not relevant here. For instance: verbs are inflected for tense-aspect while nouns are not; reflexive, ditransitive and impersonal verbs have no direct parallel in nouns nor has the opposition between unaccusative and unergative intransitive verbs (see LAUNEY 1994).

intuitions about polysynthesis more precisely and see how word compounding works in Nahuatl. Compounding is not as widely developed into very long words as in, say, Eskimo, but it is very common nonetheless. Among the 4 possible combinations (N+N, N+V, V+V and V+N) the latter is totally unattested (the verb stem must first undergo a nominalization of some sort), and in V+V structures, the second V is an "auxiliary" verb, which I only mention here because such verbs do not lie within the scope of this article[10]. But we can get some interesting hints if we compare compound nouns (N+N) and incorporation (N+V); note that in both cases the first N refers to noun stems: the absolute suffix which systematically appears in the autonomous use of nouns is dropped here, since it can never occur before any other morpheme, except the vocative. These forms of compounding were broadly examined in LAUNEY (1986), but much to my shame it did not occur to me then to compare them within the omnipredicative hypothesis. Moreover, while still agreeing with the typology of incorporation I gave there, I now realize that my classification of N+N compounding was a bit clumsy. Interestingly enough Marianne Mithun, in her paper on incorporation (MITHUN 1984), which I had not read then, seems to give up building a formal classification (the example she gives is: in *alligator shoes* the semantic relationship may be 'made for', 'shaped like', 'made of', 'meant to go with...').

3.1 Incorporation: typology

With regard to incorporation, the classification I suggested distinguishes saturating incorporation, which affects the valency of the verb, and a number of subtypes of modifying incorporation, which leave valency unaffected. We now discuss these types in more detail.

3.1.1 Saturating Incorporation ('Incorporation saturante')

This corresponds to Type I (Lexical compounding) in MITHUN (1984). The incorporated noun stem is in an object relationship to the verb stem, so the valency is reduced (a transitive V forms an intransitive N+V). Examples are:

(11a)	*ni-tana'-chi:wa*	*ni-tlakwalis-ka:wa*
	1S-basket-make	1S-food-leave
	'I make baskets'	'I fast'
	n-ista-namaka	*ni-naka-kwa*
	1S-salt-sell	1S-flesh-eat
	'I sell salt'	'I eat meat'

10 Moreover, the first V is always followed by a synchronically meaningless suffix -ti- which may be a former nominalizing morpheme. So it may be the case that after all verbs can never occur as first element of a compound word.

(11b)	*ni-tla:ka-chi:wa*	*ni-siwa:-tlani*
	1S-human-make	1S-woman-ask_for
	'I give birth'	'I am grieving for a woman'

There are also apparent cases of valency reduction from intransitive to impersonal. Impersonal verbs are verbs which only appear in 3rd person singular and cannot occur with an argument phrase,[11] so derived impersonal verbs formed as N-V compounds (13) behave just like primary impersonal ones (12a-b) (though see below for an alternative interpretation).

(12a)	*ø-kiyawi* 'it rains'	
(12b)	**ø-kiyawi in N/V*	
(13)	*(ø-)tla:l-olin*	*(ø-)se:-wetzi*
	(3S)-earth-move	(3S)-frost-fall
	'There is an earthquake. '	'It freezes. '

3.1.2 Modifying Incorporation ('Incorporation modifiante')

This corresponds to Type II (Manipulation of Case) in Mithun's typology, and was already recognized by BENVENISTE (1966). The verb valency, which may be intransitive as well as transitive, remains unaltered. There are four subtypes:

3.1.2.1 Adverbial modifiying incorporation

Here the incorporated noun stem is in an adverbial relationship with the verb, typically instrumental, but also spatial, temporal, comitative etc. This occurs with both transitive and intransitive verbs. (At this point only examples with the noun incorporated will be given. Corresponding examples with unincorporated nouns will be given in § 4.1).

(14a)	*ni-k-te-wi:teki*	*ni-k-kwaw-miktia*
	1S-3O-stone-hit	1S-3O-wood-kill
	'I hit him with a stone'	'I kill him with a stick'
(14b)	*ni-k-ma:-ka:wa*	*ni-k-tlan-koto:na*
	1S-3O-hand-leave	1S-3O-tooth-tear
	'I let it fall'	'I tear it with my teeth'

11 In LAUNEY (1986, 1994) I show that the impersonal paradox (i.e. that the "null-subject" is also a 3rd person definite subject in these verbs) comes from the fact that these verbs are a special case of intransitive verbs where the class of possible subjects is reduced to one member ('it rains' is something like 'the rain rains' or 'the nature rains' or 'what rains rains'). Hence both the definiteness and the impossibility of an argument phrase, since there is no referential ambiguity. This is why I postulate a zero prefix occurs in these examples.

(14c) *ni-k-to:p-tema* *ni-k-tla:l-akia*
1S-3O-chest-cram 1S-3O-earth-put_in
'I fill my chest with it' 'I put it into the earth'

(15a) *ø-to:nal-wa:ki* *ni-to:nal-kochi*
3S-sun-become dry 1S-sun-sleep
'It dries up in the sun' 'I sleep during the day'

(15b) *n-a:-miki* *ni-siwa:-kochi*
1S-water-die 1S-woman-sleep
'I am thirsty' 'I sleep with a woman'

(15c) *ø-sen-ya:w*
3S-one-go
'He is leaving forever'

Note that if what is said in note 11 about impersonal verbs is correct, then (13) is no more than a subcase of (15): 'there is moving by the earth, there is falling by the frost', or 'it (i.e. the nature, the situation...) moves by the earth...'

3.1.2.2 Comparison and secondary predication as modifying incorporation

Here the incorporated N expresses a point of comparison ('like an N') or a secondary predication ('as an N', 'being N') for the subject or the object:

(16a) *ø-tla:ka-tla'toa* *ni-koyo:-ne'nemi*
3S-human-speak 1S-coyote-walk
'(the parrot) speaks like a man' 'I walk like a coyote'
(16b) *ø-we:we'-miki-'* *ø-ki-pitzo:-neki*
3S-old-die-PL 3S-3O-pig-want
'they die old' 'he desires (gold) like a swine'

(16c) *ø-ki-tetzaw-mati* *ni-k-kwal-itta*
3S-3O-omen-feel 1S-3O-good-see
'he considers it an omen' 'I approve it'

3.1.2.3 Possessor Raising

The incorporated N refers to a possessed object or a possessed subject, and the possessor is raised to object (or, in the second case, subject) position. Constructions where it is the possessed subject that is raised are much less frequent.

(17a) *ni-k-kech-koto:na* *ni-mitz-tla'to:l-kaki*
1S-3O-throat-cut 1S-2O-words-hear
'I cut his throat' 'I understand you'

(17b) *ni-k-yo:l-la:lia* *ni-no-ma:-po'po:wa*
1S-3O-heart-put 1S-REFL-hand-wash
'I comfort him' 'I wash my hands'

(18a) *n-el-wa:ki* *ø-yo:l-miki*
1S-liver-go_dry 3S-heart-die
'My breast is dry' 'He faints'

(18b) *n-ikši-miki*
1S-foot-die
'My feet are benumbed'

3.1.2.4 Agent incorporation

This case is very uncommon and occurs in the passive voice only; the incorporated N refers to the agent:

(19) *ø-ko:wa:-kwa-lo* *ø-eka-tok-o*
3S-snake-eat-PASS 3S-wind-follow-PASS
'He gets bitten by a snake' 'He is pushed by the wind'

3.2 Noun compounding: typology

Turning now to noun compounding, the following is the classification I proposed in LAUNEY (1986). I prefer, for grammatical as well as for intuitive reasons, to use N1-N0, which shows that N0 is the head of the compound word (just as the compound word N-V is a verb), rather than N1-N2, which only refers to morpheme order and could lead to deceptive paraphrases.

3.2.1 NC (Class intersection)

Here we have 'combined properties' (motivating my original abbreviation 'NC (comb)') or class intersection, i.e. the N1-N0 compound noun refers to an entity which is both N1 and N0, at least to some extent, e.g. by resemblance.

(20a) *ø-siwa:-pil-li* *ø-tlaso'-kwi:ka-tl*
3S-female-noble-ABS 3S-precious-song-ABS
'(She is a) noble lady' ('It is a) sacred song'

(20b) *ø-ayo:-to:ch-tli* *ø-tetzaw-ko:wa:-tl*
3S-turtle-rabbit-ABS 3S-omen-snake-ABS
'(It is an) armadillo' ('It is a) *diadophis regalis*'

One subtype is where N1 refers to a material:

(21) *ø-teo:kwitla-ko:ska-tl* *ø-e:wa:-chi:mal-li*
3S-gold-necklace-ABS 3S-skin-shield-ABS
'(It is a) golden necklace' '(It is a) leather shield'

3.2.2 Adverbial compounds

Here N1 is in an adverbial relationship with N0, and generally refers to a place where N0, or a subclass of N0, can be found:

(22) *ø-a:-toto:l-in* *ø-kwaw-kimich-in*
3S-water-turkey-ABS 3S-wood-mouse-ABS
'(It is a) pelican' '(It is a) forest mouse'

One subtype is where N1 refers to a part of the body:

(23) *ø-i:-kwa:-kwaw* *ø-kech-omi-tl*
3S-Pos3S-head-wood (POS) 3S-neck-bone-ABS
'(It is) its horn(s)' '(It is a) neck bone'

Another subtype is when the relationship can be expressed in terms of use, purpose or possession:

(24a) *ø-mo:l-kaši-tl* *ø-tle-ma:yi-tl*
3S-sauce-plate-ABS 3S-fire-hand-ABS
'(It is a) *molcajete*' [12] '(It is) a poker'

(24b) *ø-tla'to'ka:-kal-li* *ø-ma:sa-tlakwal-li*
3S-king-house-ABS 3S-deer-food-ABS
'(It is a) palace' '(It is) horse food'

12 Sort of little mortar used to grind tomatoes, chili etc.

A third subtype is where there is an instrumental semantic relationship. This occurs mostly with deverbative nouns (action or state nouns):

(25a) *ø-its-miki-lis-tli*
3S-obsidian-die-NOM-ABS
'(That is) death by obsidian'

(25b) *ø-tla'tlako:l-katza:wa-ka:-yo:-tl*
3S-sin-get_dirty-NOM-ABSTR-ABS
'(It is the) stain of sin'

A fourth specific case occurs with deverbative object nouns (which refer to the resulting object); N1 can refer to the agent:

(26) *ø-to:ka:-tzawa-l-li* — 3S-spider-weave-NOM-ABS — '(It is a) cobweb'
ø-teo:-tla-'to:-l-li — 3S-God-INDEF-say-NOM-ABS — '(These are) God's words'

3.2.3 Restrictive compounds

This latter type is uncommon. There is a restrictive relationship, which can be of two kinds: either, as in (27a) the entity denoted by the compound is a subspecies of N0 which is characterized by N1; or as in (27b) the entity is N0 only as regards N1.

(27a) *ø-ko:ska-kwa:w-tli* — 3S-necklace-eagle-ABS — '(It is a) king vulture'
ø-ketzal-ko:wa:-tl — 3S-feather-snake-ABS — '(It is) Quetzalcoatl (the feathered snake deity)'

(27b) *ø-kwa:-pitza:wa-k* — 3S-head-get_thinner-NOM — 'He has a thin head'
ø-yo:llo'-te-tl — 3S-heart-stone-ABS — 'He is brave-hearted'

4. Paraphrases and morphosyntactic structures

The above classification of compound nouns, at least, now seems very unsatisfactory to me and needs some revamping. The striking syntactic similarities between nouns and verbs in an omnipredicative language like Nahuatl may lead us to expect that incorporation and noun compounding should at least to some extent be parallel. This parallelism, if it exists, cannot be expressed in intuitive terms: we have to work with explicit syntactic structures. I know

that many linguists will be reluctant to accept the idea of an internal syntax within compound words, although I think that it is hardly a theoretical scandal, and that it may give us very interesting insights[13]. Whatever our position on this issue, we can turn to the explicit syntax of paraphrastic sentences or phrases, following MITHUN's (1984) remark that "all languages which exhibit such morphological structures also have syntactic paraphrases". In the frame of the omnipredicative hypothesis as it appears in § 2, I will follow two methodological principles:

(a) intransitive verbs and absolute nouns are one-argument predicates, while transitive verbs and possessed nouns are two-argument predicates.

(b) there is no direct relationship between a phrase and an argument position marked by an affix, but only between two argument positions.

Let us begin with incorporation, which is easier to deal with.

4.1 The morphosyntax of incorporation

In the saturating incorporation type, when objects are involved, paraphrases are sentences with an indefinite or generic object, which in Nahuatl is expressed with no determiner, e.g. *ni-naka-kwa* (11b), literally 'I-meat-eat' is paraphrased by *ni-k-kwa nakatl*, with an object phrase which can be symbolized by (28).

(28) *ni$_i$-k$_j$* *-kwa* *ø$_j$-naka-tl*
1S-3$_j$O -eat 3$_j$S-meat-ABS
'I eat meat.'

The paraphrase rule can be expressed more generally as (29), where ">" does not mean 'is rewritten as', but rather 'may be paraphrased by'.

(29) s_i-o_j-V + s_j-N > s_i-N-V
('X Vs Y and Y is an N' > 'X N-Vs')

The adverbial type of modifying incorporation, exemplified by *ni-k-te-wi:teki* literally 'I stone-hit him' in (14a), is somewhat paradoxical, because the incorporated N shares no ar-

13 See the notion of "syntactic compression" as used by Seiler (1977). In an admittedly different approach, the very wide use of the term incorporation in Baker (1988) allows for many bridges between morphology and syntax.

gument coreference with the verb. This is the characteristic feature of adverbial relationship (in French: *relation circonstancielle*) which is linked to the sentence as a whole.[14] The paraphrase will show the N inside an adverbial phrase of some kind, either by derivation (30a), or as in (30b), as the possessor of a possessed "preposition" (this term is inappropriate in Nahuatl syntax, but I use it here for brevity's sake).

(30a) *te-tika* *ni-k-wi:teki*
stone-INSTR[15] 1S-3O-hit
'I hit him with a stone. '

(30b) *ni-k-wi:teki* *i:$_i$-ka* *ø$_i$-te-tl*
1S-3O-hit Pos3S-with 3S-stone-ABS
(= 30a)

The paraphrase rule for this type is symbolized in (31):

(31) s_i-(o_j-)V + $_{Adv}$[... s_k-N ...]$_{Adv}$ > s_i-(o_j-)N-V
('X Vs in some N circumstances' > 'X N-Vs')

In the type of modifying incorporation expressing comparison or secondary predication, there is argument coindexation with the subject (e.g. 32a) or object (e.g. 32b). The comparative meaning can be made explicit by *iwki* ('similar') and/or *ik* ('the way in which').

(32a) *ø-iwki* *(in)* *ni$_i$-koyo:-tl* *in* *(ik)ni$_i$-ne'nemi*
3S-similar (DET) 3S-coyote-ABS DET (how) 3S-walk
Lit.: 'Similar to[16] I am a coyote (is how) I walk'. Cf. *ni-koyo:-ne'nemi in* (16a).

(32b) *ø$_i$-kwal-li* *in* *(ik)* *ni$_j$-k$_i$-itta*
3S-good-ABS DET (how) 1S-3O-see
'I approve it', lit.: 'it is good (is how) I see it'. Cf. *ni-k-kwal-itta in* (16c).

This paraphrase relation is summed up in (33a-b), which deal respectively with incorporated predicates on the subject and object.

(33a) s_i-(o_j-)V + (*iwki*) s_i-N > s_i-(o_j-)N-V[17]
('X Vs as if X were an N' > 'X N-Vs')

---

14 I will not show here how this fits into the frame of omnipredicativity; see Launey (1994), § 3.4.

15 This is more a derivative than an instrumental case suffix.

16 *iwki* is a predication on the whole clause in *(ik) ne'nemi.*

17 The optional object slot is needed here to account for examples like *ki-pitzo:-neki* (16b).

(33b) s_i-o_j-V + (*iwki*) s_j-N > s_i-o_j-N-V
('X Vs Y as if Y were an N' > 'X N-Vs Y')

In the Possessor-raising type of modifying incorporation, there is a shift of the external argument position (object of transitive verbs, subject of intransitive verbs) onto the possessor of the object (e.g. 'I wash his head' becomes 'I head-wash him', 'his heart stops' becomes 'he heart-stops'). This gives us the paraphrase of possessor-raised transitives like *ni-mitz-tla'to:l-kaki* in (17a) as (34a), and of possessor-raised intransitives like *n-el-wa:ki* in (18a) as (34b).

(34a) *ni_i-k_j-kaki* *in* *$ø_j$-mo_k-tla'to:l*
1S-3O-hear DET 3S-2POSS-words
'I understand you', lit.: 'I hear it, that which is your words. ' (cf. 17a)

(34b) *$ø_i$-wa:ki* *in* *$ø_i$-n_j-el*
3S-go_dry DET 3S-1POSS-liver
'my breast is dry', lit.: 'it goes dry, that which is my liver'

Note, however, that another (and maybe more usual) paraphrase of (17a) is (34c), where the so-called applicative voice is used. This voice is usually marked by the suffix -(i)lia and forms ditransitive verbs from transitive ones by adding a new argument position with a benefactive or detrimental meaning.

(34c) *ni_i-$mitz_k$-[*k_j-*[18]*]kaki-lia* *in* *$ø_j$-mo_k-tla'to:l*
1S-2O-[*3O-]hear-APPLIC DET 3S-2POSS-words
(= 34a)

The first two of these paraphrases are symbolized, respectively, as (35a) and (35b):

(35a) s_i-o_j-V + s_j-$poss_k$-N > s_i-o_k-N-V
('X Vs Y which is the N of Z' > 'X N-Vs Z')

(35b) s_i-V + s_i-$poss_j$-N > s_j-N-V
('The X - which is an N - of Y Vs' > 'Y N-Vs')

The *agentive* type of modifying incorporation is tricky. Although (36) is an attested paraphrase for *ø-kowa-kwa-lo* in (19), I am not sure that (37), however tempting it may be, is really the source paraphrase.

---

18 In ditransitive verbs, if both objects are definite, only one of them (namely, the "dative") is marked.

(36) *ø$_i$-ki$_j$-kwa* *ø$_i$-ko:wa:-tl*[19]
3S-3O-eat 3S-snake-ABS
'He gets bitten by a snake.'

(37) s_i-o_j-V + s_i-N > s_j-N-V-PASS

Since s_i is indefinite in such a case, the paraphrase should be more or less 'something Vs Y and something is an N' > 'Y is N-Ved'. We are here in a quandary: if s_i is unknown (not just as a token, but also as a type), we can build a passive, with the subject properties transferred to the second argument. However, since the type is unknown, we cannot apply a noun predicate to this unknown entity. On the other hand, if we do apply a predicate to this entity, it will be expressed by an argument phrase, as in (36), but in that case we cannot form a passive. Nahuatl passives are strictly intransitive verbs which appear when the agent is totally excluded from the argument structure: no affix stands for the agent, and it cannot be expressed by any kind of phrase. But there is an escape hatch: actually there is no specifically agentive type of modifying incorporation, but just a subtype of adverbial incorporation in which N is in an adverbial relation to the passive verb. So first a passive verb is formed; then (and as it was said, uncommonly), an N which has an agentive meaning can be incorporated *precisely because in that case the agent is outside the argument structure.* So there is nothing like (37); rather, we are dealing with a variant of (31) with an intransitive verb, comparable to (15a-c).

4.2 The morphosyntax of noun compounding

We have seen that there are five (or one plus four) types of incorporation, and that each of them corresponds to a particular argument structure, which can be symbolized by (29), (31), (33a-b), (35a-b) and maybe (37), but as we said this latter is implausible. Let us now see if we find the same structures in noun compounding, after replacing N-V by N1-N0, and the object position by a possessor position.

The nominal equivalent of the alternation between external nominals and saturating incorporation, i.e. of the paraphrase summarized by (29), is illustrated by the following examples, which give 'expanded' versions of (24), (25) and (26):

(38a) *ø$_i$-i:$_j$-kaš* *ø$_j$-mo:l-li* > *ø$_i$-mo:l-kaši-tl*
3_iS-3_jPOSS-plate 3_jS-sauce-ABS 3_iS-sauce-plate-ABS
'(it is) the plate of sauce'

19 Note that this sentence is ambiguous: if the subject of the subordinate noun is coindexed with the object of the verb, it means 'he eats a snake'.

(38b) *ø$_i$-i:$_j$-kal* *ø$_j$-tla'toa:ni*[20] > *ø$_i$-tla'to'ka:-kal-li*
3$_i$S-3$_j$POSS-house 3$_j$S-king(ABS) 3$_i$S-king-house-ABS
'(it is) a king's house, (it is) a palace'

(38c) *ø$_i$-i:$_j$-tla'to:l* *ø$_j$-teo:-tl* *(> ø$_i$-teo:-tla'to:l-li)*
3$_i$S-3$_j$POSS-word 3$_j$S-God-ABS 3$_i$S-God-word-ABS
'(these are) God's words'

The paraphrase rule for these compounds is:

(39) s_i-poss$_j$-N0 + s_j-N1 > s_i-N1-N0
('X is someone's/something's N0 and that someone/something is an N1' >
'X is a N1-N0')

Note that in the paraphrase N0 is in its possessed form, which is parallel to transitive verbs, while the compound N1-N0 is in the absolute form, parallel to intransitive verbs. Note also that this formulation does not apply to forms like *itz-miki-lis-tli* ((25a), and see (42) below), where N1 is instead in adverbial position.

The nominal equivalent of adverbial incorporation (see paraphrase rule in 31) would be locative adverbial compounds like (22), which have well-attested paraphrases like the following:

(40) *ø-kimich-in, in kwaw-tla'* *ø-nemi* > *ø-kwaw-kimich-in*
3S-mouse-ABS DET wood-LOC 3S-live 3S-wood-mouse-ABS
'it is a mouse that lives in the woods'

This suggests the following paraphrase rule:

(41) s_i-(poss$_j$-)N0 + $_{Adv}$[...s_k-N1...]$_{Adv}$ > s_i-(poss$_j$-)N1-N0
('X is N0 in some N1 circumstances' > 'X is a N1-N0')

This also accounts for cases like (25) which do not fit the pattern of paraphrase in (38-39). Thus corresponding to *itz-miki-lis-tli* '(it is) death by obsidian' we do not find **ø$_i$-i:$_j$-miki-lis ø$_j$ itz-tli* 'obsidian's death', but rather:

(42) *itz-tika* *ø-miki-lis-tli*
obsidian-INSTR 3S-die-NOM-ABS
'(it is) death by obsidian'

---

20 This is the absolute form for a stem which otherwise appears as *tla'to'ka:-*.

The nominal equivalent of incorporation of a comparison or second predicate on the subject is illustrated by (43a, b), which are analytic paraphrases of (20a) and (21).

(43a) *ø$_i$-siwa:-tl,* *ø$_i$-pil-li* *(> ø$_i$-siwa:-pil-li)*
3$_i$S-female-ABS 3$_i$S-noble-ABS 3$_i$S-female-noble-ABS
'she is a woman (and) a noble person'

(43b) *ø$_i$-e:wa:-tl,* *ø$_i$-chi:mal-li* *(> ø$_i$-e:wa:-chi:mal-li)*
3$_i$S-skin-ABS 3$_i$S-shield-ABS 3$_i$S-skin-shield-ABS
'it is leather, (and) it is a shield'

The paraphrase rule for this type can be formulated as:

(44) s$_i$-(poss$_j$-)N0 + s$_i$-N1 > s$_i$-(poss-)N1-N0
('X is an N0 and X is (like) an N1' > 'X is an N1-N0')

Corresponding to incorporation of secondary predicates on the object, on the other hand, are examples like (45), with its first person possessor. Its paraphrase would effectively be: 'It is my house and I am (like) a king'.

(45) *ø-no-tla'to'ka:-kal*
3S-1POSS-king-house
'(it is) my palace'

The paraphrase rule for compounds of this type is:

(46) s$_i$-poss$_j$-N0 + s$_j$-N1 > s$_i$-poss$_i$-N1-N0
('X is an N0 of Y and Y is (like) an N1' > 'X is an N1-N0 of Y')

This latter type only occurs in some case of possessive constructions, where N1 can be applied to the possessor as a predicate. For the possessed form of (24a), for example, i.e. *ø-no-mo:l-kaš* 'it is my sauce dish', this would be impossible, since it would be ludicrous to say *ni-mo:l-li* 'I am sauce'. The only interpretation for *ø-no-mo:l-kaš* is that the scope of the possessor is the whole compound. But this latter interpretation might as well apply to (45), which is thus an ambiguous form. What appears here is the well-known ambiguity of genitives, which is sometimes transferred to noun compounding (see the above quotation of Mithun 1984); this can be expressed in syntactic or in semantic terms according to one's theoretical background. But there is one interesting point. We have seen before that the absolute forms of (24) may be symbolized by (39), that is to say, the nominal equivalent of saturating incorporation. This feeds back into our analysis of modifying incorporation involving comparison and secondary predicates. This appears to be related to saturating in-

corporation by adding an argument place which specifies that a certain entity is considered an N. Indeed, in many Nahuatl grammars (if I am not mistaken, the first occurrence is in Carochi 1645) the following pair can be found, where the same N-V compound is in the first case an example of saturating incorporation, and in the second case an example of comparison incorporation:

(47a) *Ni-šo:chi-te:moa*
1S-flower-seek
'I seek flowers'

(47b) *Ni-k-šo:chi-te:moa ø-kwi:ka-tl*
1S-3O-flower-seek 3S-song-ABS
'I seek songs like flowers'

The nominal equivalent of possessor-raising incorporation with an intransitive verb (35b) is illustrated by the following expanded version of the restrictive nominal compound in (27):

(48) *$ø_i$-te-tl (in) $ø_i$-$i:_j$-yo:llo' > $ø_j$-yollo'-te-tl*
3S-stone-ABS (DET) 3S-3POSS-heart 3S-heart-stone-ABS
'(like) stone is his heart'

This conforms to the paraphrase rule:

(49) s_i-N0 + s_i-$poss_j$-N1 > s_j-N1-N0
'X is a N0 and X is a N1 of Y' > 'Y is a N1-N0'

Corresponding to possessor-raising incorporation with transitive verbs we would expect the following paraphrase relationship:

(50) s_i-$poss_j$-N0 + s_j-$poss_k$-N1 > s_i-$poss_k$-N1-N0
'X is a N0 of Y and Y is a N1 of Z' > 'X is a N1-N0 of Z'

But once again we run into difficulty with finding a compound pattern corresponding to incorporation with a transitive verb. (50) would fit for the possessed forms of (23), since we can paraphrase *$ø_i$-$i:_k$-kwa:-kwaw* 'it is its horn(s), i.e. it is its head-wood' by 'it_i is its_j (i.e. the head's) wood, and it_j is its_k (i.e. the cow's) head', or 'it is its (i.e. the cow's) wood as far as its head is concerned'. But the absolute forms themselves are rather of type (39), otherwise used to paraphrase compounds of use, purpose and possession: thus *ø-kwa:-kwawi-tl* 'it is a head-wood, a horn', or *ø-kech-omitl* 'it is a neck-bone' (see (23)) which can be paraphrased by *$ø_i$-i_j-omi-w in $ø_j$-kech-tli* 'it is its bone (of) the neck'. Here once more this feeds back into our analysis of incorporation, since possessor-raising noun incorporation is in a

certain way related to saturating incorporation by raising the possessor (genitive) of the object: we probably could have a form like *ni-tla'to:l-kaki* 'I hear words' related to (17a) *ni-mitz-tla'to:l-kaki* 'I hear your words' > 'I word-hear you', i.e. 'I understand you'. In short, saturating incorporation seems in some sense to be the prototypical form of incorporation.

Now let us turn to the agentive incorporation type. Here we face the problem of finding a nominal equivalent for the passive. But this indeed exists, namely the so-called *possessive nouns*, which are formed by suffixing *-wa'* or *-e'* (details are irrelevant here) to the noun stem, and which were first translated in the early Nahuatl grammars as 'owner of N' (in Spanish 'dueño de'), e.g.:

(51) *ni-kal-e'* 'soy dueño de casa', 'I am (the) owner of a house'

This translation is obviously awkward. Actually, possessive nouns are the most straightforward translations of sentences with *have* in languages which do have such a verb: if we want to preserve their nominal nature, we could say that where English says 'I have a house', Nahuatl says something like 'I am house-provided'. And these nouns are remarkably parallel to Nahuatl passives, which are intransitive verbs derived from transitive ones when the agent is indefinite and therefore must remain unexpressed. This is clear from the English glosses below:

(52a) 'he sees me': *ø-ne:ch-itta*
(52b) 'someone sees me' > 'I am seen': *n-itta-lo*
(53a) 'it is my house': *ø-no-kal*
(53b) 'something is my house' > 'I am house-provided': *ni-kal-e'*

So there should be no difficulty for a nominal parallel to (37), which would be:

(54) s_i-$poss_j$-N0 + s_i-N1 > s_j-N1-N0-*provided*

But we are in the same quandary as for (37). We could indeed find some grammatically correct forms like 'I'm provided with a wooden house, I'm wood-housed' (*ni-kwaw-kal-e'*), but the derivational path is certainly not (54), but rather a possessor noun built on a pre-constructed compound noun. So it may be the case that, just as there is no real agent incorporation, there is nothing like its nominal equivalent.

There remains a problematic case, namely (27a), which seems to have no verbal equivalent. An underlying morphosyntactic structure could be:

(55) s_i-N0 + s_j-$poss_i$-N1 > s_i-N1-N0
('this$_i$ is an eagle' and 'something$_j$ is its$_i$ necklace' > 'it is a necklace-eagle').

(55) is much like (49), except that the coindexation of the subject and the possessor of N1 is reversed, and that we need an indefinite paraphrase for the subject of N1 (we meet again a *have*-relationship: 'something is his N' > 'he is provided with N'). A verbal equivalent would be:

(56) s_i-V + s_j-poss$_i$-N > s_i-N-V
('X V-s and something is X's N' > 'X N-V-s' or 'X V-s as far as his N is concerned')

This again is (35b) after reversing the argument indices of N. And after all, it could be an alternative paraphrase for (18a-b), where there is a part-to-whole relationship between the incorporated N and the subject. If my feet die, I also die in this respect. (See EVANS (1995:90-91) for a similar analysis of body-part incorporation in Gun-djeihmi).

5. Conclusion

The combination of head-marking, omnipredicativity and compounding makes Nahuatl a polysynthetic language *par excellence*. We have seen that the parallelism between nouns and verbs in their syntactic behaviour is extended to their ability to form compound words, and that the structures which underly incorporation and noun compounding (or, in a theoretically more neutral way, their possible paraphrases) also are remarkably parallel. However, we have seen some discrepancies. One of them is conspicuous: the head element of these compound words can be a noun as well as a verb, but the first element can only be a noun. This is one more of the features which oppose verbs and nouns, and it must be added to the few ones listed in note (9). Maybe this constraint has something to do with the presence of tense-aspect in the verb, which provides a link with the situational reference of verbs, so that prototypically (and even if they may have a generic present or a gnomic aorist among their forms) verbs are apt to express particular events. The aspectual stability of nouns, on the other hand, enables them to express tenseless properties and to have a generic reference. We are thus compelled to say that aspectual stability is the condition for the type of semantic and syntactic modification conveyed by both incorporation and noun compounding.

The second difference is that, while a parallelism always seems easy to draw between an absolute N1-N0 noun and an intransitive N-V verb, difficulties arise as soon as we deal with possessed nouns. It appears that if transitivity is a normal (and probably prototypical[21]) status for verbs, nouns are prototypically one-argument predicates, and that their two-argument form, if any, results from the secondary development of an argument (genitive) place, which in its turn is semantically many-sided. We meet here an unexpected and non-trivial side-effect of the tense-aspect parameter.

---

21 This idea is congruent with the unaccusative hypothesis, see LAUNEY (1994) § 5.

ABBREVIATIONS

ABS = Absolutive suffix
ABSTR = Abstract noun suffix
DET = Determiner
LOC = Locative suffix
NOM = Nominalizer
O = Object
PASS = Passive
PL = Plural
POSS = Possessive
REFL = Reflexive
S = Subject

REFERENCES

ABNEY, S. (1987). *The English NP in its Sentential Aspect.* MIT. Unpublished Ph.D. dissertation.

BAKER, M.C. (1988). *Incorporation.* Oxford & New York. Oxford University Press.

BAKER, M.C. (1996). *The Polysynthesis Parameter.* Oxford & New York. Oxford University Press.

BENVENISTE, E. (1966). Convergences typologiques. *L'Homme* N° 2: 5-12. Paris.

BOAS, F.(ed.). (1911). *Handbook of American Indian Languages.* Washington. D.C.. U.S. Government Printing Office.

CAROCHI, H. (1645). *Arte de la lengua mexicana.* Mexico.

CRYSTAL, D. (1987). *The Cambridge Encyclopedia of Language.* Cambridge University Press.

DUBOIS, J. *et al.* (1994). *Dictionnaire de linguistique et des sciences du langage.* Paris. Larousse.

EVANS, N. (1996). The Syntax and Semantics of Body Part Incorporation in Mayali. In: Chappell. H. and W. McGregor(eds.).*The Grammar of Inalienability: A Typological Perspective on Body Part Terms and the Part-Whole Relation.* Berlin. Mouton de Gruyter

FOX, D. (1966). Quiche grammatical sketch. in MAYERS (Ed.). *Languages of Guatemala.* The Hague. Mouton.

JELINEK, E. (1984). Empty categories, case, and configurationality. *Natural Language and Linguistic Theory* 2:39-76.

LAUNEY, M. (1986). *Catégories et opérations dans la grammaire nahuatl.* Thèse d'Etat. Université de Paris IV.

LAUNEY, M. (1994). *Une grammaire omniprédicative.* Paris. CNRS Editions.

MITHUN, M. (1984). The Evolution of Noun Incorporation. *Language* 60 (4): 847-894.

NICHOLS, J. (1986). Head-marking and dependent-marking grammar. *Language* 62: 56-119.

OSBORNE, C.R. (1974). *The Tiwi Language: Grammar, Myths, and Dictionary of the Tiwi Language Spoken on Melville and Bathurst Islands, Northern Australia.* Canberra. A.I.A.S.

QUENEAU, R. (1965). Connaissez-vous le chinook? In: *Bâtons, chiffres et lettres.* Paris. NRF.

SAPIR, E. (1921). *Language.* New York. Harcourt Brace Jovanovich.

SEILER, H. (1977). *Cahuilla Grammar.* Banning (California). Malki Museum Press.

SPENCER, A. (1991). *Morphological Theory.* Oxford. Blackwell.

Dependent-Head synthesis in Nivkh - with an outlook on polysynthesis in the Far Northeast

JOHANNA MATTISSEN

1. Preliminary remarks

Morphological synthesis phenomena have been described within the frameworks of polysynthesis, noun incorporation, or verb root serialization (MITHUN 1988, FORTESCUE 1994, BAKER 1996, DROSSARD 1997; MITHUN 1984, BAKER 1988, ROSEN 1989, GERDTS 1996; BISANG 1995). While noun incorporation and verb root serialization focus on single domains of syntax, polysynthesis comprises a range of heterogeneous phenomena. In this paper a structural principle will be presented which organizes syntax in a homogeneous way through systematic synthesis of head and dependent. This structural principle will be referred to as dependent-head synthesis, with the concepts of head and dependent used here in the sense of NICHOLS (1986). A language which is consistently organized in this way is Nivkh (Gilyak).

2. Basic typological characteristics of Nivkh

Nivkh is an endangered language isolate (usually counted among the Paleosiberian language group) spoken on the lower reaches of the Amur river and on Sakhalin Island in Eastern Siberia in several dialects.

The predicate in Nivkh is clause-final. Syntactic relations are signalled by word order (central participants), case suffixes (peripheral participants), postpositions (local relations), and, most importantly, dependent-head synthesis: the primary object (in the sense of DRYER 1986) synthesizes with its governing verb. Agent or subject agreement is only found on some non-matrix predicate forms (see (9) below for an example). Clausal nexion operates on a converbal basis.

Nivkh distinguishes number and has numeral classification. There is neither a gender nor an article system. Grammatical categories are encoded through inflection of the agglutinative type (number, case and focus; causative, temporal, aspectoid and modal categories) and through auxiliaries (negation, phases of states of affairs, modalities).

3. Nivkh morphophonemics

Morphophonemics play the decisive role for recognizing dependent-head synthesis in Nivkh. At the boundaries of adjacent lexical and/or grammatical morphemes initial and final consonant sounds are subject to morphophonemic alternations reminiscent of mutation in Celtic (cf. BALL & MÜLLER 1992). As the basis for the description of these alternations we take that form of a morpheme which represents the lexeme in a construction containing the morpheme in question alone, e.g. *təf* 'a/the house (NOM)'. In this case its initial sound is in absolute initial position and its final sound in absolute final position. This form serves as the citation form of a nominal lexeme (see SAVEL'EVA & TAKSAMI 1970). The citation form of a verbal lexeme, whose root cannot stand alone as a free form without at least one suffix, contains the indicative suffix *-ʂ ~ -ʃ* (Amur dialect), *-d ~ -nt* (Sakhalin dialects) as its only suffix. This suffix assimilates with respect to voice to a root-final consonant (in case there is one) without inducing any sound changes in the root, e.g. *ev-ʂ* 's.o. takes s.th.' vs. *esp-ʃ* 's.o. stabs s.o.'.

Due to the different types of morphophonemic alternations presented in (i) to (iv) below, lexemes and affixes have several allomorphs:

(i) Voiced root-final fricatives manifest final devoicing in absolute final position (for examples see JAKOBSON 1971:83).

(1) *təv-* 'house': *təf* (free form) vs. *təv-ux* ABL, *təv-mi* 'inside the house'[1]

(ii) A morpheme-initial fricative or plosive alternates depending on the final sound of a preceding morpheme. Occlusive-initial nouns generally begin with a plosive, occlusive-initial transitive verbs with a fricative. This initial sound alternates with a homorganic occlusive in two sets.

alternation set A[2]

p	t	ʈ	k	q
~	~	~	~	~
v	r	z	ɣ	ʁ
~	~	~	~	~
b	d	ɖ	g	ɢ
~	~	~	~	~
f	ř	s	x	χ

alternation set B

pʿ	tʿ	ʈʿ	kʿ	qʿ
~	~	~	~	~
f	ř	s	x	χ

1 Transcriptions and glosses have been changed in most examples by the author for the sake of consistency.

2 <ʈ>: [ʧ], <ɖ>: [ʤ] pattern like plosives, <pʿ> etc.: [pʰ] etc., /r/ and /r̃/ ([r̲]) pattern like fricatives.

A voiceless or aspirated plosive is found after a final fricative of an immediately preceding morpheme, a voiced or voiceless fricative after a plosive, "strong" nasal, glide, or vowel, a voiced plosive or aspirated plosive after a "weak" (eliding) nasal, and a voiceless fricative after a voiceless sound final prefix. Table 1 displays an overview of these alternations in the Amur dialect. As a rule of thumb, the alternation patterns like dissimilation after obstruents and like assimilation after vowels and nasals. An example is:

(2) (COMRIE 1981:267)
citation form: *rəmzuq* 's.o. dropped s.th.'
paχ-təmzuq 's.o. dropped a stone'
lep-rəmzuq 's.o. dropped bread'
ŋir(ŋ)-dəmzuq 's.o. dropped a [wooden] dish'

preceding morpheme	following morpheme (head of construction)	
final sound	initial sound (alternation set B)	initial sound (alternation set A)
plosive, "strong" nasal, glide, vowel	voiceless fricative	voiced fricative
fricative	aspirated plosive	voiceless plosive
"weak" nasal		voiced plosive
voiceless prefix	voiceless fricative	voiceless fricative

TABLE 1

(iii) Singular pronouns and pronominal prefixes are subject to vowel harmony.
(iv) Root initial /h/ elides after a prefix.

The alternations described above do not show up arbitrarily, but rather occur if the adjacent morphemes are in certain grammatical relations to each other: primary object and verbal root, adverbially modifying root and verbal root, attributive nominal or verbal root and nominal head, and possessor or determiner and head noun (as will be shown in detail in sections 5 and 7). In other words, they occur if two adjacent morphemes are in a government or modifying relationship, or, in NICHOLS' terms, if they are dependent and head. Dependent and head then form a complex. Furthermore, the alternations occur at the morpheme boundary of a lexical and a grammatical morpheme (prefix, number, case, focus, tense/aspectoid

suffixes) or between two suffixes. For an example, compare the allomorphy of the 'highlighting focus' morpheme in (3) with example (2):

(3) *-ta* / {fricative}+ _
-da / {"weak" nasal}+ _
-ra / {plosive, "strong" nasal, glide, vowel}+ _

4. Dependent-head synthesis and wordhood

The morphophonemic alternations suggest a certain degree of synthesis of the adjacent morphemes in a complex. And indeed BONDARKO & ZINDER (1962) proved by experiment that there are chains of adjacent morphemes which form single accentual units. Their oscillograms show that in

(4) (BONDARKO & ZINDER 1962:89)
ñivx *ləɣi-ŋəŋ-ɻ*
person salmon-hunt-IND
'man catches salmon'

there are two amplitudinal peaks (not three), one on the [i] of *ñivx* and the other on the [ə] of *ləɣi*, which means that there are two accentual units. In

(5) (BONDARKO & ZINDER 1962:89)
ləɣi-ŋəŋ-ñivx
salmon-hunt-person
'salmon hunter'

there is only one peak on the first [ə]. In other words, the whole construction is one accentual unit. There are no morphophonemic alternations here as an additional proof on the morphological level since there are no critical consonants involved. Parallel constructions involving morphophonemic alternations behave analogously, however (see BONDARKO & ZINDER 1962:85-89). Besides BONDARKO & ZINDER (1962), KREJNOVIČ (especially 1937, 1958, 1966), JAKOBSON (1971:78, 80), and WATANABE (1983:63; 1992:185) acknowledge that constructions in which the abovementioned relations hold are word-like forming a single unit from the point of view of accent and pause, and they use hyphens between the morphemes of such a complex. KREJNOVIČ (1966:37-38) and NEDJALKOV & OTAINA (1988:136) distinguish between free and bound variants of lexical morphemes, with the latter being those which occur in the complexes. PANFILOV, although not disputing the accentual unity of the constructions, writes their elements as separate words because, according to his view,

words are not defined morphophonologically, but are lexical-grammatical entities characterized by an independent, atomic meaning (1965:23-30). Without comment, most authors of Nivkh transcribed texts and grammatical sketches follow his example, so that complexes are "invisible" in the absence of morphophonemic alternations.

The morphological cohesion of the dependent-head synthetic constructions like the ones in (4) and (5) is an argument in favor of their interpretation as single words. They are not interruptible and are inflected as wholes: case or mood suffixes are only allowed at the end, not within, and take the synthesized elements in their scope. The order of the elements is rigid. Since accentual patterns, morphophonemic alternation domain, and cohesion domain are coextensive, we could consider the emerging entities to be single words.

As an alternative to the recognition of single words, "phrasal sandhi" or the concept of a "phonological phrase" (NESPOR & VOGEL 1986:ch.6) seem to offer convenient interpretations, but are not adequate on closer examination. Sandhi processes are rather loosely defined (cf. ANDERSEN 1986:1-2), and apply not only within units greater than the word, but also within the word, as in Sanskrit. Hence, to describe the morphophonemic alternations in Nivkh as sandhi does not imply anything about their status. Sandhi also tends to depend on the register (cf. SPENCER 1996:200-201), but in Nivkh the morphophonemic alternations are obligatory.

The phonological phrase (as described by NESPOR & VOGEL 1986) is the domain of phonological phenomena which are both phonologically and syntactically conditioned, but the phonological phrase does not coincide with a constituent. In NESPOR & VOGEL's (1986) examples from Italian "raddoppiamento sintattico" and from French "liaison", the relevant domains for the phonological rules are the complexes [auxiliary-preverbal adverb-verb] and [quantifier or determiner-prenominal adjective-noun]. As the phonological rules are blocked if an attribute is postnominal, NESPOR & VOGEL claim that the domain of the phonological phrase consists of the head and all elements of its phrase on the non-recursive side (1986: 168). In Nivkh, however, morphophonemic alternations apply to domains containing the head and its recursive side, and they apply to VPs (verb and object) and NPs. Furthermore, as modifying or governed elements cannot stand to the other (right) side of their heads in Nivkh, there is no non-recursive side in the first place. Additionally, the Italian phonological phenomenon optionally applies even across syntactic constituent boundaries (NESPOR & VOGEL 1986:172-173) and is thus partially syntactically blind, while the Nivkh rule applies strictly within phrases only, i.e., it is attentive to syntactic relations. In particular, Nivkh morphophonemic alternations apply between elements in a dependency relation only, i.e. not under coordination (cf. MATTISSEN & DROSSARD 1998:53). This means that the domain of Nivkh alternations can be smaller than a syntactic counterpart, in which characteristic it again parallels the domain of the phonological phrase (NESPOR & VOGEL 1986:41-42).

In their application to NPs and VPs, Nivkh alternations seem to parallel Celtic mutations, but in Welsh, for example, mutations are not phonologically conditioned (BALL & MÜLLER

1992:4-5), occur under non-adjacency in "VPs" (BALL & MÜLLER 1992:146-160, 128, 134, 137), under both adjacency and non-adjacency in NPs, and they can be blocked by intervening elements (BALL & MÜLLER 1992:155, 161-163). Besides, in Nivkh the same alternations which apply to [object-verb] and [modifier-noun] apply to [lexical morpheme-inflectional suffix], i.e. between uncontroversionally bound elements as well. Welsh mutations within a word, however, i.e. under composition or inflection, are (because they are fortitions) different in type from mutations involving "VPs" and NPs, which are lenitions (cf. BALL & MÜLLER 1992:284-286).

For all these reasons, the simplest analysis of the relevant constructions in Nivkh is to treat them as single units at both the phonological and morphological levels.

5. The Nivkh verb complex

A first domain of dependent-head synthesis in Nivkh is the verb complex. A verb root synthesizes with (i) a noun bearing an object (or marginally other) relation to it and/or with (ii) another verb root modifying the verb root in an adverbial manner[3] or functioning like a modal verb. The second phenomenon constitutes a kind of verb root serialization:

> "Root serialization is formally different from verb serialization in as much as we find only combinations of verbal roots, which in most cases cannot form sentences on their own." (Bisang 1995:177).

Here is a Yimas verb form for an example of this:

(6) (Foley 1991:179)
impa-mpu-yakal-irɨm-tay-ntut
3dlU-3plA-TAM-stand-see-REM.PST
'they stood watching the two of them.'

In Yimas, verb root serialization forms single words, as evidenced by inflectional affixes to both sides of the complex (Foley 1991:278-279).

---

3 There are no adjectives and only very few (non-synthesizing) adverbs in Nivkh, cf. MATTISSEN & DROSSARD (1998).

5.1. Object-verb-synthesis

Basically, Nivkh verbs can be divided into two classes: those which synthesize with a participant (a patient or a recipient/goal) as evidenced by morphophonemic alternations (= (di)transitives), and those which do not (= intransitives). These two classes are distinguished formally by their verb-initial sounds. Most intransitive verbs begin with a plosive (which does not alternate) or a vowel in their citation form. Most transitive verbs begin with a fricative, /i/, /e/, or /j/ (and do alternate). Two pairs illustrating this are:

(7) *ərpʧ* 's.o. chokes, suffocates' *j-ərpʧ* 's.o. inserts s.th./plugs s.th.up'
tʿa-ʤ 's.th. roasts' (intr.) *ǩa-ʤ* 's.o. roasts s.th./it' (tr.)

Before furnishing an explanation for this, let us have a look at the morphosyntactic properties of transitive verbs. These obligatorily form a synthetic unit with their primary object if it is adjacent (cf. SAVEL'EVA & TAKSAMI 1970:507). The primary object may either be encoded by a personal (1st to 3rd singular) or reflexive prefix or by a noun. The synthesized object saturates one place of the verb's valency and can represent one of the verb's arguments. There is no transitive verb without an overt object. In other words, the object marker is a dependent which its head in fact requires, contrary to what NICHOLS (1986:108) predicts. Its referential status ranges from generic as in (8) to fully discourse-referential (in the sense of being an antecedent for anaphoric reference, as in (9)).

(8) (OTAINA 1978:34)
porχ tuñ-daγa-ʤ
bar finger-equal_in_thickness-IND
'the bar is as thick as a finger'
raγaʤ ~ -taγaʤ ~ -daγaʤ 's.th. is as thick as s.th.' (transitive verb)

(9) (BEFFA 1982:87)
hoŋgutʿumke mūv-ñaqr pal-rox mər-ra
living_thus day-one_CLF mountain-ALL ascend-ENU:3s

qʿotr-kʿu-ra ʁoʁorot lər-γət-ra.
bear-kill-ENU:3s then divide-CPL-ENU:3s

tʿu-in ʁoñʤi-ror vəγi-ǩ məγ-ʤ.
sledge-LOC 3s:load-CV:after 3s:drag-CV:3s descend-IND
'Living thus, one day he went up into the mountain woods and killed a bear and then cut it fully up. He loaded it onto the sledge and dragging it along descended.'

Now, the initial /i, e, j/ of (di)transitive verbs are the allomorphs of the 3s undergoer prefix in the context of vowel or consonant cluster initial verbs.

(10) *i-vrə-d* 's.o. tidies up s.th.'
j-aq-ʈ 's.o. cuts s.th.'
e-mχaqu-ɻ 's.o. shortens s.th.'

The fricative initials of verbs beginning with a single consonant can be considered the effect of consonant alternation of a plosive (plosive-initial roots being suggested by non-alternating intransitive verbs, see (7)) due to a preceding vowel or glide which has been elided (cf. JAKOBSON 1971:86; HATTORI 1962). Non-singular pronouns, personal/reflexive prefixes, and nominals are mutually exclusive in the synthesized object slot - i.e. there is only one slot in a complex verb form for a governed dependent - but each triggers its own appropriate consonant alternations (see KREJNOVIĆ 1934:209). They are furthermore in complementary distribution to a fricative initial, as can be seen in the following examples:

(11) *j-uruɻ* 's.o. reads s.th.' : *pitʏə-uruɻ* 's.o. reads a book'
i-rləɻ 's.o. pulls s.th.' : *kʿe-xləɻ* 's.o. pulls a net'
ǐaɻ 's.o. fries s.th.' : *ʈus-tʿaɻ* 's.o. fries meat'
: *ʃo-ǐaɻ* 's.o. fries fish'

(12) *ñ-za-ɻ* 's.o. hit me'
ʃ-sa-ɻ 's.o. hit you'
pʿ-sa-ɻ 's.o. hit self'
za-ɻ 's.o. hit him/her/it/s.th.'
ñəŋ-ʐa-ɻ 's.o. hit us' etc.

In these synthesized N-V complexes (with N being a noun or pronoun) the verb is the head of the construction and governs the nominal, its object. The object is the patient of a transitive (see examples above) and the recipient or goal of a ditransitive verb (note that the patient is not case-marked here):

(13) (OTAINA 1978:34)
objezdʃik kʿe atak-asqam-ɻ
bay_watcher net grandfather-take_away-IND
'the bay watcher took the net away from grandpa'
free form: *jasqamɻ* 's.o. takes s.th. away from s.o.'

(14) (PANFILOV 1962:126-127)
əmək k'uva nux-t'ə-ɣ̊
mother thread needle-insert-IND
'mother inserted a thread into the needle'
free form: *řəɣ̊* 's.o. inserts s.th. into s.th.'

If *atak* 'grandfather' and *nux* 'needle' were not synthesizing, we should expect a 3sU marker (*j-* and fricative initial, respectively) on the verbs.

Besides for primary objects, synthesizing is possible for body parts bearing a non-agentive subject relation to the verb:

(15) (KREJNOVIČ 1934:221)
ñaγr mγe-ɣ̊ ŋəki-zaŋGa, əf-ʃaŋGa
rat row-IND tail-rise_aloft whiskers-rise_aloft
'the rat rowed, his tail rising aloft, his whiskers rising aloft'

(16) (KREJNOVIČ 1960:80)
ŋiv-ʁo-ɣ̊-ra
heart-ache-IND-HILI
'my heart is aching'
free form: *qoɣ̊* 's.th. aches'

Compare an alternative intransitive construction to (16):

(17) (SAVEL'EVA & TAKSAMI 1970:143)
ñ-uin təmk qo-ɣ̊
1s-LOC arm ache-IND
'my arm is aching'

Up to this point, the findings look like noun incorporation. Noun incorporation is understood to be the integration of nominal roots which can form free word forms into one accentual unit with otherwise free verbal roots (following GERDTS 1996). This excludes denominal verbs as in Greenlandic (free N, bound "V"), "lexical affixes" as in Spokane (bound "N", free V), and "noun stripping" (no accentual unit, cf. GERDTS 1996:93) as in Polynesian and Micronesian languages.

Nivkh object-verb synthesis is different from noun incorporation in three important respects. Firstly, because unlike "standard" noun incorporation which picks out direct objects over indirect objects, it integrates recipients, as shown above. Secondly, the lexical range of possible "incorporees" is much broader than is usually the case in incorporating languages.

Interrogatives (PANFILOV 1965:23-25) and even proper names (see JAKOBSON 1971:90, KREJNOVIČ 1934:192, 193) synthesize with a verb, as is evidenced by morphophonemic alternations (contrary to what MITHUN 1984:864 and SPENCER 1995:453 observe to occur with noun incorporation):

(18) a. *aŋ χa-ɻ*
who shoot-IND
'who shot at it?'

vs. b. *if aŋ-qʿa-ɻ*
3s who-shoot-IND
'who did he shoot at?'

(19) a. *Xor za-ñivx*
X. hit-person
'the man Xor beat'

vs. b. *Xor-ɻa-ñivx* [4]
X.-hit-person
'the man who beat Xor'

Thirdly, it is normal for incorporated nouns to lack inflection (cf. MITHUN 1984). But in Nivkh nouns synthesizing with verbs can be marked for (i) plural or comitative (which can be combined with case markers but not with number markers and thus behaves rather like number than like case) and (ii) suffixes expressing concepts parallel to KÖNIG's (1991) "focus particles", i.e. 'only', 'even' etc. Again, morphophonemic alternations prove the synthesis, compare *ʧif ~ -zif ~ -ɻif* 'path, track', *indəɻ ~ -nřəɻ* 's.o. sees s.th.', and *ñivx ~ ñivγ-* 'person', *-ti ~ -ri ~ -di* 'even', *iγɻ ~ -kʿuɻ ~ -xuɻ* 's.o. kills s.th.' with the forms printed in **bold** type in the examples below (*ñivx* and *-ti* trigger nasal alternation):

(20) (SAVEL'EVA & TAKSAMI 1970:527)
*ñi ŋaqr-ux kʿeq-xo həjk-xo-**zif**-ku-**nřə**-ɻ*
1s snow-ABL fox-DCOM hare-DCOM-track-PL-see-IND
'in the snow I saw the tracks of foxes and hares'

4 The alternation reflects an elided final nasal (see JAKOBSON 1971:90).

(21) (PANFILOV 1965:75)
*iγlulaq̇ **ñivγ-di-k'uxu-q̇***
tiger person-even-kill:RED-IND
'a tiger even kills people'

Case or (highlighting or question) focus inflection is prohibited on synthesizing nouns, however. Note that with respect to inflection, Nivkh is clearly different from noun-stripping languages, that it would resemble if one failed to acknowledge the accentual unity of dependent-head constructions.

At any rate, the object relation is encoded by an extensive dependent-head synthesis evidenced by morphophonemic alternations signalling the syntactic relation in the absence of central case-marking.

5.2. Verb root serialization

Dependent-head synthesis is not only operative in the case of a government relation, but also in the case of a modificational relation between two verb roots, with the preceding, uninflected one modifying the following, fully inflectable one.

(22) (PANFILOV 1965:75)
*pal-uin hās **vaχʃvaχʃ-eʁ-q̇***
mountain-LOC clothes tear.tear-fast-IND
'clothes tear quickly in the mountain woods'

(23) (KREJNOVIČ 1934:220)
*tə-**G**an **ve-ur-q̇***
this-dog run-good-IND
'this dog runs well'

(24) (PANFILOV 1965:31)
*ñ-mu ʃ-mu-**ək** **mγe-hiñγəj-q̇-ra***
1-boat 2s-boat-COMP row-easy-IND-FOC
'my boat is easier to row than yours'

Note that the modifying elements are not adverbs but roots which can form free verbs. The few existing Nivkh adverbs form a closed class and never coalesce to anything. The exemplified construction is in fact a serialization of verbal roots (cf. BISANG 1995:177) with each root contributing its full lexical meaning. Verb root serialization is rarer than object-verb synthesis, and more often than not an analytic construction containing an adverbial is encountered:

(25) (BEFFA 1982:88-89)
hoʁor *eʁ-gu-r* *məɣ-ɻ*
therefore fast-CST-CV:3s descend-IND
'therefore he descended quickly'

Verb root serialization and its analytic counterpart in Nivkh are fairly parallel to the Chukchi situation:

(26) a. verb root serialization
CHU (SPENCER 1995:455)
nə-tur-tejkə-qinet *nelgə-t*
3p.S-new-make-PRSII:3p>3p.O skin-PL
'they are making skins again'

b. analytic
(SPENCER 1995:455)
nə-tur-ew *nə-tejk-əkinet* *nelgə-t*
CST-new-CST 3p.S-make-PRSII:3p>3p.O skin-PL
'they are making skins again'

The same principle of verb root serialization is exploited for the synthesis of verb root and modal verb root. In this case the lexical root precedes the modal verb root and can be inflected for causative and/or tense/aspectoid categories (cf. KREJNOVIČ 1934:217-218; contrary to what BAKER 1996:31-32 predicts).

(27) (PANFILOV 1965:162)
kʻe-vərk-ku-i-gər-t *seu-tʻa-ɻ*[5]
net-rot-CST-PRS-unwilling-CV:pl dry-HAB-IND
'not wanting the net to rot they dry it'
kerɻ 's.o. is unwilling'

(28) (PANFILOV 1965:161)
imŋ *əyřagin* *tol-dox* *məɣ-ksu-ɻ-ɣu*
3p never water-ALL descend-not_know-IND-PL
'they have never gone down to the water'

Note that verb roots used as modal verbs exist as free verbs as well:

5 Both word forms are written as single words by PANFILOV.

(29) (PANFILOV 1965:71)
hə-ñivx p'-otɬla-əŋzu-γət-ra
that-person REFL-child-not_know-CPL-HILI
'that person did not know his child at all'
free form: *jəŋzuq̣* 's.o. does not know s.th.'

Some of the modal verbs are construed in this way; others are used with a complement clause (cf. PANFILOV 1965:105, 162-163). Finally, reduplication in Nivkh takes the form of root or stem serialization. Incidentally, all reduplicative forms are written in one word by PANFILOV (1965:74pp).

(30) (PANFILOV 1965:78)
qan-gu ʃo-ñi-γət-ñi-γət-ʃ
dog-PL fish-eat-CPL-eat-CPL-IND
'the dogs always eat up the fish'

In all cases of root/stem serialization mood inflection is only possible on the last root or stem, not within the synthetic unit.

6. Is Nivkh a polysynthetic language?

Object-verb synthesis and the integration of adverbial and modal elements feature among the "ingredients" of polysynthetic languages, such as Greenlandic, Mohawk, Ket, etc. So it is worthwhile to check whether Nivkh fulfills any other polysynthesis criteria.

Although there is no ultimate definition of polysynthesis to date, there is a largely agreed upon set of criteria (MITHUN 1988, FORTESCUE 1994, DROSSARD 1997), according to which a polysynthetic language is basically characterized by

(i) "word-sentences", i.e. complex word forms which translate whole sentences in other languages because they comprise encoding of a predicate's arguments (polypersonalism), adverbial concepts, and what would be auxiliaries in more analytic languages, and optionally noun incorporation

(ii) complex verb forms featuring a high number of morphemes or potential slots per word

(iii) morphology playing a prominent role due to the existence of a fair amount of bound morphemes with rather concrete ("lexical") meanings, and a significant entanglement of derivation and inflection.

In addition to what was presented above, a Nivkh verb complex has a range of potential slots, most of which need not be (and usually are not) filled (cf. FORTESCUE 1994:2601):

(31) (primary object)-(serial slot-(causative)-(temporal/aspectoid suffix))-verb root-(TR)-(aspectoid suffix)-(causative)-(temporal/aspectoid suffix)-(modality)-mood/focus/converbal suffix-((PL)-("focus particles"/highlighting focus)) (auxiliary)

A fairly complex form as far as Nivkh circumstances are concerned features in the example below.

(32) (PANFILOV 1965:78)

ţi	*maŋg-ʁař*	*jāř*	*lax*	*pʿ-əřp-γət-ku-řa-ḑ-ŋa*
2s	strong-CNE	why	cloud	REFL-wrap-CPL-CST-HAB-IND-Q

'if you are strong, why do you let clouds wrap you up?'

Besides phases of states of affairs and modal concepts (as exemplified in section 5.2), locational elements (which are bound morphemes) are encoded within a verb form, as would be expected of a polysynthetic language:

(33) (KREJNOVIČ 1960:86)
mu-he-ţ-ţ
boat-water→riverside-pull-IND
's.o. pulls a boat up onto the riverside (out of the water)'
free form: *jeţţ* 's.o. pulls s.th. out of the water'

(34) (KREJNOVIČ 1986:160)
mu tʿas-kʿu-ḑ
boat where-located-IND
'where is the boat?'
xuḑ ~ *-kʿuḑ* 's.th. is located somewhere'

Finally, productive word formation operates on the basis of recursive chaining of nominal and verbal roots (with the products written as single words by PANFILOV 1965).

(35) a. (KREJNOVIČ 1934:196)
ŋa-ŋəŋ-ñivx
beast-hunt-person
'hunter'

b. (JAKOBSON 1971:80)
pəi-mu-meñ-vo-ñivx
fly-boat-helm-take-person
'pilot'

c. (SAVEL'EVA & TAKSAMI 1970:170)
pəi-urla-ləx
fly-good-weather
'good weather for flights'

An entanglement of derivation and inflection is visible particularly in the systematic ambiguity of the nominalizing and "indicative" *ɻ*-form. On the one hand, the morpheme *-ɻ* on verbs makes them compatible with all nominal inflection. This mechanism is fully productive, though a number of these nominalized forms have become lexicalized (cf. PANFILOV 1962:69).

(36)	*nəɻ*	'do'; 'doing'	*nəɻ*	'deed, work'	
	humɻ	'be there, live'; 'living'	*humɻ*	'life'	
	lerɻ	'play'; 'playing'	*lerɻ*	'game'	
	ʃolaɻ	'poor'; 'being poor'	*ʃolaɻ*	'poverty'	
	iññ	'eat'; 'eating'	*iññ*	'food'	
	muɻ	'die'; 'dying'	*muɻ*	'dead person'	
	qalaɻ	'green; envy'; 'green'	*qalaɻ*	'vegetables, greenery'	etc.

On the other hand, the *ɻ*-form of verbs is the most frequently occurring matrix predicate form of any verb in a paradigm with other mood forms, and is therefore glossed IND here. It is still compatible with nominal number and focus suffixes, but it is not case-marked. In contrast to nominal predicates it can bear additional modal markers (like *-rej* indicating mocking surprise of the speaker).

(37) (BEFFA 1982:82)
if vi-ɻ-rej
3s go-IND-REJ
'he is gone' (in a mocking surprise)

This systematic ambiguity of a *ɻ*-form leads to the possibility of various interpretations for one and the same construction (cf. PANFILOV 1965:67-68):

(38) *if hum-ɻ-həjm-ɻ*
3s be_there-IND-know-IND

a.	's/he knows life'	(*humɻ* 'life')
b.	's/he knows the one who is there'	(*humɻ* 's.o. is there')
c.	's.o. knows that s/he lives, is there'	(*if humɻ* 's/he lives, is there')

Finally, demonstratives are formed from deictic roots + *-ɻ* and are inflected like nouns: *hə-* 'that', *həɻ* 'that one', etc.

Up to this point, Nivkh seems to resemble polysynthetic languages to a sufficient degree. But the range of concepts encoded within a polysynthetic verb form, namely manner, degree, chronology, quantification, etc. (see MITHUN (1988:443), FORTESCUE (1994:2602), and DROSSARD (1997)), is not equalled in Nivkh. Nor does Nivkh have classificatory elements (for type of action or type of participants) like e.g. Iroquoian. The most intriguing problem is, however, that Nivkh, strictly speaking, lacks polypersonalism. As we saw in section 5.1, the object is obligatorily encoded on the verb, but the finite mood forms are not marked for their subject (as in (38)), except in the imperative. The subject is marked, however, on two classes of verb forms:

(i) on the most frequent matrix form, the *ɻ*-form, in the form of (non-obligatory) number-of-referent marking

(ii) on one class of converbs through person marking.

The most frequent matrix predicate form, the "indicative" *-ɻ*-form, usually bears the nominal plural suffix if its subject is human plural, thus cross-referencing the subject. The plural suffix on the predicate is not grammaticalized agreement, however.

(39) (PANFILOV 1962:192)
tə-ñivγ-gu *mu-tʻom-γir* *ʧo-ŋəŋ-ɻ-γu* *həɻ-γu*
this-person-PL boat-5_CLF-INS fish-hunt-IND-PL that_one-PL

ŋamk-γir *ʧo-ŋəŋ-ɻ-γu*
seven-INS fish-hunt-IND-PL
'these people went fishing in five boats, those went fishing in seven'

Nivkh converbs form two classes: those which encode their subject and those which do not. The subject marker is *-t* or *-ta* (depending on the converb) cross-referencing 1s, 1p, 2p, and 3p, and *-r* or *-ra* cross-referencing 2s and 3s in the Amur dialect.

(40) (cf. TAKSAMI & POLET'EVA 1992:50)

ñi mɣe-ta,	*ʃi*	*meñ-vo-ra*	*ha-ja*
1s row-ENU:1s	2s	helm-take-ENU:2s	be_so-IMP.s

'let me row and you take the helm'

The enumerative converb, which is exemplified in (40) together with a form of the dummy matrix predicate *haq* 'be so' (inflected for mood) at the end of the clause chain, is more often than not used without this dummy and as such looks like an independent mood form (see (41), after which Beffa puts a full stop).

(41) (BEFFA 1982:86)

hoŋguʃumke	*aʃik*	*ərk*	*pil-ra*
thus_living	younger_brother	already	big-ENU:3s

pal-rox	*mər-ra*	*ʃolŋaj-xu-ra*	*tʿoχ-kʿu-ra*
mountain-ALL	go_up-ENU:3s	stag-kill-ENU:3s	elk-kill-ENU:3s

qʿotr-kʿu-ra
bear-kill-ENU:3s
'living thus the brother was already big and he went up into the mountain woods and killed stags, elks and bears'

Now, word forms like *ʃo-ŋəŋ-q-ɣu* 'they hunt fish' in (39) or *ʃolŋaj-xu-ra* 'he killed a stag and ...' in (41), with both object-verb synthesis and a subject marker, look quite polypersonal. What is more, the object and subject markers together form a bracket around the whole construction, contributing to its wordhood. Maybe we are witnessing polypersonalism in statu nascendi in Nivkh.

But even with polypersonalism being absent, a language can be polysynthetic (in the tradition of Boas 1911:75). Whereas MITHUN (1988:446-447), BAKER (1996:14), and FORTESCUE (1994:2601) more or less require a polysynthetic language to be head-marking or double-marking, DROSSARD (1997:252) allows for what he calls "non-sentential" verb complexes as in languages like Haida or Awtuw. In these languages, the verbal word does not constitute a minimal sentence and lacks participant cross-referencing (participants have to be encoded at least in the form of free pronouns). Nivkh is somewhere in between both extremes, as the verbal word does constitute a minimal sentence while being neither polypersonal, nor requiring overt expression of participants. Thus, the polypersonalism criterion is itself questionable and does not exclude Nivkh from polysynthetic languages.

At any rate, polysynthesis is understood to be a phenomenon rotating around a verbal complex only. In Nivkh, however, there is dependent-head synthesis with both verbal and nominal heads.

7. Nivkh nominal complexes

While ideas on polysynthetic verb complexes are fairly congruent, there is no unanimity as to what nominals should be like in polysynthetic languages. Mithun, working on American Indian languages, points out that nouns are generally not inflected for number, gender, or case, and that there are no obligatory NPs or number agreement (1988:444). BAKER, restricting the domain of polysynthesis considerably (and excluding languages such as Greenlandic; 1996:18-20), adds that word order is free and that there are no true quantifiers or determiners in polysynthetic languages. NICHOLS (1986:105) explicitly states that "there are no polysynthetic nouns", relegating complex noun forms in languages like Eskimo, Algonkin, Sioux, and Iroquoian to word formation, and considering the "adjective-noun" relation to be "least prone to be head-marked" (1986:76). FORTESCUE, on the other hand, recognizes incorporation of adjectival stems into nouns and marking of the possessor on nouns as polysynthesis criteria (1994:2601), particularly with Greenlandic in mind.

The structure of Nivkh nominal forms is closer to what FORTESCUE describes, but does not involve bound morphemes with fairly concrete "adjectival" meaning as Greenlandic does, e.g. *-(r)suaq-* 'big' cannot form a word on its own; see FORTESCUE (1984:297, 317-318):

(42) (FORTESCUE 1984:31)

GRE	*siursus-suaq*	*tassa*	*kisi-mi*
	rushing_sound-big	COP:that_is	alone-POSSR:4s
	'A great rushing sound, that alone (was heard).'		

Instead, Nivkh constructions involve roots which, otherwise, can form word forms on their own. These lexical roots enter into dependent-head synthetic forms with a noun being the head of the construction. Evidence for their synthesis lies in the fact that at the morpheme boundaries of adjacent determiner and noun, possessor and possessum, quantifier and noun, or modifying nominal or verbal root and noun the same morphophonemic alternations occur which we observed with verbal roots as heads, and in the fact that the synthetic forms are accentual units (see (5)) which are inflected as wholes and in which the order of elements is rigid. There is no possessor or attribute marking. In other words, there is an equally strong dependent-head synthesis in verbal and nominal complexes. AUSTERLITZ (1959:103) proves that there is indeed a difference between one or two accentual units with nominal heads:

(43) a. (possessive construction: one accentual unit)
ñ-əɬx
1s-old_man
'my old man'

b. (appositive construction: two accentual units)
ñi əɬx
1s old_man
'I, the old man'

Examples illustrating synthesis with various elements follow:

(44) synthesis of deictic root (determiner) - noun
(KREJNOVIČ 1966:39)
hə-dəf[6]
that-house
'that house'
free form: *təf* 'house'

(45) synthesis of possessor prefix - possessum
(KREJNOVIČ 1934:205)
free form: *ʃaqo* 'knife'
ñ-zaqo 'my knife'
ʃ́-saqo 'your knife'
pʿ-saqo 'one's own knife'
i-ʂaqo 'his/her knife'

(46) synthesis of non-singular pronoun, demonstrative pronoun possessor - possessum
a. *ñəŋ-ʂaqo* 'our knife'

b. (KREJNOVIČ 1966:39)
həʂ-rəf
that_one-house
'the house of that person'
free form: *təf* 'house'

6 Alternation is induced by a final elided nasal of the deictic element, compare the Sakhalin dialect variant with the nasal still present in (53).

(47) synthesis of interrogative - noun
(PANFILOV 1962:254)
təɣd *ǩaɣ-rəv-ŋa*
this_one who-house-Q
'whose house is this?'
free form: *təɣ* 'house'

(48) synthesis of proper name possessor - head noun
(KREJNOVIČ 1934:193)
free form: *təɣ* 'house'
Sarat-rəɣ 'Sarat's house'
Seus-təɣ 'Seus' house'

(49) synthesis of numeral - noun
(KREJNOVIČ 1934:204)
ŋax-kumusk
six-rouble
'six roubles'

(50) synthesis of quantifier - noun
(OTAINA 1978:61-62)
huin *səɣk-ru-γu* *veu-ɣ*
there all-lake-PL deep-IND
'all the lakes there are deep'
free form: *tu* 'lake'

(51) synthesis of modifying noun stem - head noun
(KREJNOVIČ 1958:30)
free form: *ʈif* 'path, track'
qʿotr-ʈif 'bear track'
kʿeq-zif 'fox track'
qan-ɖif 'dog track'

(52) synthesis of modifying verb stem - head noun
(JAKOBSON 1971:96)
Nosk *urla-ɢan-da*
N. good-dog-HILI
'Nosk is a good dog'
free form: *qan* 'dog', allomorphy of HILI see (3)

Note that the verb stem is perfectly parallel to the one used in verb stem serialization. Synthetic forms can be even more complex by comprising several of the above determiners and modifiers.

(53) (JAKOBSON 1971:80)
hun-tleulan-ʃiř[7]
that-white-hill
'that white hill'

(54) (KREJNOVIČ 1958:31)
ñ-γəla-gəlmr
1s-long-board
'my long board'
free forms: *kəlq* 's.th. is long', *kəlmr* 'board'

(55) (KREJNOVIČ 1934:194)
pˁ-loʃi-bitγə
REFL-Russian-book
'self's Russian book'
free form: *pitγə* 'book'

(56) (PANFILOV 1965:80)
hə-ŋamg-ñivx
that-seven-person
'those seven people'

(57) (KREJNOVIČ 1934:193)
ñ-ətək-rəf
1s-father-house
'my father's house'
free forms: *təf* 'house'

From the above evidence we can conclude that a nominal complex in Nivkh looks like this:

(58) (determiner/possessor^{n})-(quantifier)-(attribute)-nominal root-(number)-case-("focus particles"/highlighting or question focus)

7 This is an example from a Sakhalin dialect, in which there are no voiced plosives.

Dependent-head synthesis is obligatory with nominal heads, i.e. it is not possible to rephrase the above constructions analytically.

Nivkh dependent-head synthesis results in "word-phrases", i.e. complexes corresponding to phrases in more analytic languages, rather than in "word-sentences" (although they do occur) as polysynthesis does. A nominal synthetic complex as presented in this section can form the dependent of an object-verb synthetic complex as in the following examples (observe morphophonemic alternations at the boundaries):

(59) (JAKOBSON 1971:79)
ñəŋ-asq-rəu-ja
1p.EXCL-younger_sibling-teach-IMP.s
'teach our younger sibling!'
rəuɻ ~ -təuɻ ~ -dəuɻ 's.o. teaches s.o.'

(60) (PANFILOV 1968:430)
tʻa tə-qʻav-la-ʃax-ta-ja!
PROH this-hot-LA-water-drink-IMP.s
'don't drink this hot water'
ʃax ~ -sax 'water', *raɻ ~ -taɻ ~ -daɻ* 's.o. drinks s.th.'

(61) (KREJNOVIČ 1934:194)
ʃ-vila-ɢan eʁa-ɻuz-ñi-ɻ
2s-big-dog cow-meat-eat-IND
'your big dog ate (the) beef'
free forms: *pilɻ* 's.th. is big', *qan* 'dog', *eʁa* 'cow', *ɻus* 'meat', *iñɻ* 's.o. eats s.th.'

(62) (PANFILOV 1965:77)
ñi mu-lət-ivu-ñivx-nřə-ɻ
1s boat-make-PROG-person-see-IND
'I saw a man building a boat'
indəɻ ~ -nřəɻ 's.o. sees s.th.'

As is evident from the last example, this construction type also comprises what would be relative clauses (also complement clauses) in more analytic languages. In this, Nivkh again differs from noun-stripping languages, in which modification and determination of a stripped ("incorporated") noun are impossible (cf. ROSEN 1989).

Dependent-head synthesis as it occurs in Nivkh has not been dealt with in detail to date. It is not, however, a unique trait of Nivkh.

8. Dependent-head synthesis in the Far Northeast

A syntactic organization very parallel to Nivkh is found in at least two other languages, which are not too distant from Nivkh geographically, namely Chukchi and Ainu. Chukchi is a member of the Chukcho-Kamchadal family which is also subsumed under Paleosiberian languages but not actually genetically related to Nivkh; Ainu is a language isolate formerly spoken on Hokkaido, the Kuril islands, and Sakhalin. At least Nivkh and Ainu have been in contact for some while in the course of their history. Chukchi and Ainu are considered polysynthetic (see especially Shibatani 1990:17 for Ainu). Both languages manifest polypersonal complex verb forms with noun incorporation, verb root serialization, and integration of adverbials and various grammatical categories (see SPENCER 1995:444-448, 455-459; SHIBATANI 1990:60-76, 80; BAKER 1996:37 (fn. 26)), as well as complex noun forms, which we will concentrate on here.

In Chukchi, a determiner, a possessive pronoun, a *wh*-word, a quantifier, and a noun or verb root can be incorporated recursively into a nominal word form, and this mechanism is described as even more productive than noun incorporation by SPENCER (1995:477, 479-480). Such a nominal complex is subject to vowel harmony (grossly low vowels of "dominant" morphemes triggering harmony of high ones, cf. SPENCER 1995:445) in the same way as a noun-incorporating verbal complex (SPENCER 1995:478), and is inflected as a whole, with inflectional circumfixes ultimately proving the wordhood of both the nominal and the verbal complexes.

(63) integration of determiner
CHU (SPENCER 1995:479)
ga-ŋoten-təlacʔə-ma
COM-this-motor-COM
'with this motor'

(64) integration of possessor
CHU (SPENCER 1995:479)
morəg-klassə-k
our-class-LOC
'in our class'

(65) integration of numeral
CHU (SPENCER 1995:479)
ga-ŋeran-wag-ma
COM-two-claw-COM
'with two claws'

(66) integration of *wh*-word
CHU (SPENCER 1995:477)
req-upicgən
what-pole
'what pole'

(67) integration of nominal modifying root
CHU (SPENCER 1995:477)
ŋinqej-kʔeli
boy-cap
'a boy's cap'

(68) integration of verbal modifying root
CHU (SPENCER 1995:478)
ga-taŋ-pojgə-ma
COM-good-spear-COM
'with a good spear'

(69) recursive integration of modifying roots
CHU (SPENCER 1995:480)
ga-taŋ-tor-kətepa-nalgə-ma
COM-good-new-ram-skin-COM
'with a good, new, ram's skin'

Corresponding analytic phrases can be construed by deriving regular relational forms of modifying nouns, *wh*-words, and verbs (SPENCER 1995:477):

(70) (SKORIK 1961, I:427)
a. N-complex
arm-aʔaçek
strong-youth
'a strong young man'

b. paraphrase
n-erme-qin *aʔaçek*
3s.S-strong-PRSII:3s youth
'a strong young man'

In Ainu, at least possessors, quantifiers, and modifying noun or verb roots are integrated into a nominal complex and can be recursive. Evidence for wordhood is of an accentual and a morphophonemic nature: assimilation and dissimilation phenomena of final and initial sounds, a reduced form of vowel harmony, and the elimination of a hiatus (SHIBATANI 1990: 74-75, 13).

(71) (SHIBATANI 1990:73)
AINU *ku-pon-turesi*
1s.S-little-younger.sister
'my little sister'

(72) (SHIBATANI 1990:73)
AINU *k-arka-sikihi*
1s.POSSR-hurt-eye
'my hurting eye'

(73) (SHIBATANI 1990:72)
AINU *usa-oruspe*
various-rumor
'various rumors'

(74) (SHIBATANI 1990:74)
AINU *e-pon-no-poro-setaha*
2s.POSSR-slight-ADV-big-dog
'your slightly big dog'

Analytic constructions are generally possible and manifest a different order of elements and no limitation on the number of modifiers (SHIBATANI 1990:74), thus a single word complex and a phrase are distinguished.

(75) (SHIBATANI 1990:68)
AINU *pon akor sapo a-at-e-uk*
young 1s.POSSR sister PASS-rope-APPL-tie
'my young sister was tied up with a rope'

(76) (SHIBATANI 1990:74)
AINU *ponno poro e-setaha*
slightly big 2s.POSSR-dog
'your slightly big dog'

Except for the possibilities of analytic constructions, Chukchi and Ainu phrasal organizations are parallel to Nivkh. For Chukchi, furthermore, there are examples of a whole nominal complex being incorporated into a verbal complex (SPENCER 1995:480). Here again, inflectional circumfixes and vowel harmony constitute evidence for the wordhood of this construction.

(77) (SKORIK 1961, I:102)
CHU *tə-ŋəron-qora-kənʔo-rkən*
1sS-three-stag-catch-PRSI:1s>3s
'I'll catch three stags'

(78) (SPENCER 1995:480)
CHU *tə-tor-taŋ-pəlwəntə-pojgə-pela-rkən*
1sS-new-good-metal-spear-leave-PRSI:1s>3s
'I am leaving a good, new, metal spear'

In Ainu, only bare noun roots can be incorporated (SHIBATANI 1990:71), though a noun root and an adverbially modifying root can be incorporated at the same time (SHIBATANI 1990: 72).

(79) (SHIBATANI 1990:72)
AINU *pinne kamuy kiraw-riki-kur-roski*
male god horn-high-EXPL-raise
'the male (dragon) god raised the horns high'

All in all, with noun incorporation, verb root serialization, and complex noun forms Chukchi and Ainu display (to different extents) a range of syntactic parallels to Nivkh.

9. Typological classification

In conclusion, we have observed dependent-head synthesis as the basic principle of Nivkh syntactic organization. Dependent and head are adjacent morphemes at whose common boundary morphophonemic processes are operative, with the final and initial consonants of morphemes being different in absolute final or initial position and in bound position, respectively. The synthetic forms are accentual units inflected as wholes. Within the complexes only a limited number of inflectional categories can be specified. Dependent-head synthesis is operative in government and modificational relations between a head and a dependent, i.e. in the following relations:

dependent	**head**
primary object	verb
adverbial modifier (verb root)	verb
determiner	noun
quantifier	noun
possessor	noun
modifying noun root	noun
modifying verb root	noun

Two distinct dependent-head synthetic complexes can be distinguished on the basis of their heads and slots, each of which offers slots in a fixed order for dependents in two different relations to the head:

(80) a. object-serial slot-V-TAM inflection
b. determiner/possessor-quantifier-attribute-N-number/case/focus inflection

The last lexical root in a synthetic form is the head of the construction and determines the type of inflection. Structurally, the dependent-head complexes correspond to phrases in more analytic languages. Since the subject is not consistently integrated into dependent-head synthesis (i.e. there is no fully grammaticalized polypersonalism), the result are "word-phrases" rather than "word-sentences". From the point of view of function, both synthetic constructions in Nivkh serve equally well as predicates and as arguments.[8]

The synthesis of heads and dependents in Nivkh signals the existence of a dependent-head relation without specifying the kind of relation. Within NICHOLS' typology of head/dependent-marking, Nivkh must be classified as head-marking, as are noun incorporation and polysynthesis in general. The synthesis is head-marking in the sense that the dependent is fully specified on the head. NICHOLS herself recognizes a "head-marked pattern" in connection with "sandhi" in Nivkh (= Gilyak) (1986:100), but overlooks the importance of this pattern and thus considers the language "largely isolating" on the whole. Even disregarding the phenomenon discussed here, Nivkh has a considerable amount of morphology and is more agglutinating than isolating.

So Nivkh manifests a head-marked pattern and a high degree of synthesis, with a range of phenomena reminiscent of polysynthesis. Is there a relation between this type of dependent-head synthesis and polysynthesis?

The common denominator of all languages which have been described as polysynthetic is a complexity of verb forms with integration of material that forms separate constructions in more analytic languages. Apart from this, the languages exhibit quite different sets of

8 Converbs and verbs in mood forms (except the *d*-form) serve as predicates only.

traits which apply out of the pool of polysynthesis ingredients (compare typologies of DROSSARD 1997 and FORTESCUE 1994), some features being more salient than others. In this sense, BAKER (1996:338) observes that languages with rich noun incorporation are "poor(er) in other complex predicates" (e.g. Mohawk) and vice versa (e.g. Yimas, which has verb root serialization but no noun incorporation). Viewed from this angle, Nivkh could be specialized in yet another way. As we have seen, it does not stand alone, but shares a considerable number of organizational characteristics with Chukchi and Ainu, two languages of its wider geographical region. The three languages manifest (to different degrees):

(i) a dependent-head noun-verb synthesis close to "compound noun incorporation" in the sense of ROSEN (1989)
(ii) complex forms around a nominal head, with synthetization of modifiers and determiners
(iii) synthetization of complex noun forms with verbs
(iv) synthetization of verbs with adverbial and modal elements via verb root serialization.

In contrast especially to some Native American polysynthetic languages they exhibit:

(v) a moderate number of potential or filled verb slots
(vi) a low number and small variety of bound morphemes (vs. Greenlandic)
(vii) an absence of lexical affixes (bound (not lexical) elements characterizing the type of participant involved) (vs. Mohawk).

Chukchi and Ainu have been classified as polysynthetic languages elsewhere, and the difference from Nivkh seems to be rather one of degree than of kind. Besides, Nivkh fits into the polysynthesis classification by DROSSARD (1997:262):

There is integration of adverbial material (which is a necessary and sufficient condition), the verb form is sentential (vs. polysynthetic languages such as Awtuw) though not polypersonal, and there is integration of nominal material (vs. Yimas). Integration of nominal material is in the form of a lexical noun root (not a bound element like in Quileute) incorporated into a lexical verb root (not a bound one like in Eskimo).

Approaching the question of relatedness of various forms of synthesis from another perspective, we could try and establish a typology of factors bringing about a high degree of synthesis in different languages, which would cluster in various ways to be constitutive of polysynthetic, incorporating, or verb root serialization languages. The polysynthesis typology by FORTESCUE (1994:2602) could serve as a basis, which enlarged and rearranged to a classification of synthesis parameters reads as follows:

A high degree of morphological synthesis in a language can come about by way of

A: derivation, taking the form of

(i) a complex verb form characterized by a multitude of fixed slots which need not all be filled at once (including lexical affixes), e.g. Navaho, Haida, Awtuw, Ket

(ii) a complex verb form or noun form resulting from recursive affixing out of a pool of a multitude of bound morphemes with lexical and grammatical content (including denominal verbs), e.g. Eskimo, Kwakwala;

B: composition, taking the form of

a complex verb form or noun form resulting from integration of otherwise lexeme-forming morphemes, with the subtypes

(i) noun incorporation only (Mohawk, Onondaga)

(ii) noun and verb integration into each other (Chukchi, Ainu, Nivkh)
(noun incorporation and verb root serialization)

(iii) verb root serialization only (Yimas)

Languages may exhibit several of these parameters at the same time. For instance, Mohawk has lexical affixes as well as noun incorporation and some verb root serialization (cf. BAKER 1996), Chukchi and Nivkh have noun incorporation/synthesization, verb root serialization, and some denominal verbs (cf. SPENCER 1995:462, MATTISSEN & DROSSARD 1998:64), but Greenlandic sticks to only one parameter. If a language has noun incorporation only or verb root serialization only, it shows synthetic traits but in my view does not qualify as polysynthetic.

All in all, we have observed a syntactic organization in Nivkh which seems intermediate between languages where syntax is more prominent than morphology in phrase/clause constitution (e.g. European languages) and languages where morphology is more prominent than syntax (e.g. Iroquoian languages).

ABBREVIATIONS

A	agent	INS	instrumental
ABL	ablative	LA	morpheme *-la-*
ADV	adverb marker	LOC	locative
ALL	allative	NML	nominalization
APPL	applicative	NOM	nominative
CHU	Chukchi	O	object
CLF	classifier	PASS	passive
CNE	counter-expectative converb	PL	plural
COM	comitative	PO	primary object
COP	copula	POSSR	possessor
CPL	completive	PROG	progressive
CST	causative	PROH	prohibitive
CV	converb	PRS	present
ENU	enumerative converb	Q	question marker
EXCL	exclusive	RED	reduplication
EXPL	expletive	REFL	reflexive
FOC	focus marker	REJ	modal marker
GRE	Greenlandic	REMPST	remote past
HAB	habitual	S	subject
HILI	highlighting focus	TERM	terminative
IND	indicative	TR	transitivizer
IMP	imperative	U	undergoer

REFERENCES

ANDERSEN, HENNING 1986: "Introduction: Sandhi". In: ANDERSEN, HENNING (ed.): *Sandhi Phenomena in the Languages of Europe*. Berlin et al.: Mouton de Gruyter. 1-8

AUSTERLITZ, ROBERT 1959: "Semantic components of pronoun systems: Gilyak". *Word* 15 (1): 102-109

AUSTERLITZ, ROBERT 1990: "Typology in the service of internal reconstruction: Saxalin Nivx". In: Lehmann, Winfred P. (ed.): *Language Typology 1987: Systematic Balance in Language*. Amsterdam: Benjamins. 17-33

BAKER, MARK C. 1988: *Incorporation*. Chicago: Chicago University Press

BAKER, MARK C. 1996: *The Polysynthesis Parameter*. New York etc.: Oxford University Press

BALL, MARTIN J. & MÜLLER, NICOLE 1992: *Mutation in Welsh*. London: Routledge

BEFFA, MARIE-LISE 1982: "Présentation de la Langue Nivx suivie de l'analyse linguistique d'un paragraphe du conte". *Études mongoles et sibériennes* 13: 49-98

BISANG, WALTER 1995: "Verb serialization and converbs - differences and similarities". In: Haspelmath, Martin & König, Ekkehard (eds.): *Converbs in Cross-Linguistic Perspective*. Berlin: Mouton de Gruyter. 137-188

BOAS, Franz 1911: "Introduction". In: Boas, Franz (ed.): *Handbook of the American Indian Languages*. Part 1. Washington: Bureau of American Ethnology. 5-83

BONDARKO, L.V. & ZINDER, L.R. 1962: "Foneticeskaja xarakteristika razlicnyx tipov sintaksiceskix socetanij v nivxskom jazyke". *Voprosy Jazykoznanija (VJa)* 11 (4): 84-89

COMRIE, BERNARD 1981: *Languages of the Soviet Union*. Cambridge: CUP

DROSSARD, WERNER 1997: "Polysynthesis and polysynthetic languages in comparative perspective". In: Palek, B. (ed.): *Proceedings of Linguistics and Phonetics 1996*. Prague: Charles University Press. 251-264

DRYER, MATTHEW S. 1986: "Primary Objects, Secondary Objects, and Antidative". *Language* 62: 808-845

FOLEY, WILLIAM A. 1991: *The Yimas Language of New Guinea*. Stanford: Stanford University Press

FORTESCUE, MICHAEL 1984: *West Greenlandic*. London etc.: Croom Helm

FORTESCUE, MICHAEL 1994: "Morphology, Polysynthetic". In: Asher, R.E. (ed.): *The Encyclopedia of Language and Linguistics*, Vol. 5. Oxford etc.: Pergamon Press. 2600-2602

GERDTS, DONNA B. 1996: "Incorporation". In: Zwicky, Arnold M. & Spencer Andrew J. (eds): *Handbook of Morphology*. Oxford: Blackwell. 84-100

HATTORI TAKESHI 1962: "Versuch einer Phonologie des Südostgiljakischen II - Alternation". *Journal of Hokkaido Gakugei University*, Section I-A, Vol. 13 (2): 29-96

HATTORI TAKESHI 1988: "Giriyâkugo". In: *Gengogaku Daijiten*, Vol. I. Tokyo: Kenkyûsha. 1408-1414

JAKOBSON, ROMAN 1971: "Notes on Gilyak". "Texte guilyak". In: *Selected Writings* II. The Hague: Mouton. 72-102

KÖNIG, EKKEHARD 1991: *The meaning of focus particles*. London etc.: Routledge

KREJNOVIČ, E.A. 1934: "Nivxskij (giljackij) jazyk". In: Al'kor, Ja.P. (ed.): *Jazyki i pis'mennost' narodov Severa* III. Leningrad: Institut Narodov Severa (= Trudy po Lingvistike, Vol. 3). 181-222

KREJNOVIČ, E.A. 1937: *Fonetika nivxskogo (giljackogo) jazyka*. Moscow-Leningrad: Institut Narodov Severa (= Trudy po Lingvistike, Vol. 5)

KREJNOVIČ, E.A. 1958: "Ob inkorporirovanii v nixvskom jazyke". *VJa* 7 (6): 21-33

KREJNOVIČ, E.A. 1960: "Vyraženie prostranstvennoj orientacii v nivxskom jazyke (k istorii orientacii v prostranstve)". *VJa* 9 (1): 78-89

KREJNOVIČ, E.A. 1966: "Ob inkorporirovanii i primykanii v nixvskom jazyke". *VJa* 15 (3): 33-51

KREJNOVIČ, E.A. 1986: "Ob imenax prostranstvennoj orientacii v nixvskom jazyke". In: *Paleoaziatskie jazyki*. sbornik naučnyx trudov. Leningrad: Nauka. 157-167

MATTISSEN, JOHANNA & DROSSARD, WERNER 1998: *Lexical and syntactic categories in Nivkh (Gilyak)*. Theorie des Lexikons, Arbeiten des Sonderforschungsbereichs 282 Nr. 85. Düsseldorf: Heinrich-Heine-Universität

MITHUN, MARIANNE 1984: "The evolution of noun incorporation". *Language* 60: 847-893

MITHUN, MARIANNE 1988: "System-defining Structural Properties in Polysynthetic Languages". *ZPSK* 41 (4): 442-452

NEDJALKOV, V.P. & OTAINA, G.A. 1988: "Resultative and Continuative in Nivkh". In: Nedjalkov, V.P. & Comrie, B.: *Typology of Resultative Constructions*. Amsterdam: Benjamins. 135-151

NESPOR, MARINA & VOGEL, IRENE 1986: *Prosodic Phonology*. Dordrecht: Foris

NICHOLS, JOHANNA 1986: "Head-Marking and Dependent-Marking Grammar". *Language* 62: 56-119

OTAINA, G.A. 1978: *Kačestvennye glagoly v nivxskom jazyke.* Moscow: Nauka

PANFILOV, V.Z. 1954: "K voprosu ob inkorporirovanii. Na materialax nivxskogo (giljackogo) jazyka". *VJa* 3 (6): 6-27

PANFILOV, V.Z. 1958: "Složnye suščestvitel'nye v nivxskom jazyke i ix otličie ot slovosočetanij (k probleme slova)". *VJa* 7 (1): 105-111

PANFILOV, V.Z. 1960: "Problema slova i "inkorporirovanie" v nivxskom jazyke". *VJa* 9 (6): 50-58

PANFILOV, V.Z. 1962, 1965: *Grammatika nivxskogo jazyka.* 2 Vols. Moscow-Leningrad: Nauka

PANFILOV, V.Z. 1968: "Nivxskij jazyk". In: *Jazyki Narodov SSSR*, Vol. 5: *Mongol'skie, Tunguso-Man'čžurskie i Paleoaziatskie jazyki.* Leningrad: Nauka. 408-434

ROSEN, SARA THOMAS 1989: "Two Types of Noun Incorporation: A Lexical Analysis". *Language* 65: 294-317

SAVEL'EVA, V.N. & TAKSAMI, Č.M. 1970: *Nivxsko-Russkij Slovar'.* Moscow: Sovetskaja Enciklopedija

SHIBATANI MASAYOSHI 1990: *The Languages of Japan.* Chicago: CUP

SKORIK, PJOTR 1961: *Grammatika čukotskogo jazyka.* Vol. I. Moscow: Izdatel'stvo Akademii Nauk

SPENCER, ANDREW 1995: "Incorporation in Chukchi". *Language* 71: 439-489

SPENCER, ANDREW 1996: *Phonology.* Oxford: Blackwell

TAKSAMI, Č.M. & POLET'EVA, C.F. 1992: *Raduga.* Kniga dlja dopolnitel'nogo čtenija v 1-2 klassax Nivxskix škol. St. Peterburg: Prosveščenie

WATANABE MICHIKO 1983: "Giriyâkugo". *Gengo* 12 (11): 62-69

WATANABE MICHIKO 1992: "Giriyâkugo tadôshibun no tokuchô". In: Miyaoka Osahito (ed.): *Kita no Gengo: Ruikei to Rekishi.* Tokyo: Sanseido. 179-190

Clause combining in Apache[1]

DAGMAR JUNG

1. Introduction

Clause combining in Apachean languages typically consists of paratactic structures combined with weakly embedded hypotactic structures. This characteristic represents an areal trait in North America (MITHUN 1999). The division between paratactic and hypotactic constructions in the Apachean languages parallels the structuring of events into successive v. simultaneous events, respectively (cf. LONGACRE 1985). The focus in this paper is the range of clausal constructions which make use of the enclitic *go* in Apache and which can be shown to be related by the concept of simultaneity.

There are no non-finite verb forms in Apachean, in the sense that every verb form exhibits full inflection for Person and Tense/Aspect. There are no paradigmatic restrictions on the range of inflections occurring in embedded clauses. This lack of non-finite verbal constructions is regarded as a typical characteristic of polysynthetic languages (BAKER 1996, MITHUN 1984). Moreover, the complex structure of the Apachean verbs allows for the expression of additional aspectual and modal meanings that in other languages requires analytic constructions such as auxiliaries plus participles. Example (1) shows a verb in the inceptive aspect, examples (2) and (3) modal expressions[2]:

---

1 Without the expertise on the Jicarilla Apache language of Mrs. Patricia Torivio from Dulce, New Mexico, this work would not have been possible. Melissa Axelrod's comments were very valuable. Examples are taken from the author's fieldwork on Jicarilla Apache (1993-1999), unless otherwise indicated. The second source are HOIJER'S unpublished Jicarilla Apache texts from ca. 1930.

2 Examples are given in the practical orthography. Abbreviations in the interlinear transcription: ADV = adverb, ARE = areal, ASP = aspect, COM = comitative, CONT = continuative, DIR = directional. DIST = distributive, DL = dual, EMPH = emphatic, FUT = future, H = human, INC = inceptive, IND = indefinite, IO = indirect object, ITER = iterative, LEX = lexical prefix, LINK = linker, LOC = locative, NARR = narrative, NEG = negative, NONS = nonsingular, O = object, P = postposition, PFV

(1) *ashíį́'* ***nkee****-nááná-Ø-dzá-ná* *dagi* *ts'ętso'-yé*
then INC-again-3O-walk-NARR up meadow-DIR
'Then he started to walk up to the meadow' (HOIJER 3:1)

(2) *ń-di- gh-Ø-ł-hash*
2sO-LEX-FUT-3S-V/V-bite
'It might bite you'

(3) *dá-'i- i-sh- hásh*
EMPH-INDO-ASP-1sS- sleep
'I have to sleep'

In the following, the four main types of clause combining techniques are described. First, examples of coordination are presented ranging from mere juxtaposition of clauses to discourse structuring connectives and adverbial connectives. Next, in section 3, a type of embedded construction in which postpositions function as subordinators with adverbial meanings is illustrated. In section 4, I describe the nominalizing/relativizing construction which appears in complement structures. Finally, in section 5 I discuss the constructions employing the enclitic *–go*. These constructions include temporal, conditional and causal adverbial clauses, complement clauses, analytic causatives and finally resultative constructions., all of which are related by this enclitic's general function of marking a simultaneous temporal relationship among two or more entities or events.

2. Coordination

Clauses in longer stretches of discourse that relate to a succession of events are most frequently introduced by particles translated as 'then, and then' (cf. also Scollon 1985 on sequencing of events in Chipewyan). Example (4) is representative of the discourse schema in narratives:

(4) *adą* *yi-na-' -iis-Ø-zo*
and then 3O-around-INDO-ASP-3S-make.a.line
'And then he circled it.'
Adą *da'kwéé* *i-naa-y -íį-Ø-'íį*
and then over there DIR-ITER-3O-ASP-3S-give
'And then they took it again.'

= perfective, PL = plural, PROG = progressive, S = subject, SER = seriative aspect, SPEC = specific, V/V = voice/valence.

Much less common is the mere juxtaposition of two coordinate clauses without an overt particle, such as in example (5):

(5) *ądą tl'o' ná-'i- Ø-t'i íi-Ø- ł- ts'e*
then outside back-ASP-3S-walk ASP-3S-V/V-made noise
'And then (the horse) was walking back outside and made noises.'

A clause type that could be expected to be marked is also juxtaposed without further indication of its status: instances of reported or indirect speech are not marked with a complementizer, as can be seen in example (6)[3]:

(6) *ha'daodá a- nn-lé daa-mi-ł- dńń- 'ni*
something LEX-ASP.2SS-make PL-3O-COM-1NONSS- tell

nda doo ha'dao á-i-Ø--lé yį-ká-Ø-t'į
but NEG something LEX-ASP-3S-make 3O-LEX.ASP-3S-want
'we told him to do something about it, but he didn't want to do anything'

Reported speech in Apache cannot be distinguished formally from direct speech. The pronominal marking of the embedded verb and the matrix verb are the same in direct and indirect uses, and there is no complementizer. Example (6) also exemplifies the use of a connective particle *(e)nda* at the beginning of the second line to conjoin the first and second complex clause. The construction of the desiderative verb seen in the second complex clause without a complementizer will be discussed below in section 4.

Connectives relaying adverbial meanings are employed to emphasize the kind of relationship between two events. The use of these connectives is not related to other overt means of subordination, such as word order differences or restricted paradigms. Example (7) shows a temporal conjunction *iskisna* 'until' that restricts the duration of the event expressed in the first clause.

(7) *adą bįįshzhéé-'í yi-ł-há-di-Ø-de dą́,*
then fawn-SPEC 3O-COM-ADV-ASP-3S-chase PST

iskisna' *bįįshzhéé-'í nn-ee-'ł-de*
until fawn-SPEC ADV-ASP-3S-v/v-tired
'and then they were chasing the fawn, until the fawn became tired'

3 This is actually not totally correct, since in actual discourse the speech event can be marked off by intonational pattern for direct speech.

3. Local postpositions as adverbial 'subordinators'

Another construction employed in clause combining is the use of locative postpositions following the verb to form a complex event. The intimate historical connection of adpositions and subordinators has been attested in many languages (cf. e.g. Genetti 1991). This construction is rather infrequent in Jicarilla texts. The first example below shows the ablative postposition *-shį* used in its local meaning. No additional nominalizing morphology is necessary to combine the verb with the postposition/subordinator:

(8) *naa- sh-Ø-zhee-**shį*** *naa-Ø-dáł*
CONT-ASP-3S-hunt -**ABL** back-3S-come.PRG
'he returned from hunting' ('he hunted–from there–he came back')

The next example shows the locative postposition *–ee* with a temporal meaning 'right at the time when':

(9) *adą **dá-** ye 'ii-kaas-**ee*** *iłkįį* *na-i -dnn-ką*
then **EMPH**-inside-3S-come-**LOC** rifle up-3O-LEX.PFV-3S-handle
'And then, **as soon as** he came inside, he picked up a rifle'

The emphatic prefix *dá-* in this example supports the punctual event expressed by the first clause. The directional postposition *–yé* adds a concessive reading in the next example:

(10) *dádoo chish há -ji-ł -'á'-**yé*** *chish g-á -nń-Ø-ghį-ná*
NEG wood LEX-IND.S-V/V-send-**DIR** wood 3O-BEN-ASP-3S-carry -NARR
'Although nobody had sent him for wood, he carried the wood for him.'(GAM:17)

The use of the same directional element in example (11) is ambiguous with respect to a locative and a purposive reading:

(11) *k'adi díí kǫka-híí* *dáxaa'* *chish g-á-n-ch'-ó-ghé -**éé***
now this camp-SPEC anywhere wood INDO-BEN-LEX-INDS -ASP-carry-**LOC**

sh-ee-náá-'-gǫǫ-'ai *'áshdla-di* *'ii-s -'įį -go.*
1sO-P-ITER-INDO-LEX-2SS.ASP-bet five-times INDO-ASP-give-LINK
'Will you wager with me for five points to carry wood to anyone in this camp?'
(GAM:10)

In summary, this section illustrated that the postpositions/subordinators are used as a kind of locative nominalization (as in (8)), but are also extended in their use to form adverbial clauses (although it is not yet clear how their meaning is extended in the various contexts).

4. Nominalization

The next type of construction of embedded clauses concerns the uses of nominalized verbs. The nominalizing suffix *–i* is employed as a marker of relativizing constructions and it may also serve as a marker of complementation by extension of its nominalizing function. The nominalized verb keeps all of its inflection and is combined with the suffix expressing either human 'the one(s), who' in example (12), or non-human 'that one', as in example (13):

(12) *ká -ná -da -' -ts'i-gis -'í*
water.into-ITER-DISTR-INDO-INDS-turn-**NOM**
'people who do their washing'

(13) *béshi-ye'-kǫh'-n-Ø-tsaa-'í*
metal-inside-fire-ASP-3S-big-**NOM**
'stove' (' big fire inside the metal')

Example (13) represents a common type of word formation process in Apache, i.e. lexicalized expressions of nominalized descriptive verbs (here derived from a stative verb *n-tsaa* 'it is big').

The nominalizing suffix may occur at the end of a complex modifying clause to refer back to the head noun, as the following conversational example shows:

(14) *ayí*	*mangáani-'ąą*	*ká -ná -da-' -ts'i-gis -'ee*
DEM	white.person -TOP	water-ITER-DISTR-IND.O-INDS-turn-LOC
that	white person	at where one wrings it (clothes) in the water
sh-ee-di-Ø-nii		*d-ń -ń-ni -í*
1sO-P -LEX-3S-dislike		LEX-PFV-2sS-say -**NOM**
she didn't like me		you said – **the one**

'That white person at the laundry who you said didn't like you?'

In some constructions, the nominalizer can be equally analyzed as a relative marker or a complementizer:

(15) *łį́* *mi-ł-na-dii-Ø-dee' -í* *mi-ye-gǫǫ-Ø-si-ná*
horse 3O-COM-CONT-LEX.ASP-3S-chase-**NOM** 3O-LEX-LEX.PFV-3S-be.mad-NARR
'(The fact) that the horse chased him made him mad'
or: 'The horse that chased him made him mad'

The nominalized verb in (15) can be analyzed as a complex nominal expression in subject function. The noun *łį́* is unmarked toward its status as either head of a relative clause or the subject of an embedded complement clause.

In example (16) the nominalization indicates an embedded question:

(16) *adą́* *didé -'í* *yi-ł-k'é-i-n-o-'ą́-na*
and then people-SPEC 3O-P-LEX-3o-ASP-3S-tell-NARR

ya'hįį́ka *naa -'ash -'í*
for what reason CONT-go.DL-**NOM**
'And then they told the people why they had come/the reason for their coming.'

Finally, the nominalizer is used with verbs of saying if there is a second verb of saying as complement of the first one:

(17) *bįį́shzhéé* *na-i-n-Ø-ł-kį́-'í* *á-yii-ł-ni* *goo-ł-ni*
fawn back-3O-ASP-3S-V/V-bring-**NOM** thus-3O-3S-v/v-say INDO.ASP-3S-V/V-say
'She said to her: she means that they brought a fawn back'

Because of the lack of other subordinating characteristics besides the suffix, the nominalizing construction in Apache can be considered a superficial type of embedding. Still, the interpretation of a complementing structure in (14) implies a categorical change from verbal clause to a subject complement within a complex clause. This construction is the strongest candidate in Apache in the search for embedded clauses.

Historically, the nominalizing suffix is related to the modern nominal suffix *–i* denoting specificity (as in example 18), or in a discourse context identifiability of the noun.

(18) *dii* *chish-í* 'this particular tree'
this tree –**SPEC**

All of these uses probably find their origin in a Proto-Athabaskan demonstrative that cliticized to its preceding unit (SAPIR 1922, JUNG 1995).

5. *go*-Constructions

The most frequent construction used in Apache for clause combing – besides coordinating structures – is the use of a fully inflected verb with the linker *–go*. As is shown below this linker functions primarily in temporal and adverbial expressions, but is also found as a kind of linker in auxiliary and complementizing constructions as well.

5.1. Temporal and other adverbial relations

The adverbial uses of two clauses combined by *–go* are centered around the temporal notion of simultaneity of events. Usually the first clause is marked by *–go* (as in example 19), but it may also occur as second clause (as in example 20).

(19) *łao*	*dibéi-zhą́ą́*	*íi-Ø-ł-hosh-**go***	*k'ą-hi-sh -ei*
sometimes	sheep-small	ASP-3S-V/V-sleep -LINK	across-ASP-1sS-come

'Sometimes, when the lamb was sleeping, I came across it'

(20) *naa-da-' -n-dzoo*	*isgwéela*
CONT-PL –INDO-IMV-herd.out	school
we herded sheep	school

*ha-na-daa-ts'i-is-kai-**go***
out-back-PL-INDS-PFV-go.PL-**LINK**
when one came out of there
'Coming out of school we herded sheep'

In (19) the event denoted by the second clause takes place while the event of the first clause marked by *-go* is still going on. In (20) this is not strictly the case, since the event marked by the linker must have occurred before the first clause. The context of habituality (the described events took place on a daily basis for a long period of time) might facilitate the temporal frame of (20).

When there is a construction of the type [X does Y-*go* X does Z], the interpretation is one of total temporal overlap of the two events. The verb marked by the linker specifies the event expressed by the unmarked verb.

(21) *da'ko*	*anngo*	*g-á-na-i-ní-ł-t'ii-**go***	*á -go-o-ł-ni-néh:*
then	DEM	INDO-P-CONT-3O-ASP-V/V-look-**LNK**	LEX-IND.O-ASP-V/V-say-QUOT

'Looking at it he tells them this:'

(22) *didé* *mi-kee* *y-e-' -hi-naa-Ø-ł -ąął-**go***
man 3O-foot 3O-P-INDO-SER-LEX-3S-V/V-measure-**LINK**

n -a-Ø-'aash
CONT-ASP-3S-go.DL
'They went around measuring men's feet.'

Examples (21) and (22) each consist of two verbs whose subject arguments are coreferential. In (21) 'his looking' and 'his telling' is understood as taking place simultaneously. In (22), the seriative aspect characterizes the 'measuring' event as taking place repeatedly, one after the other. The motion verb basically adds a path to this activity, resulting in one complex overall event. If the subject arguments are non-coreferential, the interpretation of this construction can be one of temporal overlap, but not one of a complex predicate.

A problematic case for the analysis of the linker *–go* can be seen in the next example from HOIJER'S texts. The first clause 'then he put up a pole' is introduced by a sequential connective, *yéé'a* 'then', and is marked as being concluded by the narrative suffix *-ná* that follows the verb. The second verb 'he was thinking' is the one marked with the linker and supposedly expresses an event that overlaps temporally with the following event. But the next clause is again introduced by a sequential particle, that sets off the third clause from the one before:

(23) ***yéé'aa*** *maashk'a'* *ya -i-dn-Ø-tsi-**ná*** *hiní-ne-i-Ø-ł-t'i-**go***
then pole up-3O-LEX-3S-stick-**NARR** LEX-CONT-3O-3S-V/V-think-**LINK**

yéé'aa *dá'yádaa* *doo-ná -' -ó-ɬį-Ø-'aa* *át'é-**ná***
then something NEG-back-INDO-LEX-ASP-3S-wager 3.COP-**NARR**
'Then he put up his pole – he was thinking that there was nothing left he could wager.' (GAM:9)

HOIJER'S translation suggests that the last clause is a complement of 'thinking', although there is no formal sign that this clause is not to be an independent unit. The appropriate analysis of the *go*-construction in (23) is probably to be sought in its status as being embedded in a larger discourse unit rather than analyzing the sequence as well-defined clausal units (cf. further examples of discourse embedding below in section 5.2.).

Besides the temporal uses of the above examples the *go*-construction can take on additional adverbial meaning, extending the temporal meaning to causal and conditional interpretations. In these cases, both, temporal and causal or conditional readings are simultaneously present. A causal extension is exemplified in (24):

(24) *adą́* *nǫ́ǫ́* *mi-ts'ą́* *hi -Ø-kas -o*[4] *yi-ke'-yeh*
and then over there 3O-away ASP-3S-run.SG-**LINK** 3O-after-DIR

i -ɬa-haa-Ø-t'ash *yi-ke'-i-ɬaa-haa-Ø-t'ash*
DIR-COLL-ASP-3S-go.DL 3O-after -DIR -COLL-ASP-3S-go.DL
'And then it was running away from it over there, because they (the dogs) were jumping after it, they were jumping after it'

Here, the result is presented in the first clause marked by the linker, with the second clause presenting the cause.

A conditional reading extended from the temporal reading is frequently found. Here, the linker marks the event which is conceptualized as the condition for the resulting event. In (25) interpretation is ambiguous as to a temporal or conditional reading.

(25) *kǫghąyaame'ɬiayɬí* *mi-ye'* *go-ghą- **go***
teepee 3O-inside INDO-live-**LINK**

géh *ɬts'ą'-ye* *'de-'-daa-ɬ-je*
just middle-in fire-INDO-PL-build
'if you live in a teepee, then you build a fire in the middle'

A clear conditional reading is found in examples (26) and (27) below. Here, the event of the unmarked clause will only take place under the condition of the *go*-clause (in both examples a desiderative verb).

(26) *m-á-na-'-i-ɬ-dė* *yį-ka-Ø-t'į -o-'ąą*
3O-BEN-LEX-INDO-ASP-V/V-wash 3O-LEX-3S-want-**LINK**-TOP
'Wash it for them if they want it.'

(27) *yi-k'é-dą́ą-Ø-géésh* *yi-ká-daa-Ø-t'i-**go*** *hń-Ø-zi-ná*
3O-for-PL.LEX-3S-look 3O-LEX-PL.ASP-3S-want-**LINK** ASP-3S-think-NARR
'They can come to see it if they want to, he thought.' (HOIJER 8:39)

Despite the different implications in the relationship of two clauses, there is no additional linguistic information (such as e.g. a conjunction) to differentiate between temporal and conditional uses of the go-construction.

4 The linker *-go* occurs phonetically reduced as *–o* especially following vowels, but also in fast connected speech.

5.2. Discourse embedding

Verbal predicates marked with the linker *–go* do not always occur in a clausal combination together with an unmarked verbal predicate, but are also found as structurally and intonationally independent units in discourse. These clauses can be said to be dependent not on another single clause, but rather to be embedded in a larger discourse unit.

(28) *shi shįį shi -shdą́zha łį m-ee-naa-t'aash-**go***
1s and 1sO-little sister horse 3O -P-CONT -go.DU-**LINK**
'Me and my little sister were on horseback.'

Example (28) relates background information on the general situation in this narrative.

Another possibility is the chaining of several verbs marked with *–go*, as in the following example:

(29) *doo aíí di dá-hį-dał-**go** ą-'ee aíí*
NEG DEM 2S EMPH-2SGS.CNJ-go.SG-**LINK** there-LOC DEM

*mi-ł -d-n-tłi -**go** m-ee-kaa-náá-' -į -gis-**go***
3O-COM-LEX-2SS-tremble-**LINK** 3-INST-water.DIR-ITER-INDO-2SS.ASP-wring-**LINK**
'Not that you go over there, try to do everything at once, and wash with it again.'

In this conversational example, all three verbs are equally marked by the linker. Although a sequence of events is implied, the overall perspective on this discourse unit is to express the sequence as one overall event. The negative particle introducing this unit has scope over all three predicates, which contributes to the interpretation as one event unit.

5.3. Auxiliary constructions

Jicarilla Apache does not employ many auxiliaries in order to express temporal, aspectual, or modal specifications in addition to a main predicate, and is in this respect more conservative than Navajo or Western Apache. Two constructions that conceivably belong to this grammatical domain both employ the linker *–go* suffixed to the lexical verb. One of the auxiliary-constructions expresses past habitual aspect (note that *–o* and *–yo* are variants of *–go*):

(30)a. *daa-sh-di-ł -tłi -**o** ni-ghai*
PL-INDS-LEX-TR-burn-**LINK** PFV-years.pass.PFV
'They used to burn them for light'

b. *na- '- ni- s- soo-yo* *ní -i-ghai*
CONT-INDO-LEX.ASP-1sS-herd-**LINK** PFV-1sS-years pass.PFV
'I used to herd sheep'

The verb used as auxiliary *n-ghai* 'time/years pass(es)' combines with the lexical verb and the linker to express a simultaneous event with habitual meaning. The second auxiliary-construction consists of the copula and a lexical verb marked with the linker. This combination expresses an epistemic modal meaning of uncertainty, as in example (31):

(31) *enda* *mi-'máá* *m-a-na-' -ń-t'é-o* ***át'é***
but 3O-mother 3O-BEN-BACK-LEX-ASP-come-**LINK** 3S.COP
'but its mother may have come back to it'

The copula only appears in its third person form. This auxiliary construction can be used interchangeably with the adverb *naate* in the meaning 'maybe'.

5.4. Analytic causative constructions

There are two alternative constructions to express causativity in Apache: a synthetic one and an analytic one. The analytic construction consists of the actual lexical verb, fully inflected and marked by the linker *–go,* followed by an equally fully inflected form of the verb *á-...ɬ-'į* 'to make thus', as seen in the next set of examples:

(32)a. *hi-sh-ą-ǫ* *á-shį-nn-la*
ASP-1sS-eat-**LINK** LEX-1sO-2sS.ASP-make
I eat you make me
'You make me eat'

b. *di-Ø-ɬ-jį-ǫ* *á -sh -į*
LEX-3S-V/V-sneeze-**LINK** LEX-1sS-make
she sneezes I make (her)
'I make her sneeze'

(33) *ko-'i* *dásade* *n-ɬ-wosh-go* *á-ch'i-ɬ-'įį*
water-SPEC already ASP-V/V-boil-**LINK** LEX-IND.H.S-V/V-make
'You already started to boil the water' ('you made it thus that it came to a boil')

The causer and the causee are both marked as pronominal arguments on the second verb, while the first verb marked by the linker takes the causee as subject argument and expresses the caused event.

5.5. Complements of desiderative and cognitive verbs

The construction of complex events that consist of a desiderative verb and its complement depends on the coreferentiality, or non-coreferentiality, respectively, of the subject arguments of both clauses. If the subject arguments are not coreferential, then the linker will be used between the lexical verb and the desiderative verb, as in (34 a). But if the subject arguments are coreferential, then no overt marking of any type intervenes between lexical verb and desiderative verb (34b).

(34)a. *díijį́ t'aiisą́ǫ n -**n**-dei-**o** há -**sh**-t'į́*
today early afternoon back-ASP.**2sS**-come-**LINK** LEX -**1sS**-want
'**I** want **you** to come home early today'

b. *díijį́ t'aiisą́ǫ na -n-**sh**-dai há -**sh** -t'į́*
today early afternoon back-ASP-**1sS**-come LEX -**1sS**-want
'**I** want to come home early today'

The two distinct constructions occur in elicited data (34 a and b) as well as in spontaneous speech (example 35 a and b), so that we can assume that this difference in marking represents a robust pattern.

(35)a. *adą́ doo-n'-łah-hę́ę́-Ø-t'aash yį́-k-a-Ø-t'į́*
and then NEG-REVERS-COLL-ASP-3S-go.DL 3O-LEX.ASP-3S-want
'And then they didn't want to stop running'

b. *dii da-'áá-Ø-ni-**o** yį́-ka-Ø--t'į́*
this EMPH-LEX.ASP-3S-true-**LINK** 3O-LEX.ASP-3S-want
'he wanted this to be true'

The desiderative verb constructions can be related to the properties of GIVON'S (1980) binding hierarchy of complements; coreferential expressions (34b and 35a) exhibit a tighter semantic relationship than non-coreferential expressions (34a and 35b), and this relationship is mirrored in the lack vs. presence of overt grammatical marking.

As opposed to the desiderative construction, other combinations have a fixed pattern; they either always employ the linker or they don't. A verb which never employs the linker is the cognitive verb *yi...-li* 'believe'. In example (36) the subjects arguments of the two verbs are not coreferential, and still there is no extra morphological marking to signal the combination.

(36)a *i -naa-Ø-dzaa* *y -ą́ą-yó -Ø-li*
DIR-LEX.PFV-3S-go.SG 3O-about-LEX.PRG-3S-believe
'she believed that he left'

b. *dii* *da-'áá-Ø-ni-* *y -ą́ą-yó -Ø-li*
this EMPH-LEX.ASP-3S-true 3O-about-LEX.PRG-3S-believe
'he believed this to be true'

Example (36b) compared to (35b) is unmarked as to the relationship of the two verbs. Another cognitive verb – *n...zį́* 'think' – also employs no overt marking when conjoined with a complement verb:

(37) *dii* *da-'áá-Ø-ni-* *ń-Ø-zį́*
this EMPH-LEX.ASP-3S-true ASP-3S-think
'he thinks that this is true'

A different construction with a related verb 'think of' can be seen in (38):

(38) *moosha* *m-aa-shį́-nį́-o* *shi-di-i-Ø-ł-jį́*
cat 3O-about-1sO.ASP-think-**LINK** 1sO-LEX-ASP-3S-v/v-sneeze
cat I think of it it makes me sneeze
'Thinking of the cat makes me sneeze'

Here, the linker is part of a construction in which the verb 'think of' plus the linker can be either argued to represent a temporal relationship ('When I think about the cat it makes me sneeze'), or they could represent a nominalized complement of the main clause ('The thinking of the cat made me sneeze'). The main verb here is a synthetically expressed causative (as opposed to the analytic causatives discussed above in 5.4.). To analyze the linker as a nominalizer does not seem feasible, though, since there are no distinct occurrences in which the linker would straightforwardly function as such.

There is one more type of construction including the linker *–go* in the domain of cognitive verbs: it seems that the expression of 'to know' differs from all the above examples, in that it is always the one marked by the linker, as in (39):

(39) *da-mégo-sį́-ǫ* *diijį́* *t'aiisą́ǫ* *na-n-sh-dai*
EMPH-LEX-1sS.know-**LINK** today early afternoon back-ASP-1sS-come
'I know that I'll come home early today'

This construction is also exceptional in its word order: the supposed main verb precedes the complement clause. This differs from the otherwise systematic word order exemplified

above, in which the main verb always follows its complement verb that may then be marked by the linker.

5.6. Resultatives and verbs of perception

A similar construction to the analytic causative construction introduced above is used to express resultatives, except that the object of 'make' represents the result of this activity [X makes Y Z-*go*]. There is no separate word class of adjectives in Apache, instead stative verbs are employed.

(40) *n-Ø-keeɬ-**go*** *á-ch'i-ɬ-'ɻɻ-**go*** *a-Ø-ɬ-t'a-da-**o***
ASP-3S-flat-**LINK** LEX-IND.H.S-V/V-make-**LINK** LEX-3S-V/V-thin-ADV-**LINK**

*á-**sh**-aa-ɬ-iɬ*
LEX-IND.H.S-ASP-V/V-make.PRG
'You make it flat, you make it thin.'

The verbs 'be flat' and 'be thin' are suffixed by the linker; the first occurrence of 'make' also exhibits this marking, but this belongs to the general construction of temporal simultaneity between the two events 'make the flat AND thin'.

A last word class employing a construction with the linker is the class of perception verbs. If there is an argument sharing between the verb of perception and its complement of the type that the object of the perception verb is also the subject of the complement verb, then the embedded verb will be marked by the linker, e.g. [X saw Y doing Z-*go*].

(41)a *dlǫ'* *iǹkįní* *yi-bąąshį* *si-Ø-zi-**go*** *daa-hɻɻ-ɬ-tsą*
prairie dog road 3O-beside ASP-3S-stand-**LINK** PL-ASP.INONSS-saw
'we saw the prairie dog standing beside the road'

This construction type differs from an alternative relative clause construction that would employ the nominalizing suffix

(41)b *dlǫ'* *iǹkįní* *yi-bąąshį* *si-Ø-zi-**i*** *daa-hɻɻ-ɬ-tsą*
prairie dog road 3O-beside ASP-3S-stand-**NOM** PL-ASP.INONSS-saw
'we saw the prairie dog that was standing beside the road'

While the *go*-construction focuses on the embedded event as a whole, the nominalizing construction merely modifies or identifies its head noun and has no direct semantic relationship to the event of 'seeing' as is the case in (41a).

5.7. Non-verbal constructions with *–go*

Besides the use of *–go* in clause combining and complex predicate constructions, there is a related use of it in the nominal domain: suffixed to nouns or demonstratives it is usually translated as 'also', as in example (42):

(42) *adą́ chinii doo-naa-go-ɬį́ dakoo ą-shį́ ɬį-**go** géh*
then dog NEG-CONT-LEX –bark now there-from horse-**LNK** just

*nzhǫn-na-kai dibé -**go** n'ą́ą́ géh nzhǫn-e-o*
good-CONT -go.PL sheep-**LINK** over there just good-ADV-LINK

naa -daa-kai
around-DIST-go.PL
'Then the dogs were not barking now, and the horses were standing calmed down, the sheep were also standing around over there, all calm'

The different animals being present in the situation are listed in his example. The first clause talks about 'dogs', and then the 'horses' of the second clause as well as the 'sheep' of the third clause are suffixed by *–go*.

(43) *Adą́ iyáná' na-i-Ø-ɬ-'į́ d-nn-'nį́-ǫ*
then what CONT-3O.ASP-3S-V/V-bark LEX-1NONSS-say-LNK

*náhi-**o** géh na-ɬ-ha-gǫ́ǫ́-tso*
1NONS-**LINK** just 1NONSO-COM-LEX-ARES.ASP-surprise
'Then we said: what is he barking at?, it just surprised us.'

Here, the independent emphatic pronoun *náhi* occurs with the linker. This use should not be necessary to link the last clause to the previous one, since the preceding verb 'say' already is marked for it to express simultaneous events. There is also no reading of the pronoun as part of a list comparable to example (42). Still, the use of the linker with the demonstrative adds a reading of a temporal ongoing event, and sets off the pronoun as an independent discourse unit.

Finally, the linker *–go* occurs in often lexicalized expressions denoting temporal notions, such as the particles *ɬa**go*** 'sometimes'(< *ɬa'* 'some' + *go*), *ha'**go*** 'when', or *hai**go*** 'in the wintertime', and *dą́ą́**go*** 'in the springtime'. Many temporal adverbs can be actually related back to a verb plus the *–go*. Usually the linker has been reduced to just occur in its reduced as a vowel in these expressions. The change can be seen by comparing some present Jicarilla adverbs (example 44) with GODDARD'S (1911) (example 45) spelling.

(44)a.	*yiiskąǫ*	'tomorrow' (< 'when the night ends')	
b.	*shá'ï'ąǫ*	'in the evening' > (< 'when the sun sets')	
(45)a.	*yiiskągo*	'tomorrow'	
b.	*shá'ï'ągo*	'at sunset'	(Goddard 1911)

An additional area in the lexicon in which the linker is found is the formation of abstract nouns by combining the linker with verbs.

(46)*go-ɬ-n-zhǫ-go*	<	*mi-ɬ-n-Ø-zhǫ*
INDO-COM-ASP-good-**LINK**		3O-COM-3S-ASP-good
'love'		's/he loves him/her'

As can be seen in (46), the a change in the pronominal object in the experiencer marking accompanies the abstraction of the expression: an indefinite object pronominal is used instead of the normal pronominal object. The use of the *go-construction* in the formation of new lexical items from inflected verbs takes us back to its clausal functions, especially as indicator of adverbial temporal relations.

6. Summary

Clause-combining in Jicarilla Apache does not exhibit strong subordinating or embedding tendencies. All verbal predicates are fully inflected for person and tense/aspect and there are no restrictions of any morphological kind on the putative subordinate predicates. Besides coordinating structures the other two main constructions to combine clauses are the nominalizing construction and the one discussed here in more detail, the *go*-construction. This latter construction exhibits a whole range of uses in clause combining, from temporal to other inferenced adverbial relations such as conditional marking. It is also integral part of the formation of some complex predicates, such as causatives, resultatives, and some complementing structures. It is pervasive throughout the grammar, and as hinted at with the last examples, also well-represented in the lexicon. This wide range of occurrence is due to its general function of marking a simultaneous relationship of two or more entities.

REFERENCES

GENETTI, CAROL. 1991. From postposition to subordinator in Newari. Approaches to grammaticalization, ed. by Elizabeth Close Traugott and Bernd Heine. Amsterdam: Benjamins, 227-255.

GIVÓN, TALMY. 1980. The binding hierarchy and the typology of complements. Studies in Language 4:333-78.

GODDARD, P.E. 1911. Jicarilla Apache Texts. New York: Anthropological Papers of the American Museum of Natural History, 8.

HOIJER, HARRY. n.d. Jicarilla texts. Ms. Philadelphia: American Philosophical Society Library.

HOIJER, HARRY. 1975. The history and customs of the Lipan, as told by Augustina Zuazua. Linguistics 161.5-38.

JUNG, DAGMAR. 1994. Functions of the nominalizer *-i* in Jicarilla Apache. Santa Barbara, Ms.

JUNG, DAGMAR. 2000. Word order in Apachean. The Athabaskan Languages: Perspectives on a Native American Language Family, ed. by Theodore B. Fernald & Paul R. Platero. Oxford: Oxford University Press, 92-100.

LONGACRE, ROBERT E. 1985. Sentences as combinations of clauses. Language typology and syntactic description, vol.II: complex constructions, ed. by Timothy Shopen. Cambridge: Cambridge University Press, 264-65.

MITHUN, MARIANNE. 1984 How to avoid subordination BLS 10, 493-509.

MITHUN, MARIANNE. 1999. The languages of Native North America. Cambridge: Cambridge University Press.

POTTER, BRIAN. 1997. Wh/Indefinites and the structure of the clause in Western Apache. Dissertation, University of California Los Angeles.

SCHAUBER, ELLEN. 1979. The syntax and semantics of questions in Navajo. New York & London: Garland.

SCOLLON, RONALD. 1985. The sequencing of clauses in Chipewyan narrative. Grammar inside and outside the clause, ed. by Johanna Nichols and Anthony C. Woodbury. Cambridge: Cambridge University Press, 113-131.

Infinitives in Polysynthesis: the case of Rembarrnga[1]

RACHEL NORDLINGER and ADAM SAULWICK

Abstract

It is a well-known cross-linguistic generalisation that polysynthetic languages tend to avoid infinitival clauses (HEATH 1975, MITHUN 1984). In BAKER's (1996) framework this follows directly: polysynthetic languages must encode arguments on the verb (to comply with the Morphological Visibility Condition) and, due to general theoretical assumptions, this is possible only in tensed clauses. For BAKER (1996), therefore, non-finite clauses in polysynthetic languages are an impossibility. In this paper we challenge this conclusion with data from Rembarrnga, a polysynthetic language of Northern Australia. We show that Rembarrnga does indeed have infinitival clauses, both with and without pronominal affixes, and discuss the structure of these clause types in some detail. Furthermore, we argue that Baker's attempts to reanalyse these clauses as either nominalisations or fully finite forms are inaccurate and lead to a skewed analysis of the data. We conclude that the absence of non-finite clauses can not be definitional of polysynthesis, which must encompass the possibility of infinitives in a language like Rembarrnga.

1. Introduction

BAKER's (1996) well-known theory of polysynthesis predicts that non-finite clauses should be impossible in polysynthetic languages (p. 452). This follows from two central assumptions: (i) non-finite verb forms cannot contain agreement morphology. Agreement morphemes can only be adjoined to heads that have structural Case-assigning features (p.

---

1 We wish to thank the following Rembarrnga speakers who provided the language data cited here recorded by Adam Saulwick: Jackie Bun'garniyal, Wayne Campion, Ronnie Martin, Smiler Martin, Brian Nyinawanga. We would also like to thank Nick Evans for inviting us to write this paper and for discussion of many of the issues herein.

192), and while tensed INFL is one such head, tenseless INFL (following general assumptions within the GB/Principles and Parameters tradition) is not (p. 472); (ii) definitional of polysynthetic languages is their total adherence to the Morphological Visibility Condition (MVC) (1), which requires each thematic role assigned by the verb to be associated with a morpheme in the verb itself, either through agreement morphology or incorporation.

(1) *The Morphological Visibility Condition (MVC)* (BAKER 1996: 17)
A phrase X is visible for *Q*-role assignment from a head Y only if it is coindexed with a morpheme in the word containing Y via:
(i) an agreement relationship, or
(ii) a movement relationship

From these two assumptions it follows that a non-finite (tenseless) verb must necessarily violate the MVC, making infinitival clauses in polysynthetic languages an impossibility.

Unfortunately for Baker, this is simply not the case. In this paper we challenge Baker's theory with data from the polysynthetic Australian language Rembarrnga, which has infinitival clauses, both with and without pronominal agreement morphology, as exemplified in (2) and (3) respectively:[2]

(2) *Nginy-waralh-miny* *guwa* ***nginy-ro-ngœ***
1>2-ask-PP PURP 2-go-INF
'I asked you to go.' (Saulwick 4-1-14)

(3) *Mala-galitj* *yarr-nœtti-ny* *mahgun* ***ma-ngœ-gan***[3]
all-other 1a>3-keep-PC again get-INF-DAT
'We used to keep the other half (of bullock coaches) again to get more. (While the others used to go back to the station.)' (Saulwick 1-70-8)

---

2 The large majority of our examples are taken from the unpublished field notes of the second author. Numbers after each example refer to the location in the original texts. Abbreviations used in the glosses are: 1>2 = indicates first person acting on second person, etc. (unless otherwise stated pronominal number is minimal); 12 = 1st person inclusive; a = augmented; ABL = ablative; ALL = allative; BEN = benefactive applicative; CAUS = causative; COM = comitative; CON = connective; DAT = dative; DEM = demonstrative; DM = discourse marker; DUP = reduplication; EX = exclusive; FUT = future; INF = infinitive affix; INTERROG = interrogative; IRR = irrealis; masc.pro = masculine free pronoun; NEG = negative; NML = nominal; NP = non-past; PT = past; PC = past continuous; PROG = progressive; POSS = possessive; PP = past perfective; PRES = present; PURP = purposive; RR = reflexive/reciprocal; SEQ = sequential; STM = stem-forming affix; TEMP.LOC = temporal locative.

3 Note that pronouns in Rembarrnga distinguish number according to a minimal/augmented contrast (see CONKLIN 1962, among others) in which augmented is equivalent to plural for all person categories except first person inclusive for which it marks more than three.

Each of these examples contains both a main verb, inflected for tense and aspect, and a non-finite subordinate verb which is inflected with a distinct, tenseless suffix we gloss INFinitive. Main verbs always carry pronominal agreement prefixes. Infinitival verbs, on the other hand, can appear with pronominal prefixes (as in (2) or without (as in (3)). There is no empirical reason to treat these subordinate verbs as anything other than non-finite infinitival clauses: they are clearly tenseless, as we discuss in more detail below, and perform the same general functions as infinitival clauses in other, non-polysynthetic, languages.

Part of the problem lies in determining what falls within the definition of "infinitive". Traditional definitions (e.g. Dictionary of Language and Linguistics (HARTMANN & STORK 1972), The Oxford English Dictionary (SIMPSON & WEINER 1989), The Oxford Encyclopedic English Dictionary (HAWKINS & ALLEN 1991)) define infinitives as verb forms unspecified for person, number and tense. According to this definition, the subordinate clause in (3) is infinitival, but that in (2), which contains person marking, is not. However, the existence of infinitival verb forms inflected for subject in various Romance languages (RAPOSO 1987, LEDGEWAY 1998, VINCENT 1998, MENSCHING 2000) shows this definition to be too restrictive. Rather, as VINCENT (1998) concludes, in the definition of finiteness "it seems clear that the availability of person/number marking has no role to play" (p. 147); it is possible to have non-finite verb forms inflected for person. This is exactly what we find in the Rembarrnga data, as shown in (2).

The existence of such infinitival clauses in a polysynthetic language like Rembarrnga, poses serious problems for BAKER's (1996) claim that infinitival clauses are incompatible with polysynthesis. Moreover, the presence of inflected non-finite forms such as in (2) undermines Baker's assumption that pronominal agreement morphology can only co-occur with tensed verbal forms, an assumption that has already been seriously questioned by work on inflected infinitives in Romance languages.

BAKER (1996: 475-476) is aware of the potential problems presented for his theory of polysynthesis by these clause types in Rembarrnga as described in an earlier grammar by MCKAY (1975). However, in an attempt to maintain his position, he claims that these verbs are not infinitives, but should be reanalysed as either fully finite verbs (2) or nominalisations (3), thereby conforming to his theoretical framework after all. The primary purpose of this paper is to challenge this position and show that such reanalysis is inadequate and leads to a skewed representation of the empirical facts. Rather, the possibility of infinitival clauses must be accounted for in any comprehensive theory of polysynthesis.

In the following section, we begin with an overview of the relevant aspects of Rembarrnga grammatical structure, focussing particularly on the structure of the 'verbal word' and the tense/aspect/mood paradigms. We turn to a detailed discussion of the nature of infinitival clauses and their ramifications for Baker's theory of polysynthesis in section 3.

2. Typological overview of Rembarrnga

Rembarrnga is a Gunwinyguan language of Central Arnhem Land in Northern Australia. Like most other Gunwinyguan languages (including Bininj Gun-wok (EVANS, this volume), and Ngalakgan (BAKER, this volume)) it is highly polysynthetic, with verbs having multiple affix slots, including pronominal cross-referencing, valence-changing affixes, and the possibility of incorporated nominals and adverbials.

Table 1 presents the basic morphological template for the Rembarrnga verb.[4] Note that some of the slots in the template are possibly further divisible. For example: (i) the pronominal prefix slot is generally glossed as a portmanteau but may sometimes be analysed into subject and object forms also bearing tense; (ii) more than one verbal applicative may occur in a single verbal word; (iii) on occasion more than one adverbial may be strung together; (iv) verb roots may be reduplicated and also incorporated; (v) some aspectual forms compound with anomalous verbs to form aspect markers.

Table 1: A simplified structure of the Rembarrnga verbal word

bnd pro	numb	app	adv	IN	app	verb	stem suff	caus/ trans	rflx/ recip	tns/asp/ mood	numb	case
prefixes						roots	suffixes					

Abbreviations: bnd pro = bound pronominal, numb = pronominal number, app = applicatives, adv = adverbials, IN = incorporated nominal, stem suff = stem-forming suffix, caus = causative, trans = transitiviser, rflx/recip = reflexive/reciprocal, asp = aspect, tns/asp/mood = tense/aspect/mood.

The following examples exhibit some of the complexities suggested by this template, including the use of benefactive (4) and comitative applicatives (5), nominal incorporation (6), and adverbial incorporation (7).

(4) *Dow* *wurrhwurrungu* *yarran-ba-bak-rdi-rditj-mœ-rri-niny*
hunting.fire old.people 3>1a-a-BEN-DUP-return-STM-RR-PP
'The old people would all return to us (from) fire hunting.' (Saulwick 2-83-42)

4 In fact, the structure of the verbal word is more complex than Table 1 suggests: there are co-occurrence restrictions for some affixes such that not all affix slots may be filled concurrently, and some slots indicated here may be further divided. For full discussion see MCKAY (1975) and SAULWICK (in preparation).

(5) *Ga-bœh-re-guwamh-miny.*
NP3>3-lest-COM-depart-PP
'Lest she left with it.' (Saulwick 2-168-60)

(6) *"Garda! Nga-y-wurrpparn-marninyh-mœ-ttœ-na", ø-yi-niny.*
oh 1-COM-emu-make-STM-RR-FUT 3-say-PP
' "Oh! I will turn myself into an emu", she said.' (Saulwick 2-168-86)

(7) *Barr-bak-garra-bulhwarr-miny.*
3a>3-BEN-upwards-scatter-PP
'They flew up away from her.' (Saulwick 2-168-23)

Subject and object pronominal agreement prefixes are generally obligatory on all Rembarrnga finite verbs (they may be omitted with infinitival verbs, as discussed in § 3). The presence of these prefixes, combined with the complex morphological and incorporation possibilities of verbs, means that a single verb frequently represents a full and complex sentence (7).

Rembarrnga distinguishes four tenses: past punctual (PP), past continuous (PC), present (PRES) and future (FUT), and also irrealis mood (IRR).[5] Additionally verb roots must contain a stem suffix (STM) in order to further affix markers such as reflexive/reciprocal (RR) and some aspectual markers. As we saw in (2) and (3) above, in place of tense/aspect/mood marking, verb roots in non-finite clauses may contain an infinitival suffix (INF). Depending on the verbal conjugation, the infinitive marker is homophonous with either the present tense form, or the irrealis form. Despite this formal similarity, these affixes can be distinguished on paradigmatic and semantic grounds, as discussed in § 3.

Depending on the method of classification there are six (or seven) verb conjugation classes. A simplified paradigm containing the primary four verb classes that will concern us here is given in Table 2 below.[6]

Table 2: Partial verb conjugation paradigm.

Class	present	future	past punctual	past continuous	irrealis	infinitive
1	*ø*	*-Ca*	*-miny*	*-mœrn*	*-mœ*	*ø*
3	*-ra*	*-ra*	*-ba*	*-bœrn*	*-bœ*	*-bœ*
4	*-n*	*-na*	*-nginy* +	*-niny*	*-nœ*	*-nœ*
5	*ø*	*-ngara*	*-ya* +	*-nginy* +	*-ngœ*	*-ngœ*

5 Note that this affix is referred to as 'past counterfactual' by McKay (1975). In contrast, we follow Saulwick (in prep.) in treating it as a marker of general irrealis mood, not restricted to the past tense.

6 Forms marked with the subscript '+' actually have a number of allomorphs, of which we provide only the most common here. For a more extensive paradigm see McKay (1975:132).

Full conjugations of two class 5 verbs are given to exemplify the paradigm given in Table 2: *ma* 'get' (8) to (12) and *ra* (with allomorphs *ro* ~ *ri*) 'go' (13) to (17). Subscripted numerals following the English gloss of verbs show verb class membership.

(8) *Naganh* *durra-na* *yarr-ma-ø...*
this body-NML 3a>3-get$_5$-PRES
'We get this body...' (Saulwick 1-81-16)

(9) *Munmunh* *da-ma-ngara.*
grass.sp 2>3-get$_5$-FUT
'You'll get "munmunh" grass.' (Saulwick 1-12-2)

(10) *Matjjih* *mundu* *nga-mi-ya...*
CON stone 1>3-get$_5$-PP
'And I got a "mundu" stone.' (Saulwick 1-12-17)

(11) *Gali-na* *mambart* *yarr-ma-nginy...*
big-NML billy.can 1a>3-get$_5$-PC
'We would (always) get a large drum...' (Saulwick 1-70-22)

(12) *Nœnda-ji-ma* *mœlak* *jundu* *yarr-ma-ngœ* *gali-na-wala.*
that-TEMP.LOC-DEM NEG stone 1a>3-get$_5$-IRR big-NML-ABL
'At that time, we didn't get any money from the government.' (Saulwick 1-8-20)

(13) *Yeneh-gah* *nginy-ra-ø?*
INTERROG-ALL 2-go$_5$-PRES
'Where are you going (now)?' (Saulwick 1-0-8)

(14) *Yeneh-gah* *nginy-ro-ngara?*
INTERROG-ALL 2-go$_5$-FUT
'Where will you go?' (Saulwick 1-77-2)

(15) *Mitjjindah* *yarra-ri-ya* *yarr-da-nginy* *mutbu.*
long.ago 1a-go$_5$-PP 1a>3-wear$_5$-PP hair.belt
'In former times, we went, we wore hair belts.' (Saulwick 2-13-12)

(16) *Mœlak* *banh-wala* *birri-ro-ngœ...*
NEG here-ABL 3a-go$_5$-IRR
'They didn't go from here...' (Saulwick 3-46-3)

(17) *ø-rdurl-ga-bœ-yi-ro-nginy.*
PT3>3-catch.fire$_1$-CAUS$_3$-STM-PROG$_5$-PC
'He would be lighting fires.' (Saulwick 3-40-9)

Note that causative *–ga*-derived verbs take Class III conjugation. This is glossed with a subscript 3 following the causative marker. Additionally, some verb roots such as *ra* 'go' (see (17)) may function as aspectual suffixes and thus require subsequent affixes of their conjugation class.[7]

3. Infinitival clauses in Rembarrnga

Infinitival clauses in Rembarrnga are used in purposive and jussive clauses and are exemplified by (2) and (3) repeated from above.

(2) *Nginy-waralh-miny* *guwa* *nginy-ro-ngœ.*
1>2-ask-PP purp 2-go$_5$-INF
'I asked you to go.' (Saulwick 4-1-14)

(3) *Mala-galitj* *yarr-nœtti-ny* *mahgun* *ma-ngœ-gan.*
all-other 1a>3-keep-PC again get$_5$-INF-DAT
'We used to keep the other half (of bullock coaches) again to get more. (While the others used to go back to the station.)' (Saulwick 1-70-8)

In these clauses the main verb carries regular finite tense/aspect/mood suffixes (see § 2), while the subordinate verb is inflected with the infinitival suffix, the form of which differs depending on the conjugation of the verb (see Table 2). These can be distinguished from finite verbs on a variety of grounds.

Firstly, while finite verbs almost always contain pronominal agreement morphology, infinitival verbs need not, as in (3).[8] Secondly, a verbal word inflected with the infinitival

---

7 We assume that the aspectual use of these verbs results from reanalysis of an earlier associated motion construction, in which the main verb root *ra* 'go' incorporated another predicate.

8 In fact, it also possible, though extremely rare, for finite verbs to occur without bound pronouns, as evidenced by the following example:

(i) *Jacket* *nga-bak-juyh-wa* *guwa* *da-ø.*
1>3-BEN-send-FUT PURP wear$_5$-PRES
'I'll send him a jacket to wear.' (Saulwick 4-1-6)

Such examples pose an additional problem for Baker's Morphological Visibility Condition in that they contain a fully finite verb that does not include a morphological representation of each

suffix cannot function as a main verb without reinterpretation of its temporal semantics to the finite category that the infinitival suffix is homophonous with: either irrealis mood (for class 3, 4 or 5 verbs) or present tense (for class 1 verbs). For example, if the infinitival verb *yarr-marra-ngœ* in (18) is cast in a main clause as the main clause predicate, the temporal semantics it contributes has to be different from that in (18) – i.e. it has to be interpreted as irrealis mood ('we would spear a devil'). This interpretation is not possible when it functions as an infinitival predicate.

To mark its purposive function the infinitival clause may use a particle *guwa* translatable as 'so' or 'in order', or the verbal word may be additionally inflected with the dative suffix. (The use of dative case to mark the verb of purposive subordinate clauses is extremely common in Australian languages (BLAKE 1977, DIXON 1980)).[9] There are at least two dialects in Rembarrnga, one which inflects the infinitive verb with the dative suffix, and one which doesn't. Compare (18) and (19), from the northern and southern dialects respectively.[10]

(2) *Yarra-many* ***guwa*** ***ngayang*** ***yarr-marra-ngœ.***
1a.ex-went PURP devil 1a>3-spear$_5$-INF
'We went to spear a devil.' (Saulwick 4-1-36)

(3) *Ngayang-gan* *yarra-ri-ya-ma*[11] ***yarr-marra-ngœ-gan.***[12]
dead.man-DAT 1a-go-PP-DM 1a>3-spear$_5$-INF-DAT
'We went (in revenge) for the dead man, to spear (the man responsible for his death).' (McKay 1975:321 3.6-12)

Although the purposive clauses in these two examples are functionally and semantically equivalent, the latter contains the dative case affix, while the former does not.[13] This

---

thematic role associated with it. However, since this is the only example of such a finite verb in the present corpus, further investigation is required.

9 Finite verbs in purposive clauses may also be inflected with the dative suffix, so this cannot be taken as proving that these infinitival verbs are not verbal – see § 3.2 for further discussion.

10 In checking the data one speaker commented "That's the Bulman one, you put that -gan on the end." This statement refers to the dialectal difference between southern (Bulman) and northern regions.

11 The glossing on this affix is intentionally vague due to the fact that its function is not well understood. Further investigation is required.

12 McKay (1975) gives examples in a slightly modified IPA using voiceless stop symbols for all stops. Examples from McKay cited here are for purposes of consistency transliterated into the version of the practical orthography used throughout this paper. This uses voiced stop symbols syllable initially, voiceless symbols syllable finally and homorganic voiceless stop symbols for geminate stops. Rembarrnga has a sixth mid central vowel, represented here with <œ>.

difference has no bearing on the properties of infinitival clauses at issue in this paper, and we use examples from both dialects interchangeably.

As mentioned above, there are two types of infinitival verbs: those which contain pronominal agreement morphology, and those that do not. These two possibilities appear to be in free variation, with speakers readily interchanging the two within the same contexts. Since these types pose different problems for Baker's framework, we will discuss each type separately.

3.1 Infinitives with pronominal agreement

Infinitival verbs with pronominal agreement are identical to fully finite verbs except that the verb is inflected with the infinitival suffix. Compare (20) and (21) which have comparable semantics.

(20) *Balay* *ngan-waralh-miny* *guwa* *danggu-na* ***nga-bak-rniya-ø.***
already 3>1-ask-PP PURP meat-NML 1>3-BEN-cook$_5$-PRES
'She already asked me to cook her that meat.' (Saulwick 4-1-23)

(21) *Nginy-waralh-miny* *guwa* *danggu-na* ***da-bak-rniya-ngœ.***
3>2-ask-PP PURP meat-NML 2>3-BEN-cook$_5$-INF
'She asked you to cook her that meat.' (Saulwick 4-1-22)

Both (20) and (21) contain jussive clauses yet in (20) the subordinate verb is fully finite, inflected with the present tense, while in (21) it is inflected with the infinitive suffix. Furthermore, the subordinate verb in (21), while inflected with the infinitive suffix, still contains regular pronominal agreement morphology.

The existence of such pronominally-inflected infinitives does not violate Baker's Morphological Visibility Condition, since in this case the arguments selected by the verb *are* encoded morphologically in the verbal word. However, they would still be ruled out in his framework by the fact that agreement morphemes cannot be adjoined to a non-finite INFL, since it is not an assigner of structural Case (see § 1). Baker's solution, implicit in his reanalysis of Rembarrnga infinitives that will be discussed in detail in § 3.2, is to take the presence of agreement morphology as an indication that the verb is in fact finite after all. Thus, according to Baker, these are not infinitives but fully finite forms, with the agreement morphology adjoined to a tensed INFL in the phrase structure.

---

13 The presence of the particle *guwa* is independent of the presence or absence of the dative case affix. See (28) for an example of an infinitival clause containing both the dative case and the purposive particle.

However, there are serious problems with this approach. Firstly, if the theoretical hypothesis is that agreement morphology is only possible with finite verbs, then it creates a completely circular argument if we prove this hypothesis by simply taking the presence of agreement morphology in the first place to indicate that the verb is finite. Thus, given the absence of any other empirical evidence that these are in fact finite verbs, and when viewed from outside Baker's theoretical assumptions, the presence of pronominal agreement does not preclude an analysis of these clauses as infinitives, particularly since pronominally-inflected infinitives are well-attested in other languages, such as Portuguese (RAPOSO 1987), Calabrian (LEDGEWAY 1998) and Old Neapolitan (VINCENT 1998) (22).

(22) *li quali* *tu* *commanderray* *de* *liberare**nosse**.*
Who you order.FUT to free-AGR.3PL-REFL
'who you will order to free themselves' (VINCENT 1998:136, ex. (2))

Secondly, crucial to Baker's reanalysis of these infinitives as fully finite forms is the fact that the infinitival suffix must be marking some finite category. Yet it is unclear what this category might be. As noted in § 2, the infinitival form must be analysed as a distinct member of the verbal paradigm despite the fact that each conjunctional variant is homophonous with some other tense form. Consider the following examples. Examples (23) and (24) contain the class 1 verb root *yaw* 'spear' inflected with the infinitival suffix (23) and the homophonous present tense suffix (24). While (25) and (26) make the similar contrast for the class 3 causative stem *rdurlh-ga* 'catch.fire-CAUS', which is inflected with the infinitival suffix in (25) and the homophonous irrealis mood suffix in (26).

(23) *Walkkur-nggœ* *bordi* ***ø-marninyh-miny*** *guwa* *guweny* ***yaw-ø.***
son-2POSS spear PT3>3-make-PP PURP kangaroo $spear_1$-INF
'Your son made a spear to spear (a) kangaroo.' (Saulwick 4-1-7)

(24) ***Yarr-yaw-ø*** *guweny* *bartta-murrnginy-yi.*
1a>3-$spear_1$-PRES kangaroo COMIT-shovel.spear-PROP
'We spear kangaroos with a shovel spear.'[14] (Saulwick 2-31-8)

(25) *Nginy-waralh-miny* *guwa* *ngurah* ***da-rdurlh-ga-bœ.***
3>2-ask-PP PURP fire 2>3-catch.fire-$CAUS_3$-INF
'She asked you to light a fire.' (Saulwick 4-1-18)

(26) *Malak* ***ngi-rdurlh-ga-bœ*** *many* *barrppuh* *dawa* *nga-rdurlh-ga-ra.*
NEG 1-catch.fire-$CAUS_3$-IR not.yet soon now 1>3-catch.fire-$CAUS_3$-FUT
'I haven't lit the fire yet, I'll light it soon.' (Saulwick 4-1-20)

14 Note that pronominal prefixes representing lower animates or inanimates need not be marked for number. This explains the use of the singular object prefix in this example.

The semantics of the two infinitival clauses in (23) and (25) are equivalent, despite the fact that the suffix in one instance is homophonous with a present tense form, and in another is homophonous with an irrealis form. If we were not to analyse the infinitive as a distinct inflectional category for Rembarrnga verbs (as McKAY (1975) also does) we would be left with the difficult task of accounting for why it is that verbs belonging to one inflectional class require the present tense suffix to be used in these contexts, while other verbs require the irrealis mood suffix instead. This point is made even stronger by the following example (27), which has two conjoined purposive clauses, one containing a class 1 verb and the other a class 5 verb. An analysis that did not recognise the infinitival suffix as distinct would be required to treat the first subordinate verb as present tense and the second as irrealis mood, an analysis that seems both semantically and pragmatically incongruous.

(27) *Bayarr-walœ-mutti-ya* ***jarl-ø-gan*** *re* ***ma-ngœ-gan*** *me.*
1a>3a-SEQ-show-PP hunt$_1$-INF-DAT animal get$_5$-INF-DAT veg.food
'We also showed them to hunt animals (and) to get vegetable food.' (Saulwick 1-75-5)

The infinitival suffix, therefore, must be analysed as distinct from the other finite tense and mood forms that it is homophonous with. However, for Baker's reanalysis of these infinitival verbs as fully finite forms to be maintained, this infinitival suffix must be encoding some finite tense/aspect/mood category. This category must be compatible with the semantics of purposive and jussive clauses, yet distinct from future tense, present tense and irrealis mood, all of which already exist in paradigmatic opposition to these forms. So, what category is it? Perhaps Baker could argue that it is just a general 'tensed' category that does not specify any further temporal semantics, yet here again we are faced with the fact that the only reason to consider it finite at all is to account for the presence of the agreement prefixes.

On the other hand, if one removes the shackles of Baker's theoretical assumptions, the facts are perfectly clear: these verbs are not tensed at all, they are simply non-finite, infinitival verbs that contain agreement prefixes.

3.2 Infinitives without pronominal agreement

These clauses are identical to those discussed in § 3.1 except that the infinitival verb does not contain any pronominal agreement morphology. The subject of the infinitival clause is usually interpreted as co-referential with an argument of the main clause (28 & 29), but may be supplied by context instead (30).

(28) *Jurla* *nga-ma-ngara* *guwa* *rdom-ø-gan.*
water(NOM) 1>3-get-FUT PURP drink$_1$-INF-DAT
'I'll fetch some water for him to drink.' (McKAY 1975:325 3.6-20a)

(29) *ø-rdarlh-miny* *guwa* *rderh-ga-bœ-gan* *guweny.*
3>3-roast-PP PURP hard-CAUS$_3$-INF-DAT kangaroo(NOM)
'He roasted the kangaroo to harden it.' (McKay 1975:320 3.6-10)

(30) *Narr-appah-jat-yo-ngara* *warikku* *guwa* *ngomh-ø.*
2>3-a-poison-PROG-FUT immediately PURP float.to.surface.dead$_1$-INF
'You two poison the water straight away so that (the fish) will die and float to the surface.' (McKAY 1975:320 3.6-11)

These clause types are most common with third person subjects, as in the above examples, but they are not restricted to these. Examples (31) and (3, repeated from above) show their use with 1st person subjects, singular and augmented respectively.

(31) *Burppa* *rayfl* *nga-ga-nginy* *banh* ***jarl-ø-gan.***
rifle rifle 1>3-take-PP here hunt$_1$-INF-DAT
'I took a rifle to hunt.' (Saulwick 2-39-1)

(3) *Mala-galitj* *yarr-nœtti-ny* *mahgun* ***ma-ngœ-gan.***
all-other 1a>3-keep-PC again get$_5$-INF-DAT
'We used to keep the other half (of bullock coaches) again to get more. (While the others used to go back to the station.)' (Saulwick 1-70-8)

BAKER (1996:474ff) bases his discussion (and subsequent dismissal) of infinitival clauses in Rembarrnga on the following three examples taken from MCKAY (1975:324-5).[15]

(32) *Jurla* *nga-ma-ngara* *guwa* *ga-rdom-ma-gan.*
water(NOM) 1>3-get-FUT PURP NP3>3-drink$_1$-FUT-DAT
'I'll fetch some water for him to drink.' (McKay 1975-325 3.6-20b1)

(33) *Jurla* *nga-ma-ngara* *guwa* *rdom-ø-gan.*
water(NOM) 1>3-get-FUT PURP drink$_1$-INF-DAT
'I'll fetch some water for him to drink.' (McKay 1975-325 3.6-20a)

15 Note that these sentences are translated in both MCKAY (1975) and BAKER (1996) with the main clause in the past tense ('I fetched him some water to drink'). However, the presence of the future tense suffix on the main verb leads us to believe that this was a simple error.

(33) *Jurla* *nginy-bak-ma-ngara* *guwa* *rdom-ø-gan.*
water(NOM) 1>2-BEN-get-FUT PURP drink$_1$-INF-DAT
'I'll fetch you some water to drink.' (McKay 1975-324 3.6-19a)

In (32), the subordinate verb has both tense and agreement morphology, just as it would in a main clause. This is essentially synonymous with (33), in which the subordinate clause has an infinitival verb (homophonous with the present tense form for this verb class) and no agreement morphology. (34) is the same as (33) except that the subject of the subordinate clause is interpreted as 2nd, rather than 3rd person. McKay (1975:324) reports that example (34) was considered ungrammatical by one of his language consultants; whereas both (32) and (33) were accepted more generally. Given this, Baker argues that we can distinguish two dialects of Rembarrnga—one (dialect A) in which infinitival clauses like (32) and (33) are restricted to third person singular subjects, and the other (dialect B) in which they are unrestricted for person.[16] Thus, speakers of dialect A will reject examples such as (34), while speakers of dialect B will accept all three of the examples above.

For the first of these dialects, Baker reanalyses the infinitival clauses as fully finite "tenseless" clauses. In this reanalysis the infinitival suffix is in fact a "tenseless" finite suffix and the third person subject prefix that co-occurs with it has simply been reanalysed as ø. On this reanalysis this dialect of Rembarrnga has two functionally equivalent fully finite purposive clauses: one (32), which is marked with the future tense and has an overt pronominal prefix, and the other (35a), marked with the "tenseless" suffix and a zero pronominal prefix. (35b) shows the same example under our analysis.

(35) a. *Jurla* *nga-ma-ngara* *guwa* *ø-rdom-ø-gan.*
water(NOM) 1>3-get-FUT PURP 3>3-drink$_1$-?TNS-DAT
'I'll fetch some water for him to drink.'

b. *Jurla* *nga-ma-ngara* *guwa* *rdom-ø-gan.*
water(NOM) 1>3-get-FUT PURP drink$_1$-INF-DAT
'I'll fetch some water for him to drink.'

An obvious problem with Baker's analysis of (35a) is the incongruous denotation of a finite clause as being tenseless. A finite clause is by definition one which shows "formal contrasts of tense, number and mood" (CRYSTAL 1992); a "tenseless" clause must therefore be nonfinite. Baker's primary motivation in considering these clauses to be both tenseless and

16 This difference does not appear to exist among our language consultants, who all readily produce and accept such infinitival clauses with non-third person subjects. This may represent a change in the language in the 30 years since McKay's fieldwork. Nonetheless, we discuss this possible dialect anyway in the interests of more fully representing and challenging Baker's position.

finite is the presence of pronominal agreement. However, as discussed in section 1, such agreement is compatible with infinitival clauses cross-linguistically.

The analysis in (35a) is made more problematic by the fact that a zero third person pronominal prefix already exists in Rembarrnga in the past tense, as shown by the contrast between the following past (36) and non-past (37) examples.

(36) *Nihdanda* *ø-bordoh-miny* *gorngorr.*
3masc.pro PT3>3-put-PP billabong
'He put (made) the billabongs.' (Saulwick 2-72-4)

(37) *ga-bordoh-ra* *bag-gah.*
NP3>3-put-FUT bag-ALL
'He'll put it into a bag.' (Saulwick 3-110-167)

Baker's analysis therefore requires that speakers of dialect A have two zero third person pronominal prefixes: one which is used with the past tense inflection, and the other which is used with the "tenseless" suffix. Note that the "tenseless" suffix in this case is homophonous with the present tense suffix, yet the present tense form of the verb is incompatible with the zero third person pronominal prefix, as in (38).

(38) a. **Jurla* *ø-rdom-ø.*
water 3>3-drink$_1$-PRES
'(S)he drinks water.'

b. Jurla ga-rdom-ø.
water NP3>3-drink$_1$-PRES
'(S)he drinks water.'

Thus, according to Baker, speakers of dialect A have generalised the past tense form of a pronominal prefix to the "tenseless" form used in purposive clauses, yet not to the present tense verb with which this form is homophonous.

Secondly, crucial to Baker's reanalysis is the claim that the verb in (35a) is not infinitival at all, but is fully finite, just like that in (32). Thus, this analysis runs into the same problems discussed in § 3.1 concerning the correct characterisation of the (finite) category encoded by the "tenseless" suffix.

Finally, Baker's analysis is further complicated by the fact that McKay also provides a fourth example in which the overt non-past pronominal prefix *ga-*, found with the non-past verb in (32), appears with the infinitival suffix (39). And McKay notes that the speaker who rejected (34) (i.e. the speaker of dialect A) had no difficulty accepting (34).

(39) *Jurla* *nga-ma-ngara* *guwa* *ga-rdom-ø-gan.*
water(NOM) 1>3-get-FUT PURP NP3>3-drink$_1$-INF-DAT
'I'll fetch some water for him to drink.' (McKAY 1975: 325 3.6-20b2)

Thus, the special zero pronominal prefix that Baker needs to postulate for this dialect to account for (34) is clearly only optional, since it is possible for the regular overt pronominal prefix to appear with these infinitival (or "tenseless") verb forms also.

For Rembarrnga dialect B—the one in which these infinitival clause types are possible with a subject of any person category—Baker proposes a reanalysis of nominalization.[17] In this dialect he argues that the infinitival verbs have been reanalysed as nominals so that a sentence like (34), for example, actually means 'I got you some water for drinking'. He cites as evidence for this reanalysis the use of the dative case suffix on the verb in these clauses: since the dative case suffix is also found on regular nominals in main clauses, its presence on infinitival verbs can be interpreted as signalling that this verb has been nominalised (p. 476).

However, although his proposed English translation of (34) is reasonable, there are a number of reasons why this reanalysis is also implausible. Firstly, all speakers of Rembarrnga accept a sentence such as (32) as being fully grammatical. In this example a finite verb, indisputably verbal, is also inflected with the dative case. Thus, the presence of the dative case on the infinitival verb in (34) is not indicative of nominalisation at all.

Secondly, these infinitival verb forms do not exhibit any of the characteristics typical of nominals, including the following.

i. Nominals are frequently optional due to the specification of core arguments by pronominal prefixation; infinitival verb forms however are not omissible.
ii. Nominals can form noun phrases specified by a fairly rich set of demonstratives and further modified by adjectivals; whereas infinitival verb forms cannot.
iii. Nominals can suffix dative/possessive pronominals, numerals, syntactic case as well as other quantificational suffixation (e.g. comparison, 'like X', 'big X', 'small X'). These suffixes are not found with infinitival verb forms.
iv. Nominals do not typically function as predicates without the addition of verbalising morphology, whereas infinitival verb forms always function as predicates.

Furthermore, nominalisation of verb roots in Rembarrnga is restricted to a small set of verbs (see McKAY 1975:195), all of conjugation Class 1, which are nominalised by suffixation of *-yi(na)*. These translate as English adjectives such as, 'be white/light', 'be cool', 'be dry', 'be soft', 'be light in weight'.

---

17 For theory-internal reasons, nominalisations are not required to satisfy the MVC and so do not cause the same problems for Baker's analysis that infinitival clauses do (p. 476).

Thus, despite Baker's suggestions otherwise, there is no empirical evidence to support a reanalysis of these infinitival forms as either fully finite forms, or as nominalisations. We are led once more to the conclusion that they are simply infinitives.

4. Conclusion

BAKER's (1996) theory of polysynthesis, and the theoretical assumptions that lie behind it, require that a polysynthetic language not have infinitival verbs. In this paper, we have challenged this position with data from Rembarrnga, a language with both non-finite verbs and extensive polysynthetic properties. Rembarrnga has two types of non-finite clauses: those with pronominal prefixes and those without. We have shown that both of these verb types in Rembarrnga are clearly infinitival, and that Baker's attempts to reanalyse them as either fully finite verbs or nominalisations have no empirical basis.

Each of these infinitival clause types poses a different problem for Baker's framework. Those without pronominal agreement violate his Morphological Visibility Condition, the central tenet of his theory of polysynthesis, while those with pronominal agreement prove wrong his assumption that nonfiniteness must necessarily require the absence of pronominal verbal agreement. Even outside of the polysynthetic domain, this correlation has been proven false by the existence of person/number-inflected participles and gerunds in many Romance languages (LEDGEWAY 1998, MENSCHING 2000, RAPOSO 1987, VINCENT 1998).

While the absence of infinitival clauses in polysynthetic languages may be a tendency (MITHUN 1984, HEATH 1975), it is not absolute. Such clauses exist quite productively in a polysynthetic language like Rembarrnga, showing clearly that they must be allowed as a possibility in any theory of polysynthesis.

REFERENCES

BAKER, M. C. (1996). *The Polysynthesis Parameter*. New York: Oxford University Press.

BLAKE, B. (1977). *Case marking in Australian languages*. Canberra: AIAS.

CONKLIN, H. C.(1962). Lexicographical treatment of folk taxonomies. In: F. W. HOUSEHOLDER and SOL SAPORTA (eds.) *Problems in Lexicography* (Publication 21, Indiana University Research Centre in Anthropology, Folklore and Linguistics). Bloomington: Indiana University Press. Pp 119-141.

CRYSTAL, D. (1992). *The Encyclopedic Dictionary of Language and Languages*. Oxford. Blackwells.

DIXON, R. M. W. (1980). *The languages of Australia*. Cambridge. Cambridge University Press.

HARTMANN, R. R. K. & F. C. STORK (1972). *Dictionary of Language and Linguistics*. London: Applied Science Publishers.

HAWKINS, J. M. & R. ALLEN (1991). *The Oxford Encyclopedic English Dictionary*. Oxford: Clarendon Press.

HEATH, J. (1975). Functional relationships in grammar. In: *Language* 51:89-104.

LEDGEWAY, A. (1998). Variation in the Romance infinitive. the case of the Southern Calabrian inflected infinitive. In: *Transactions of the Philological Society* 96:1-61.

MCKAY, G. (1975). *Rembarrnga: A language of Central Arnhem Land.* Unpublished Ph.D. Thesis, Australian National University.

MENSCHING, G. (2000). *Infinitive Constructions with Specified Subjects*. Oxford: Oxford University Press.

MITHUN, M. (1984). How to avoid subordination. In: *BLS* 10:493-509.

RAPOSO, E. (1987). Case theory and Infl-to-Comp: the inflected infinitive in European Portuguese. In: *Linguistic Inquiry* 18:85-109.

SAULWICK, A. (In Preparation). *The Semantics of the Verbal Word in Rembarrnga.* Ph.D Thesis, University of Melbourne.

SIMPSON, J.A. & E. S. C. WEINER (1989). *The Oxford English Dictionary, Second Edition.* Oxford: Clarendon Press.

VINCENT, N. (1998). On the grammar of inflected non-finite forms (with special reference to Old Neapolitan), In: IØRN KORZEN and MICHAEL HERSLUND (eds). *Clause combining and text structure*. Copenhagen: Samfundslitteratur. Pp 135-158.

Lexicological and lexicographic problems of word families in Cayuga[1]

HANS-JÜRGEN SASSE

1. Problems with Lexical Units in Polysynthetic Languages

The question of how to write a dictionary of a polysynthetic language has become a major concern of a fair number of linguists involved in lexicographic work in co-operation with native communities on the American continent. Regardless of the individual peculiarities of each case, there is one fundamental question with which all the relevant discussions inevitably end up sooner or later: What is the most adequate basic unit of lexical representation in a language whose word forms often correspond to phrases or clauses and whose roots and affixes correspond to words in those languages on which the established tradition of lexicography is based?

As recently as six years ago CORNILLAC (1993) still recommended designing Eskimo dictionaries in terms of morphemes (roots and formatives + rules to compose words out of

---

1 This paper was inspired by experience from my involvement as an - albeit marginal - advisor in the current Cayuga Dictionary Project on the Six Nations Reserve, Ontario, Canada. I have greatly benefitted from discussions with Carrie Dyck, Michael Foster, Francis Froman, Amos Key, Alfred Keye, and Lottie Keye. The data incorporated in this paper are predominantly based on fieldwork with Evelyn Jacobs, Alfred Keye, and Hubert Skye in 1998, though material resulting from earlier fieldwork (especially with Alta Doxtador, Reginald Henry, and Stanley Skye) has also been worked in. I am indebted to all my Cayuga consultants, without whose keen interest and willingness to share their native tongue with me this work would not have been possible and I apologize for any mistake or misunderstanding which may have distorted their material.

I particularly wish to thank two institutions which have supported my work on Cayuga over the past nine years: the German Research Society for a number of generous grants, and the Woodland Cultural Centre for inexhaustible hospitality without which my stays in Ontario would have meant much less to me.

My sincere thanks also go to two people who commented on preliminary drafts of this paper: Leila Behrens (Cologne) and Nick Evans (Melbourne).

these) rather than in terms of "words". On the basis of the equations "Eskimo morpheme = European word = basic lexical unit" and "Eskimo word = European phrase = productively rule-derived grammatical unit" he concluded that Eskimo is a language "where one does not learn words but learns how to construct words" (1993:208; translation mine). For Iroquoian languages, this is established lexicographic practice, developed in the structuralist tradition in which the scientific leaders of the Iroquoianist community are/were deeply rooted. In particular, it was the influential figure of Floyd LOUNSBURY who propagated the treatment of (Northern) Iroquoian languages in strictly segmentational morphemic terms in his epoch-making "Oneida Verb Morphology" (1953). This procedure, operating with roots, stems, and affixes and defining fixed series of morpheme slots for polymorphemic word forms, has been followed in the analysis of Seneca, Mohawk, Cayuga, and other languages of this group ever since.

Many linguists working in the field of Iroquoian languages have now come to the conclusion that this traditional practice, useful as it may be for comparative purposes, is not particularly adequate for synchronic lexicographic representations in a practical dictionary, for which it has a number of serious shortcomings. First and foremost, it is grammar-based (and, hence, form- rather than meaning-oriented). Second, it usually proceeds from the assumption that the morphology of these languages can be treated in terms of concatenative, agglutinating-type morphology with a more or less perfect one-to-one correspondence between form and meaning. The starting point of a dictionary entry is the root, to which subentries for root extensions, derivations, incorporational compounds, etc., are subordinated. The linguistic motivation for this must be sought in the theoretical foundations of structuralist morphology, including the assumption that the lexical meaning of an expression is tied to its root and that the root therefore constitutes the basic unit of the lexicon. It is also implicitly assumed that the root is the main locus of homonymy or polysemy. If metaphorical semantic extensions or specializations of larger, morphologically more complex forms are taken into account, they are usually represented as subentries with reference to the "original" meanings. Thus, for example, the Cayuga verb base *-ak-ya't-awi-'t-* 'put one's dress on' would usually appear under the root *-awi-* 'be in a tube'. *-awi-'t-* is the CAUSATIVE of *-awi-* and has the meaning 'put sthg in a tube-like space; insert (e.g. into a hole bored into the wall)'. This is a member of a subclass of verbal roots which can be productively compounded with their objects, a process called "incorporation". In the case of *-ak-ya't-awi-'t-* the nominal root *-ya't-* 'body' is incoporated. The whole string is preceded by what Iroquoianists call the "SEMIREFLEXIVE" prefix *-ak-*, here referring to the possessor of the body part, which gives us the meaning of 'insert one's body in a tube-like space'.[2]

---

2 This is the way vocabulary items of Cayuga are usually explained in HENRY & MITHUN (1984). It follows traditional Iroquoianist practice. For example, in the Mohawk vocabulary of G. MICHELSON (1973), the lexical entry is the root, e.g., *-?nas* 'plume, feather, quill'. Then the nominal reference form *ò:nas* 'feather' is given. Compounds are listed as subentries: "with verb root /kvra/: *o?nashakv:ra* 'swan': with verb root /atek/: *o?nashaté:kv* 'Burnt feather' (Personal

Speakers of the languages concerned are extremely unhappy with this kind of representation. They consider it a purely etymological approach which in their view does not have anything to do with the semantic reality of their languages. For the average language user, a "word" is something like *sakyá'tawi't* 'put your dress on!', *ękakya'tá:wi't* 'I'm going to put my dress on'. For practical purposes, then, a root-based dictionary is said by these speakers to be useless: since the root is not accepted as a psychologically real unit of the language, nobody would look up an entry in a dictionary on the basis of its root. Vocabularies or dictionaries based on this principle are therefore usually rejected by the community. The objections frequently raised by language users have to do with their dislike for segmentation and concentrate on the following two points:

(a) Speakers are unable to segment words for morphological reasons. Due to the extensive morphophonemic distortions, roots and affixes are often not formally recognizable. For example, the first person singular of the HABITUAL aspect of the verb 'eat sthg' in Cayuga is *í:ke:s* 'I eat it'. In terms of a concatenative morpheme-based analysis this would be represented as /ke+k+s/ (*ke-* '1st sg.' + root *-k-* 'eat sthg' [disappearing morphophonemically in front of *s*] + *-s* 'habitual suffix'). Speakers comment that nobody would come up with the idea that the root here is *-k-*, that *-ke-* represents the first person singular, and that the increment *i-* is a meaningless extra syllable attached to any word form whose "underlying" structure is monosyllabic. Even for linguistically trained and sophisticated speakers it is often hard to break down a complex word form into its constituent parts, so that the smallest unit accessible to the speakers is a phonological word which is, at the same time, a morphological object capable of being uttered as a minimal sentence.

(b) Speakers are unable to segment words for semantic reasons. There is a large number of morphologically complex word forms with idiomaticized meanings of the type illustrated by 'dress oneself' above. Many of these contain obsolete elements or elements of limited occurrence to which it is difficult to attribute meanings; only "the word as a whole makes sense". Even if such a form is morphologically segmentable, speakers do not readily associate the morphological components of an idiomaticized expression with their "literal" meanings.[3]

---

name)". Cross-reference between entries is usually avoided; information is sometimes doubled, but not systematically: the word for 'swan' also appears under *-kvra* 'white', but 'Burnt feather' does not show up in the entry for *-atek-* 'burn'.

Almost all earlier dictionaries and vocabularies are of a similar make-up. A notable exception is CHAFE (1963), which is a small thesaurus rather than a dictionary, operating exclusively with whole word forms.

3 It could be objected that the speakers' aversion to compositional analyses of figurative expressions is in contradiction with the frequent phenomenon - often commented on in the literature on

From the late 1980s on, the question of an adequate lexicographic format for Iroquoian languages has become a hotly disputed topic at the annual Iroquoianist meetings at Rensselaerville near Albany, N.Y. The Iroquoianist community began to understand that relying entirely on roots as lexical entries was insufficient and therefore developed the idea of the "base dictionary" (cf. discussions in FOSTER / K. MICHELSON / WOODBURY 1989, K. MICHELSON 1990; an early examplification of this new Iroquoian dictionary format is RUDES (1987), the first dictionary completed is RUDES (1999); the Oneida dictionary by MICHELSON / DOXTATOR is about to follow[4]). What this means is that not only roots but also compounds and derived verbs and occasionally even entire words are now represented with separate entries in the dictionary. The new strategy was guided by the insight that these larger complexes are often semantically non-compositional (cf. BONVILLAIN 1989 and especially K.MICHELSON 1990, who deals at some length with the problem of lexicalization, raising points similar to those made in this paper and proposing criteria for identifying lexicalized units.). This was accounted for by the postulation of a new fundamental lexical unit, the "base", which is defined as "a unit whose meaning and/or form cannot be derived synchronically" (FOSTER / K. MICHELSON / WOODBURY 1989:2).

The "base dictionary" adequately captures the fundamental lexical categories of the language in accrediting them separate entries: nominal roots (*-ya't-* 'body'), derived nominal stems (*-akya'tawi'thr-* 'dress'), verbal roots (*-awi-* 'be in a tube'), derived verbal stems (*-awi't-* 'insert'), compound roots (*-akya'tawi't-* 'put a dress on'), lexicalized fully inflected verb forms (*kaǫtanę́hkwih* 'horse' [< 'it hauls logs']), and uninflectives (*só:wa:s* 'dog') (cf. SASSE 1993). There can be no doubt that this device is much more realistic than the earlier root-based approach since now *-awi-* 'be in a tube', *-awi't-* 'insert', *-akya'tawi't-* 'put a dress on' and *akya'tawi'thra'* 'dress' all have their own lexical entry, though the reference system in the dictionary format proposed still uses the root as the general target of reference (e.g. one would have cross-referencing remarks such as "see *-ya't-* 'body', see *-awi-* 'be in a tube'", etc.), and the root is still considered the base par excellence ("what have traditionally been called roots in Iroquoian are by definition bases" (FOSTER / K. MICHELSON / WOODBURY 1989:2)).

Northern Iroquoian languages - that speakers readily volunteer "etymologies" of descriptive terms (such as "*tekaihstohá:kǫh* is the newspaper, what it really means is 'it is pressed by the metal'"). However, these "semantic analyses" are likely to be later linguistic abstractions offered for the benefit of the eliciting linguist and are probably confined to the fieldwork context. They do not necessarily prove that there always is an intimate connection between compositional and non-compositional senses of multimorphemic expressions in the mental lexicon. The fact that "literal" meanings may be associated with parts of utterances - even bound ones - does not imply that speakers are able to isolate or reconstruct roots and perceive them as the basic organizational units of their lexicon.

4 Both dictionaries make independent entries of roots and complex stems derived from these roots, which are called 'bases' (incorporated noun plus verb and/or derivational cotegories such as causatives, etc.)

Unfortunately, these new lexicographic proposals have not made the language users much happier. It appears that speakers ultimately aspire to a full-word lexicon, in which "words" in the sense of inflected word forms constitute the basic lexical units. This means that they do not content themselves with a lexical stem representing a conventional label for a paradigm, but aim at a lexical format in which each form in the paradigm is represented by a separate lexical entry. In other words, speakers insist on the extreme opposite position (morphological word = word form = lexical entry). How to reconcile these two positions?

The problem with root dictionaries has resided in an unbridgeable clash between the interests and the language user and methodological backgrounds of the classical Iroquoianist (and the structural linguist in general). Iroquoianists are, by their tradition, ethnolinguists with a combined interest and background in historical-comparative linguistics and structuralist morpheme analysis. This led to their assumption of huge word families clustering around a root. The morphological richness and the abundance of compositionally transparent derivations and compounds seemingly supports this assumption. From a lexical semantic point of view, however, which is obviously also the point of view of the speakers, these word families are historical word clans clustering around an ancestor root, which is hardly recognizable as a core element in all their members.

Now that this position is given up in the new concept of the "base" dictionary, the way has been paved for a compromise, provided that a number of open questions with respect to the status of words and stems in these languages can be solved. In contrast to what Cornillac assumes for Eskimo, it is not true for Cayuga (and for all the other members of the Northern Iroquoian group) that one does not learn words but learns to make words by combining smaller bound units. If this were the case, speakers should be expected to have no difficulties in recognizing morphemes as meaningful parts of words. In actual fact, however, they vehemently deny that anything which has a dash in front or at the end is a valid representational unit of their language. Casual conversations as well as formal interviews over the years of my research in Six Nations revealed that speakers unanimously claim to have learned full word forms, which they associate with the situations in which they have acquired them. The retrieval of lesser used forms of the paradigm is therefore often considered difficult. After a moment's thought speakers first come up with a decision on whether or not they have ever heard this form. Only when they think they have not heard it are they ready to construct it themselves, and then they do this by analogy to parallel forms of similar paradigms rather than by putting together "morphemes". Should there be no model for the analogy, speakers not seldomly refuse to construct the form altogether. These observations cast considerable doubt on the correctness of the assumption of morpheme-based word-formation processes and give rise to the question of whether a full-word dictionary would not in fact be more adequate for this type of language, in that it would perhaps come closer to representing the "mental lexicon" of the speakers.[5] But, one may

5 One might speculate on whether these observations reveal a true typological fact about the language or whether they must be attributed to the situation of language obsolescence with its

justly object, how is a full-word dictionary practically feasible in view of the immense number of forms available? And how to represent the productive strategies in such a dictionary?

Before turning to some general considerations for a possible solution, we will present a case study demonstrating how families of word forms are organized on the different levels of historical abstraction, transparency, and compositionality. We will illustrate this with the word clan 'be in water' in Cayuga, which would normally be presented in traditional structuralist approaches as a single lexical entry. In a "base dictionary" it would probably cover a dozen or so separate entries, linked to each other by a reference system. This is not a very spectacular example. There are numerous much larger word clans, but the one considered here is large enough to illustrate the principal phenomena involved and the problems they pose for lexicographic work.

2. The Cayuga Word Clan *-o-* 'be in water': A Brief Case Study

2.1 Occurrences of *-o-* as a Non-derived Base

The story begins with a single vowel, *-o-*, which is traditionally taken to be the root for 'be in water'. From the point of view of the synchronic morphological verbal system of the language it is, in the first place, the inflectional base of at least two quite different morphological verbs, a state verb with the meaning 'be in a liquid; be in the process of boiling', which occurs only in the STATIVE aspect (*-o-*$_1$), and a dynamic verb 'cook; soak; fish', which occurs in all three aspects (*-o-*$_2$).

Let us deal with the state verb (*-o-*$_1$) first. It is a member of verb class H3 according to the morphological verb classification in SASSE AND KEYE (forthcoming), takes so-called agent prefixes, and occurs only in the 3rd person singular neuter.[6] The key forms of the simplex verb's tense paradigm are PRESENT *í:yo'* 'it is in a liquid', PAST *í:yohk* 'it was in a liquid', FUTURE *ę́:yo:k* 'it will be in a liquid'. The remainder of the inflectional paradigm can be

well-known "desuetude" effects. I have no ready answer to this at the moment; it should be noted, however, that all of the consultants on whose linguistic behavior and intuitions the above statements are based are fluent speakers of the language.

6 Categories constituting the pronominal prefix paradigm include three persons (first, second, third), three numbers (singular, plural, dual, the latter being distinguished only in the first and second persons), inclusive/exclusive distinction in the first person dual and plural, and three genders (masculine, feminine, neuter) distinguished in the third persons singular and plural. This gives us 14 forms to the core paradigm: 1s, 2s, 3sm, 3sf, 3sn, 1pe, 1pi, 2p, 3pm, 3pf, 3pn, 1de, 1di, 2d. For details on verb classes, prefix sets, etc. cf. SASSE AND KEYE (forthcoming).

regularly constructed on the basis of these forms. The verb may occur with and without an incorporated noun root. When compounded with a noun root, the forms are *ka-* (or one of its "allomorphs") '3rd sg. neuter agent' + N root + *-o'* for the PRESENT, *ka*-N root-*ohk* for the PAST, *ę-ka*-N root-*o:k* for the FUTURE. The meaning does not convey any information on the 'manner of being in the liquid'. For example, the entity referred to by the subject prefixes of this verb may float on the surface, or may be covered by the liquid, like, say, a piece of sugar at the bottom of a cup of coffee, a fish swimming in the river, or a rusty bike lying in a pond. But note that it is a verb of position rather than of motion; in fact, it is one of two dozen or so position verbs used for the positional classification of incorporated nominal expressions. To turn *-o-* into a verb of motion, you have to use a derivational suffix *-kye'*, which is referred to as the PROGRESSIVE in Iroquoianist literature and which has the basic meaning 'move along progressively doing something'. We thus have *okye'* 'it moves along floating'; when combined with the CISLOCATIVE prefix it yields *kyokye'* 'it's floating towards me, comes floating'; when combined with the TRANSLOCATIVE *heyó:kye'* 'it's floating that way, it's floating away', etc.

Getting back to the positional *-o-*$_1$, we have to state, first of all, that it is not entirely adequate to say that *-o-*$_1$ just means 'be in a liquid'. *-o-*$_1$ has developed a number of different senses and subsenses which are etymologically – and in part also synchronically – related. These are illustrated in (1) and (2) below. It is evident that at least two distinct major senses must be established. One of them can be glossed as 'be in a liquid/in water; float'. Examples are given in (1).

(1) *-o-*$_1$ / Sense 1: 'be in a liquid/in water; float'

(a) as a simplex:

í:yo'	'it is in the water, floats'

(b) with productive incorporation:

wahkwę́:nyo'	'(the) clothes (are) in the water'
ka'nhǫ́hso'	'(the/an) egg(s) (is/are) in the water'
ka'wáho'	'(the) meat (is) in the water'
etc.	

(c) with additional "derivational" suffixes:

kakso'	'(the/a) dish (is) in the water'
kaksó:nyo'	'dishes (are) in the water' (DISTRIBUTIVE)
kanráhto'	'(the/a) leaf (is) floating in the water'
onráhtokye'	'(the/a) leaf (is) floating along' (PROGRESSIVE)
kahǫ́na'to'	'potatoes (are) in the water'
ohǫna'tó:kye'	'potatoes (are) drifting along' (PROGRESSIVE)

(d) with alternation of agent and patient prefixes:

kahnyǫ́hso'	'(the) squash (is) in the water' (agent)
ohnyǫ́hso'	'(the) squash is put in water' (patient)
kanę́ho'	'(the) corn (is) in the water' (agent)
onę́ho'	'(the) corn is put in water' (patient)
kawí:tro'	'pieces of ice in the water' (agent)
owí:tro'	'ice cubes' (patient)
katsyǫ́'to'	'(the) fish (is/are) in the water' (agent)
**otsyǫ́'to'*	'(the) fish (is/are) put in water' (patient)[7]

(e) with different types of opaqueness:

kahǫ́:yo'	'(a) boat/ship (is) in the water' (formally aberrant)
kanyóhsro'	'ferry' (opaque)
ohné:ko'	'water (is) in the pond' (non-compositional, no productive sense)

(f) idiomaticized verb forms:

tháǫ:to'	'Toronto' (< 'he has a log in the water over there', cf. *káǫ:to'* '(a) log (is) in the water')
kaná'tsyo'	'Ottawa' (< 'pail in water')

The optional articles and copula forms given in parentheses in the English translation are meant to indicate that these forms are, in principle, completely unspecific with respect to definiteness or with respect to a situational ("verbal") or an object ("nominal") reading.

Sense 2 may be glossed as 'be in the process of cooking in boiling water or broth'. A significant formal difference between this and *-o-*$_1$ / sense 1 is found in the fact that only in the first sense is it possible to use the verb without an incorporated noun stem; sense 2 'be in the process of cooking' never occurs without an incorporee (hence the "plus" sign instead of the dash in front of the root, a conventional symbol proposed in recent Iroquoian lexicography to indicate verbs that occur only with incorporees). Examples are given in (2) below.

---

7 We will use the asterisk here to indicate forms that are not accepted in the sense under discussion. This does not preclude the acceptibility of these forms in a different reading.

(2) *+o-*$_1$ / Sense 2: 'be in the process of cooking in boiling water or broth'; only with inc.):

kanráhto'	'leaf is boiling'
ka'nhǫhso'	'egg(s) is/are boiling'
kahǫna'to'	'potatoes are boiling'
ka'wáho'	'meat is cooking'
kahnyǫhso'	'squash is boiling'
kanę́ho'	'corn is boiling'
katsyǫ'to'	'fish is/are cooking'
wahyo'	'fruit are cooking'
etc.	

A closer look at the examples in (1) shows that, even in the totally transparent examples in (b), (c), and (d), there are significant differences in the inflectional potential and the associated meanings of each "base". First of all, it appears that the translation of the verbal part varies with the meaning of the nominal incorporee, e.g. *kakso'* 'dish is in the water' vs. *kanráhto'* 'leaf is floating'. These differences in translation should probably not be taken as constituting evidence for distinct sub-senses but can be inferred pragmatically from the semantics of the incorporees under the assumption of a general sense of *-o-*$_1$: since *-o-*$_1$ is "vague" with respect to a manner component, the differences are to be attributed to the different physical nature of the incorporees – heavy things sink under the surface while light things float on the surface. Nevertheless, speakers are often very decisive about the "correct" translation and it may well be that these perceptual differentiations lead to sub-senses of *-o-*$_1$ *in statu nascendi*, first triggered by the context of the incorporee but later potentially becoming independent.

Second, we note that there are certain inflectional changes associated with certain semantic effects, which are limited to a number of tokens of the pattern. For example, some speakers claim that a shift from the agent prefix to the patient prefix of the 3rd person neuter (*ka-* to *o-*) changes the meaning from 'x is in water/liquid' to 'x has been put into water/liquid' (cf. *kanę́ho'* vs. *onę́ho'*, etc.). This was accepted only for a small subset of instances, not, for example, for *otsyǫ́'to'* *'fish are put into water'. These forms look like STATIVES of a dynamic verb *-o-*$_3$ 'put into water', which does not exist, however, as a full paradigm.

Furthermore, within the examples of sense 1 we observe different degrees of compositional transparency. The patient form corresponding to *kawí:tro'* 'pieces of (natural) ice (are) in the water', *owí:tro'*, is semantically conventionalized to 'ice cubes' (< 'ice put into the water'). The word form *káhǫ:yo'* 'a boat/ship in the water' contains an idiosyncratic "stem allomorph" *-hǫy-* of the root for 'boat/ship', which is otherwise *-hǫw-*. The word *kanyóhsro'* 'ferry' is totally opaque; the element *-nyohsr-* does not occur in any other combination. The word *ohné:ko'* 'water in the pond' should literally mean 'liquid put into liquid' (*-hnek-* is the nominal root for 'liquid'). Since this doesn't make any sense, the form

is non-compositional; it is morphologically analyzable but semantically non-transparent. And so forth. At the bottom of (1) we arrive at a number of examples of a very prominent and productive lexical category of Iroquoian languages which may be called "conventionalized verb forms", i.e. forms taken out of the productive paradigm and used in a new sense, in the present case as place names: *tháǫ:to'* 'Toronto' ('he has a log in the water over there'), *kaná'tsyo'* 'Ottawa' ('a pail is in the water'). More on conventionalized verb forms in section 2.3 below.

We will now briefly consider sense 2 'be in the process of cooking in boiling water or broth'. We have already stated that *-o-*$_1$ may be read in this sense if and only if it has an incorporee. We can now further specify that this incorporee has to be a member of the semantic field of entities usually prepared for consumption by means of boiling (BOILABLES-FOR-CONSUMPTION). We thus have *kahǫ́na'to'* 'potatoes are boiling', *ka'nhǫ́hso'* 'egg(s) is/are boiling', etc., but not *wahkwę́:nyo'* '*clothes are boiling', *kȧkso'* '*a dish is boiling', etc. The semantic potential of the collocations seems to be rather strictly governed by this principle; anything that can be conceived of as BOILABLE-FOR-CONSUMPTION may fall under this rubric, while for any other things that may be put into boiling water other verbs have to be used (incidentally, 'leaf' is not an exception since leaves are boiled for medicine).

From a historical point of view, is seems quite clear that sense 2 'be in the process of cooking in boiling water or broth' is a semantic specialization of sense 1 'be in a liquid', which was narrowed down in the context of BOILABLE-FOR-CONSUMPTION incorporees. This has not, however, led to the blockage of the original combinations with *-o-*$_1$, in the first sense. Hence one encounters a considerable number of ambiguous forms, such as *kanráhto'* '1. leaf is floating, 2. leaf is boiling', *kanę́ho'* '1. corn is in the water, 2. corn is boiling', etc.

We will now turn to the dynamic verb *+o-*$_2$, exemplified in (3). It belongs to verb class H3 in the Sasse/Keye classification and possesses a full paradigm of three aspect stems, HABITUAL *-oha'*, PUNCTUAL *-o:'*, STATIVE *-o'*, of which altogether 14 tense-aspect-mood paradigms can be formed. In combination with the 14 "pronominal prefixes" referred to in fn. 5 this yields a paradigm complex of 196 word forms.

(3) *+o-*$_2$ (dynamic verb H3 [*-oha'/-o:'/-o'*])

(a) Sense 1: 'cook (in boiling water or broth)'; productive with inc.; without inc. a different verb must be used in the sense of 'boiling'

ękénrahto:'	'I'll boil leaves'
konráhtokye'	'she is going along boiling leaves, is still...'
ękhǫ́na'to:'	'I'll boil potatoes'
kohǫna'tó:kye'	'she's going along cooking potatoes, is still...'
ęké'nhǫhso'	'I'll boil eggs'

ęk'wáho:'	'I'll boil meat'
ękéhnyǫhso:'	'I'll boil squash'
akéhnyǫhso'	'I have boiled squash'
ohnyǫ́hso'	'it (animal) has boiled squash'
ękhnę́ho:'	'I'll boil corn'
akhnę́ho'	'I have boiled corn'
ękétsyǫ'to:'	'I'll cook fish'
akétsyǫ'to'	'I have cooked fish'
ękáhyo:'	'I'll cook fruit'

(b) Sense 2: 'soak dry food' (with inc.; limited occurrences)

ękéthe'thro:'	'I'll soak flour'

(c) Sense 3 (+SEMIREFLEXIVE): 'fish by means of incorporee' (with inc.; limited occurrences)

ękáta'ao:'	'I'll fish by net'
ękátahnyo:'	'I'll fish by hook' (evolving into a general term for 'to fish')

The dynamic verb $+o\text{-}_2$ is not attested as a simplex. It has at least three senses: 1. 'to cook sthg (= incorporee) in boiling water or broth'; 2. 'to soak dry food (= incorporee); 3. 'to fish by means of incorporee'.

In the first sense, it is exactly parallel to $+o\text{-}_1$/sense 2, and can be considered the dynamic agentive counterpart thereof. In other words, any compound that allows the 'boiling' reading exemplified in (2) can also occur as a dynamic cooking verb. It stands to reason that $+o\text{-}_2$ / sense 1 is simply a dynamic extension of $+o\text{-}_1$/sense 2. Nevertheless, it is clearly a different unit, which is proven by the fact that its STATIVE is not semantically identical with the STATIVE-tantum verb $+o\text{-}_1$/sense 2. While the latter has a processual reading ('x is boiling'), the former has the typical resultative-perfect reading of terminative verbs ('someone has boiled it'). This leads to another (albeit marginal) set of potential ambiguous forms in the third person singular neuter. For example, when we go down the paradigm of the STATIVE *akéhnyǫhso'* 'I have boiled squash', we arrive at the 3rd person neuter *ohnyǫ́hso'* 'it (animal) has boiled squash', which is not a very common reading, to be sure, but not totally nonsensical, for example, in the context of a fable. The same form was given in (1) with the reading 'squash is put in water'. With the "PROGRESSIVE" *-kye'* even triples become possible, such as *onráhtokye'* '1. leaf is floating along; 2. leaf continues boiling; 3. it (animal) is going along boiling leaves'.

Senses 2 and 3 of *+o-*$_2$ are limited to very few conventionalized cases. Both are semantic extensions of an obsolete sense 'put into water', which shows its vestiges in forms like *oę́ho'* 'corn is put into water' discussed above. The etymological meaning of the 'fishing' verbs is 'put the net, hook, etc., into the water', but these verbs are both semantically and formally opaque (the elements *-a'a-* and *-ahny-* are not readily associable with the meanings 'net' and 'hook' for the average speaker).

To end this part of the discussion, it is of some interest to re-examine the arrangement of the forms in the lists of (1), (2) and (3). It is immediately evident that this arrangement, which is based on the morphological make-up of the verbal system and attributes lexical semantics to single roots and bases, does not really do justice to the semantic connections we have established so far. We have seen that the development of new senses is not so much a matter of the verbal roots themselves as it is triggered by the semantics of the incorporees. In other words, it appears inadequate to attribute the sense differentiations to the morphological verbs *-o-*$_1$ and *+o-*$_2$. Rather, what happens is that the semantic groups in which the different forms can be allocated constitute a large network of alternations (static-dynamic, agentive-patientive, etc.), in which semantically defined patterns of compounds rather than single verb roots are involved. This leads to a considerable number of related sets, each with its individual "mini-grammar", which cut across the general morphological template. Any incorporee may change the sense of the host verb; and any N+V compound may as a whole serve as the point of departure for semantic extensions via metaphor, etc., i.e. for the creation of new non-compositional senses. Moreover, it does not seem to make much sense to distinguish between "syntactic" (i.e. productive compositional) and "lexical" (conventionalized non-compositional) incorporation; these are gradual rather than discrete phenomena. Some of the apparent non-compositional incorporations may open productive new channels (as in the BOILABLE-FOR-CONSUMPTION cases), others constitute small closed sets (such as the 'fishing' verbs), still others remain singular non-compositional occurrences (such as *ohné:ko'* 'water in the pond').

2.2 Derivations of *-o-*

Let us now turn to a set of forms which is traditionally considered to be the set of verbal derivations of *-o-*.

The first to be discussed is the so-called INCHOATIVE, traditionally described as being formed with a suffix *-'-*, which yields *-o-'-*, or rather *+o-'-* given that this verb occurs only with incorporees. The meaning of this complex is 'fall into the water (of an inanimate entity = incorporee)'. This meaning is not predictable from the semantics of the INCHOATIVE suffix. All compounds with *+o-'-* attested so far are entirely compositional, cf. (4):

(4) "INCHOATIVE" +*o-'-* 'fall into the water (of inanimate entities):

-hǫna't-o-'- : *hękahǫná'to'*	'the potato will fall into the water over there'
-nraht-o-'- : *ękánrahto'*	'the leaf/leaves will fall into the water'
etc.	

The next formation to be considered is *-o-h-*, i.e. *-o-* with a suffix *-h-*, an opaque element with CAUSATIVE meaning. *-o-h-* occurs as a simple verb in the meaning 'put into liquid, dip' and belongs to conjugation type S5 in SASSE & KEYE (forthcoming). It is always combined with the TRANSLOCATIVE ('thither') preverb *he-*: HABITUAL *hé:kohs* 'I put it into a liquid', punctual future *hę́:koh* 'I will put it into a liquid', STATIVE *hewákohǫh* 'I have put it into a liquid'.

The verb *-o-h-* allows productive incorporation:

(5)

-hǫna't-o-h : *hękhǫ́na'toh*	'I'll put potatoes into the water'
-at-hs-o-h- (+SEMIREFLEXIVE): *hęká:tshoh*	'I'll put my mouth into the water'
-at-ahsi't-o-h- (+SEMIREFLEXIVE): *hękatahsí'toh*	'I'll put my feet into the water'
-at-hs'oht-o-h- (+SEMIREFLEXIVE): *hękáts'ohtoh*	'I'll put my hands into the water'
etc.	

In addition, one finds a large number of separate compounds based etymologically on this verb, with all sorts of constellations from near-compositional to totally opaque:

(6) (a) Compositional though with unpredictable preverb:

-at-rę́nawę́'t-oh- (+SEMIREFLEXIVE, + DUALIC):

tękatrę́náwę́'toh 'I'll put sugar in liquid (coffee, tea, etc.)'

(b) Non-compositional though motivated:

-at-rihst-oh- (+ SEMIREFLEXIVE) 'trap' (< 'put metal in water for oneself')

(c) Non-compositional though partly motivated:

-a'kęh-oh- 'load a firearm' (= 'put ashes or gunpowder in water')

(d) Totally opaque:

-ahs-oh- 'paint' (*-ahs-* '???')

-'sk-oh- 'drown, fall into the water (animate)' (*-'sk-* '???')

From the synchronic point of view it is clear that something like *-ahs-oh-* is an incorrect representation. *-ahsoh-* is now a monolithic verb root with the meaning 'paint'. This is evidenced by the fact that it can in turn serve as a basis for productive incorporation, which proves that *-ahs-* is not felt to be an incorporated element any longer (double incorporation is not permitted). But this is equally true for cases such as *-at-rihst-oh-* and *-a'kęh-oh-*. These are not to be analyzed as compositional, although they are formally transparent and speakers may be able to interpret them "literally" upon request. Speakers usually claim that the use of these forms in the "literal" meaning is blocked so that they are non-ambiguous, but I would not be surprised if a speaker came up with a context where he/she could use verbs like these in the compositional reading.

Let us move on to (7), the regular REVERSIVE ('undoing') formation *-o-kw-* 'take out of the water':

(7) *-o-kw-* ("REVERSIVE") 'take out of the water':

(a) Simplex (verb of class S10); key forms:

HABITUAL *kokwahs*	'I take it out of the water'
FUTURE PUNCTUAL *ękó:ko'*	'I'll take it out of the water'
STATIVE *akó:kwęh*	'I have taken it out of the water'

(b) With incorporation (chiefly compositional):

akwi:tró:ko'	'I took the ice out of the water'
akenrahtó:ko'	'I took the leaf/leaves out of the water'
ak'wahó:ko'	'I took the meat out of the water'

Like the preceding, this stem takes the full aspectual paradigm of dynamic verbs (with its 196 person/tense/aspect forms); it belongs to conjugation S10 of the proposed morphological verb classification and occurs both as a simplex and with incorporated noun stems: HABITUAL *kokwahs* 'I take it out of the water', PUNCTUAL FUTURE *ękó:ko'* 'I will take it out of the water', STATIVE *akó:kwęh* 'I have taken it out of the water'; with incorporated *-witr-* 'ice' *akwi:tró:ko'* 'I took the ice out of the water', etc. The attested compounds are chiefly compositional, with the exception of *khnékokwahs* 'I take the water out of a pond, empty a pond, get it dry', which is based directly on *ohné:ko'* 'water in the pond'.

Further derivational categories, which are productive and occur with all of the verbs mentioned so far are the following:

(8) "DISTRIBUTIVE" ("several subjects/objects/actions involved"); all tokens attested are entirely compositional;
"PROGRESSIVE" ("going along doing action; action/state still ongoing"); chiefly compositional; several metaphorical examples;

"DISLOCATIVE" ("go to do the action"); all tokens attested are entirely compositional;
"BENEFACTIVE" ("do the action for or against s.o."; only with action verbs); all tokens attested are entirely compositional.

By contrast, there are a number of word-formation processes of severely limited occurrence. For example, the addition of a regular CAUSATIVE suffix is restricted to *-okw-* 'take out of water', which yields *-okw-aht-* 'use sthg to take sthg out of water'. This, in turn, may be input to further derivations. Metaphorical sense extensions are possible at every derivational step.

In sum, the examination of the "derivational" categories complicates, but does not significantly alter the picture obtained from the examination of the "families" that cluster around the simple *-o-* stems. Again, the idea that the meanings of all these expressions evolve hierarchically from the ancestor root and are based on the verbal stems of the derivatives seems inadequate.

2.3 Conventionalized Verb Forms (CVFs)

This term is used to refer to the lexicalization of verb forms. One of the inflected forms in the paradigm receives a new sense, by which process ambiguity arises between the compositional reading of the productive word form and a non-compositional reading. The creation of CVFs is one of the most productive strategies of lexical enrichment in Iroquoian languages[8]. CVFs are usually in the 3rd person PRESENT of the HABITUAL aspect:

(9)	*ha-na'tá-ǫni-hs*	1. 'he habitually makes bread'
	he-bread-make-HABITUAL	2. '(he is a)(male) baker'
	e-hyátǫ-ha'	1. 'she habitually writes'
	she-write-HABITUAL	2. '(she is a)(female) secretary, writer'
	te-ká:-tęh	1. 'it habitually flies off'
	DUALIC-it-fly:off:HABITUAL	2. 'plane'

In principle, however, any verb form may undergo this process:

(10)	*t-kat-í'trǫ'*	1. 'the animals live there'
	there-they(neuter)-live:STATIVE	2. 'stable'

8 CVFs, like all other non-nominal forms with object-denoting senses, may in turn be input to a process that Iroquoianists call "nominalization". It consists of adding one of a number of "nominalizing" suffixes (*-'t-, -'thr-, -ht-,* etc.) to the stem. Extended nominal stems so formed may be incorporated and thus yield the point of departure of a new cycle of derivations.

hya-yé:na:	1. 'let him catch you!'
he:you-catch:IMPERATIVE	2. 'tag' (game)
ka-nǫ́hǫnyǫhk	1. 'let there be thanks!'
it-thank:STATIVE:IMPERATIVE	2. 'thanksgiving speech'
shǫkwa-ya't-íhs'ǫh	1. 'he created our bodies'
he:us-body-make:ready:STATIVE	2. 'the Creator'

A considerable number of CVFs are formed on the basis of the material discussed so far. We have already referred to the place names for 'Toronto' and 'Ottawa' in section 2.1. For reasons of space, we will confine ourselves to only a few more examples.

The 3rd person forms of the HABITUAL PRESENT of *-ahsoh-* 'paint' may receive an "occupational" reading 'be a painter'; inserted into the appropriate syntactic argument slots they may simply translate the nouns '(male/female) painter(s)'. Gender and number distinctions are made possible by the respective changes in the pronominal prefixes: *hahsohs* '(he is) a painter', *ǫhsohs* '(she is) a painter', *hęnáhsohs* '(they (men) are) painters', etc.

Similarly, the 3rd person feminine singular of the causative of *-ahsoh-* + incorporated *-kah-* 'eye', *ǫtkahahsóhtha'* 'she uses it to paint her eyes with it', is used as a CVF in the sense of 'eyeliner'.

Occasionally, the CVFs may cluster in productive patterns with a small but regularly exploited input domain. A good example of this is the formation of words denoting different kinds of cooking pots. These are CVFs based on the 3rd person singular HABITUAL PRESENT of the "INSTRUMENTAL" derivation ('use to do something') of the 'cooking/boiling' verbs and can be formed from any one of these verbs:

(11)	*-'wah-o-*	'cook meat'
	-'wah-o-hkw- ("INSTRUMENTAL")	'use to cook meat'
	e'wáhohkhwa'	1. 'she uses it to cook meat with'
		2. 'pot for cooking meat'

3. Conclusions

The traditional attempt to analyze formally related sections of the vocabulary in terms of semantically more or less transparent sequences of roots, stems, and affixes (a practice frequently employed in the past) has turned out to be inadequate from a lexical point of view. Rather, portions of the vocabulary based on a single root tend to evolve into an extremely complicated array of different semantically unrelated "word clans" and "word

families" in such a way that the idea of an "ancestor root" acting as some sort of synchronic semantic core expanded in a step-by-step fashion soon loses its explanatory power. Methods of lexical enrichment usually move back and forth between different lexical levels. The mechanisms involved are not confined to overt derivational processes and incorporation. Extended senses constituting the input for novel morphological formations also play an important role. Such extensions are often triggered by the semantic content of incorporated noun roots, i.e. by the semantic implications to which a specific noun+verb compound gives rise (as in the case of the cooking verbs, where the new sense of cooking was developed on the basis of the incorporation of food-denoting nominal roots). Moreover, lexicalization of inflected verb forms (CVFs) can be identified as a specific type of lexical enrichment characteristic of Northern Iroquoian languages, which opens up new channels for the creation of new word families. On the other hand, we have an extremely rich inventory of morphology creating semantically compositional forms. There are productive inflectional paradigms and verb classes which define the morphological exponents of tense, aspect, mood and person inflection with great regularity.

Given the complex semantic relationship among the various word forms and their correlations with a host of different types of grammatical behavior, it is understandable that speakers tend to play down the productive paradigmatic regularities of their language and come to the conclusion that only a full-word dictionary is in a position to help them out of the dilemma. But apart from the fact that such an enterprise is not practically feasible in view of the enormous number of productive morphological devices[9], it would obscure the automatic morphological processes of the language. It would not make sense to represent both the first and the second person singular of every verb as separate headwords in a dictionary, given the fact that the second person can almost always be constructed on the basis of the first by simply changing *k-* to *s-* (or vice versa).

A feasible compromise between the desire of the speakers to work with full word forms and a linguistically adequate dictionary format can be envisaged in the form of a computational lexical data base with the following characteristics:

9 Recall that 196 forms are possible for the person-inflected core paradigm of one single aspect stem (in all tense and mood forms) of an intransitive or non-human-object transitive verb. When non-third person pronominal prefixes of transitive human-object verbs are taken into account, the sum of possible forms rises to around 850. For the complete paradigm of the three aspects this has to be multiplied by three (2550). This sum can be multiplied many times over, given that the complete set of paradigms can be run through with a considerable number of productive prepronominal prefixes such as the cislocative, translocative, negation, and so on and so forth.

1. It should proceed from the concept of a "base dictionary" as described in section 1, listing all non-transparent stems and all "units whose meaning and/or form cannot be derived synchronically", including all CVFs, but also many transparent forms such as CISLOCATIVES, TRANSLOCATIVES, etc. A lot of work still needs to be done to figure out what constitutes a "base", since subtle facets of sense distinctions and their repercussions on the grammatical behavior (e.g. aspect system, paradigmatic gaps, etc.) are hardly investigated. But this does not in principle constitute an obstacle.

2. Word forms based on entirely productive morphological mechanisms such as inflection for person should not be given separate status as lexical units. Instead, the lexical data base should be expanded by a morphological component generating all regular word forms of the paradigm of a "base".

3. In order to avoid stem representations with dashes, reference forms should be defined for each lexical category (e.g. 3rd person neuter for nominal bases and state verb bases, perhaps 1st person singular for action verbs, etc.).

4. The data base should have a full-word component listing commonly used word forms of "bases", from which links are established to a corpus of examples or texts which help in understanding the contextual usage of these forms.

As a final point one might raise the question of how far a polysynthetic language like Cayuga actually raises theoretical problems of analysis different from those familiar from European languages. While dictionaries of moderately synthetic languages such as German or Russian avoid going back to the root level, they are, in principle, confronted with similar problems on the word level, which continues to be considered the basic lexicographical level for these languages. This leads to severe inadequacies of semantic representation, which have often been discussed but still largely underestimated in the past, such as the representation of idioms, complex predicates (phrasal verbs, light verb constructions, serial verbs), etc. Since polysynthetic languages show up the inadequacies of the compositional model of linguistic meaning in a more extreme way, they would constitute a good starting point for a general discussion of the proper level of lexical representation – but this would be the topic of a different paper, so we will leave it at that here.

REFERENCES

BONVILLAIN, NANCY 1989. "Noun Incorporation and Metaphor: Semantic Processes in Akwesasne Mohawk". *Anthropological Linguistics* 31:173-194.

CHAFE, WALLACE L. 1963. *Handbook of the Seneca Language*. Albany, N.Y.: The University of the State of New York.

CORNILLAC, GUY 1993. "De la nécessité de concevoir pour l'esquimau un dictionnaire sans mots". *Amerindia* 19/20: 203-209.

FOSTER, MICHAEL, KARIN MICHELSON & HANNI WOODBURY 1989. *Base and Affix Dictionary for Iroquoian Languages*. Unpublished paper.

LOUNSBURY, FLOYD. 1953. *Oneida Verb Morphology*. (Yale University Publications in Anthropology, 48.) New Haven, CT: Yale University Press.

MICHELSON, GUNTER 1973. *A Thousand Words of Mohawk*. Ottawa: National Museum of Man.

MICHELSON, KARIN 1990. "The Oneida Lexicon". In: Costa, David J. (ed.), *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistic Society, February 16-19, 1900: Special Session on General Topics in American Indian Linguistics*. Berkeley: Berkeley Linguistics Society, 73-84.

MICHELSON, KARIN & MERCY DOXTATOR (to appear). *Oneida-English; English-Oneida Dictionary*. University of Toronto Press.

RUDES, BLAIR A. 1987. *Tuscarora Roots, Stems, and Particles: Towards a Dictionary of Tuscarora*. (Algonquian and Iroquoian Linguistics Memoir # 3.) Manitoba: Voices of Rupert's Land Fund.

RUDES, BLAIR A. 1999. Tuscarora-English, English-Tuscarora Dictionary. Toronto / Buffalo / London: University of Toronto Press.

SASSE, HANS-JÜRGEN 1993. "Das Nomen - eine universelle Kategorie?" *Sprachtypologie und Universalienforschung* 46(3):187-221.

SASSE, HANS-JÜRGEN & ALFRED KEYE (forthcoming). *Cayuga Verb Morphology*. Brantford, Ont.: Woodland Cultural Centre.

Ket as a polysynthetic language, with special reference to complex verbs[1]

WERNER DROSSARD

1. Yeniseyan languages

Ket is the only surviving member of six Yeniseyan languages. According to WERNER (1994: 1-4, 1995: 1-6, 1996: 1-5 and 1997: 1-4) all of these languages were still alive during the 18th century, as is attested by the field notes of four German researchers (D.E. MESSERSCHMIDT 1720-1727, J.E. FISCHER 1768, P.S. PALLAS 1776 and J.G. GEORGI 1776-1780). Three subgroups can be reconstructed: Arin and Pumpocol, Assan and Kott, Ket and Jugian. Roughly, Ket is spoken by approximately 800 persons in an area around the Yenisei river between its eastern tributaries, Lower Tunguska and Stony Tunguska.

2. A short history of Ket research

WERNER (1997: 4-10) presents us with the most detailed survey on Ket studies from the 18th century to the present. The next in line, after the four German researches mentioned above, is the famous M.A. CASTRÉN who compiled his field notes in his *Versuch einer jenissei-ostjakischen und kottischen Sprachlehre* (1858). DONNER published the summary of his journeys between 1911 and 1916 in 1916. In 1934 KARGER wrote a grammatical sketch of some 15 pages for the series *Jazyki i pis'mennost severa* which is still of some relevance (KARGER 1934b). Referring back to CASTRÉN's and DONNER's studies BOUDA composed several papers (1957, 1968, 1971) that will be discussed later.
The sixties were marked by three prominent scholars, A.P DUL'ZON, his pupil H. WERNER and E.A. KREJNOVIČ. As we shall see later the publications of these three authors - supplemented by BOUDA's work - will be the framework for our argumentation.

---

1 I want to thank Nicholas Evans (Melbourne), Stefan Georg (Bonn, NOW-project: A descriptive grammar of Ket) and Hans-Jürgen Sasse (Cologne) for useful comments on earlier versions of this paper.

There are a lot of noteworthy studies that were written between the seventies and the nineties that can be headed under phonology, morphology, syntax, and word-formation. It turns out that problems of tonology have not received the attention they deserve (cf. WERNER:1996). Some papers of ŠABAEV (1984), VALL-KANAKIN (1990) and BELIMOV (1991) concerning coding strategies and problems of syntactic typology are discussed by WERNER (1997).

3. Problems of presentation and analysis

3.1 Different authors - different writing systems

Anyone working on an 'exotic' language that has been documented over the course of three or four centuries has to face the discrepancies between the writing systems chosen by different authors for recording the language in question.

This implies that each of the researchers mentioned in 2. chose his own way of writing, at least up to 1934. As WERNER (1997: 25) points out, KARGER was the first at this time to compose an alphabet which was based on the middle Ket dialect (KARGER 1934a). Since most Ket speak the southern dialect now, WERNER introduced a new alphabet in 1989, which uses Cyrillic letters and some diacritics.

Nevertheless, the problems we have to face remain:

a. Some of the sources we quote (DUL'ZON, KREJNOVIČ) do not use KARGER's alphabet, but their own system instead
b. BOUDA and WERNER (in their German publications) use different Latinized versions

The compromise we have made here is to try to adjust the Cyrillic systems used (despite their discrepancies) to the Latinized version used by WERNER, and also to adjust BOUDA's writing conventions to WERNER's.

3.2 Different authors - different analyses

It is a trivial fact, too, that different authors have different ways of analysing the facts of a language. In the course of the following paragraphs the reader will notice that the authors chosen as sources for our discussion often differ greatly in their analyses. The main points of disagreement are found in morphological analysis. As we will show later (emphasizing this phenomenon where pertinent) some authors tend to interpret units of three or four phonemes as one morpheme with a single meaning, whereas others segment the item into

two or three meaning components, with the interpretations of possible smaller parts also differing widely.

3.3. One author - different ways of presentation and analysis

Some irritations that result from the circumstances described above are reinforced by the fact that one and the same author will change the way of presenting or interpreting the facts of the language described. Without wanting to review WERNER's books, one of many cases of 'repairs' has been choosen to illustrate the point, namely that the presentation of personal affixes shows some discrepancies.

In his 1994 presentation (p.99-100) WERNER gives a summary of the so-called B- and D-affixes (cf. § 5.1) These surveys contain the labels for person-marking and series without any additional labels for secondary functions. WERNER (1997: 150-151), on the other hand, includes some extra information on the secondary function of subject- and object-affixes, i.e. some information on version (which is discussed in § 5.1.1). As can be seen from this paragraph we have provisionally adopted the latter variant.

WERNER himself gives the following reason for altering his presentation. In his 1994 presentation he based his analysis on KREJNOVIČ: the *a*-affix of the 3sg is taken as an affix of intransitivity, but subsequent research suggested that the *a* in question is a marker of subject version.

It is absolutely legitimate for researchers to change their minds. Due to the extremely complicated morphology of Ket, we have to give credit to everyone involved in describing this language, but it has to be stressed that the vast number of discrepancies (which we cannot list in toto) has led to some uncertainty on the part of this author. In addition, the reader has to bear in mind that we were dependent on compiling the facts collected by different authors (with controversial opinions) without having the opportunity to do field research of our own. One of the most striking facts with regards to the subparagraphs to follow is that WERNER and KREJNOVIČ did not care about the role and function of what BOUDA calls preverbial elements. It is entirely due to DUL'ZON that the importance of these morphemic units has been shown, in particular in regard to polysynthesis.

4. Polysynthesis: characteristics and definitions

4.1 Polysynthesis as defined in BOAS 1911

In an earlier paper (DROSSARD 1996[2]) we have shown that polysynthetic languages can be systematized within a framework between non-sententiality and sententiality, the minimum criterion being that a language has at least one semantically definable category (i.e. adverbial concepts) that occurs exclusively in bound form, primarily within a verbal head. This implies that some Amerindian languages are polysynthetic, though having neither polypersonal verbs nor noun incorporation. The minimum condition for this type conforms to what BOAS said in his introduction to the *Handbook of American Indian languages*

> "a large number of distinct ideas are amalgamated by grammatical processes to form a single word, without any distinction between the formal elements in the sentence and the contents of the sentence".

Within this framework some Athapaskan languages, including Haida and Tlingit, can be regarded as polysynthetic, although

> "(they) do not readily incorporate the object, but treat both pronominal subject and pronominal object as independent elements"

This means that the following Haida sentence from SWANTON's text (1911:277ff) can be taken as an example of non-sentential polysynthesis: the third person pronoun is not included in the verbal complex, but there are bound instrumental and directional prefixes and suffixes with lexical content fulfilling the minimum criterion of polysynthesis:

(1) *L!* *Lu* *-is* *-L!xa* *gí'l* *-gAn*
they by canoe -come -motion toward -landward -PRET
'They came shoreward by canoe'

---

2 The main goal of my 1996 paper was to integrate traditional and recent concepts on polysynthesis, as for instance Boas' definitions (quoted in the text), Fortescue's criteria (1994) and Baker's syntactically-based research (1996). Despite some controversies that arise from combining these approaches with each other, I propose that there are three main types of polysynthesis that differ in complexity, but nevertheless can be called polysynthestic in each case. This allows us to also include, in particular, cases like Yimas which is called polysynthetic by FOLEY (1991:277ff.), but would not fit into Baker's pattern due to its lack of productive noun incorporation, despite its polypersonalism and adverbial incorporation.

The next subtype of polysynthesis, then, is given when - in addition to the integration of adverbial concepts - polypersonalism exists. This is the case in Tsimshian as quoted by BOAS himself:

(2) *t -yuk -ligi -lo -d'Ep -daL -Et*
he -begin -somewhere -in -down -put -it
'He began to put it down somewhere inside'

Although some authors (including FORTESCUE 1994:2600) regard noun incorporation as one of the most prominent criteria for polysynthetic languages, we nevertheless want to adhere to our minimum criterion, the integration of adverbial (instrumental, modal, local) concepts into the verbal head. There is a third subtype simultaneously displaying adverbial affixing, polypersonalism **and** noun incorporation as shown in the following example from Wichita (ROOD 1976:65):

(3) *kiya:ki -riwa:c -áras -a -ri -kita -a? -hí:rik -s*
QA:3sg -large -meat -COLL:pl:PAT -PORT -top -come -ITER -IMPF
'He brought the large (quantity of) meat to the top'

Ket belongs to the third subtype. Evidence for this will be given in detail in paragraph 5. To give the reader a first impression we can refer to example (14) displaying polypersonalism, example (8) containing noun incorporation and example (37) showing bound adverbial morphology.

4.2 The structure of the lexicon in polysynthetic languages

Most authors who refer to polysynthetic languages emphasize that polysynthesis (especially the combining of bound morphemes) is concentrated around a verbal nucleus. One of FORTESCUE's main characteristics of polysynthetic morphology is "the integration of locational, instrumental and other adverbial elements into the verbal complex as affixes" (1994: 2601, characteristic e). These so called 'field affixes' modify verbal processes or actions by adding information about their 'periphery': is the action directed upward, downward, into water, upstream, etc, or is it carried out by hand, by foot etc. In surveying Amerindian languages, one will surely come across a lot of structures similar to that given in example (1). Thus, one reason for the existence of 'long words' (to be more precise: 'long verbs', in a narrower sense) is the accumulation and addition of 'adverbial material', as mentioned above. Another reason is that a lot of polysynthetic languages compose their verbal lexemes by putting together meaning units, i.e. by using adverbial material as a means of word formation. As TALMY (1985) has shown in his important paper on lexicalization patterns, some languages lexicalize a bundle of meaning features (manner, figure and the like) in the

shape of one unanalysable, non-transparent unit, whereas other languages display a step-by-step, and morpheme-by-morpheme composition, thereby possibly inflating the size (and length) of the verbal word. Compare for instance:

(4) English: to sew

(5) Eastern Pomo (MITHUN 1996:140)

bi°	*-khí*	*-ch*	*-ki*	(to sew) *-°*
draw by sharp object	-together	-move in consistent direction	-expelling pressure -SEMEL	-STAT

5. Ket as a polysynthetic language

As mentioned above Ket is a polysynthetic language with adverbial affixes, polypersonalism and noun incorporation. Each of these ingredients will be discussed in detail below. It turns out that one additional and crucial feature of polysynthesis in Ket is that verbal forms, in many cases, are compounded by two verbal roots.

Even though such patterns are characterized as **polysynthetic** on the basis of their morphological behavior, we want to show (§ 6) that the standard explanations used for serial verb constructions (such as those found in **isolating** languages) may also be applied to Ket. That is, the systematization that holds for serial verbs (characteristic of South Asian, Papuan and African languages) shall be taken as a basis for the description of Ket 'verb + verb constructions' with the additional effect of shedding more light on the language and making it amenable to more general typological comparisons. But first of all a general overview of the verbal complex must be given.

WERNER (1995: 52-53, 1997: 154-155 and 1998: 71-72) presents us with a template similar to those used in Athabaskan linguistics, exhibiting 14 positions to the left of the verbal root and three to the right. Although, in general practice, morphemes left of the root (i.e. position 0) are indexed by a 'minus', and morphemes to the right by a 'plus', WERNER turns the whole thing around. We shall adopt WERNER's notation, because the bulk of the material to be analysed is found on the left.

14 13 12 11 10 9 8 7 6 5 4 3 2 1 0 -1 -2 -3

In detail, the positions occupied are:

14 = subject pronominal prefix, 13 = root 3 (only coocurrent with position 12, e.g. nouns incorporated by verbs of position 12), 12 = root 2 (for verbs, preverbs, adverbs, nouns), 11 = derivative, 10 = causative, 9 = version, 8 = subject or object pronominal, 7 = determinative,

6 = permansive (aspectual), 5 = tense, 4 = subject or object pronominal, 3 = aspect, 2= subject or object pronominal incl. version, 1 = imperative, 0 = root 1, -1 = derivative, -2 = plural marker, -3 = subordinating suffix.

We also adhere to WERNER's approach in not taking epenthetic vowels (EV) or epenthetic consonants (EC) to be positional classes of their own.
Using this template as a basis we can show that Ket, according to the typology given in § 4, has:

a. polypersonalism
b. verbal affixes for adverbial meanings (esp. in position no. 12)
c. noun incorporation (the incorporated noun is also placed in pos. 12 (R 2)

First of all, we discuss six basic constructional patterns listed by WERNER (1998: 69-70), in order to show how Ket works. These include simple patterns with one root and, in extreme cases, three root morphemes and one determinative (the status of the latter is discussed in 5.4):

pattern 1: 0 (one root morpheme only)

(6)

14	4	0
di	*-p*	*-tet*
1sg:S	-3sg:O	-hit

'I hit it'

pattern 2: 7 0 (one root and one determinative)

(7)

14	7	5	4	EV	3	0
da	*-d*	*-o*	*-v*	*-i*	*-l'*	*-de*
3sg:f:S-DET	-PST	-3sg:n:O	-EV	-PERF	-load	

'she loaded it onto it'

pattern 3: 12 0 (two root morphemes)

(8)

14	12	5	EV	0
da	*-nan'*	*-s'*	*-i*	*-vet*
3sg:f:S-bread		-PRS	-EV	-make

'she is baking bread'

pattern 4: 13 12 0 (three root morphemes)

(9)

14	13	12	7	EV	0
da	*-nan'*	*-bet*	*-q*	*-i*	*-vet*
3sg:f:S-bread		-make	-DET	-EV	-do

'she will bake bread'

pattern 5: 12 7 0 (two root morphemes and one determinative)

(10)

14	12	7	5	2	0
d	*-al'*	*-d*	*-a*	*-d*	*-do*
3sg:f:S-beat		-DET	-PRS	-1sg:O-	chop

'she beats me up'

pattern 6: 13 12 7 0 (three root morphemes and one determinative)

(11)

14	13	12	7	0
da	*-haŋ*	*-s'el'*	*-G*	*-ej*
3sg:f:S-female		-reindeer	-DET	-kill

'she killed a female reindeer'

As we shall see later, our main emphasis in § 6 will be on pattern 3. The above example includes a noun in R2 position, but it is also possible, as indicated above, for an adverbial element to be in this slot (as a kind of preverb or prefix) or for verbal elements (which will be discussed below at length). This fact is well captured in DUL'ZON's grammar from a more large-scale perspective:

> "(within the more complex verbal forms) the modifying root can be a noun, adjective, numeral, verb, pronoun or adverb..." (DUL'ZON 1968:486)

With regards to the structure of our approach, this implies that, after having discussed polypersonalism, we first have to look at incorporation phenomena (which means that there are nouns in R2-position), then deal with adverbial elements in this position and finally discuss KREJNOVIČ's presentation (1968a), which mainly concentrates on interpreting the components of complex R2-R1- configurations as verbal. But as we shall see later it will be possible to go beyond that scope, since, founded on DUL'ZON's quote above, it is plausible to take into consideration adverbial or even nominal sources (as R2s), the possibility of which eludes KREJNOVIČ.

5.1 Pronominal affixation

5.1.1 Polypersonalism

The most relevant positions in the above patterns (1-6) are numbers 14, 8, 4 and 2. In order to determine what is what, we can again refer to two of WERNER's synopses. In his books (1994: 99) and (1997: 150) he gives a summary including the basic distinction between B- and D-affixes, B representing 1sg in one group of verbs, D representing 1sg in another group of verbs, according to a proposal by BOUDA (1957).
The first and most basic structures featuring D- versus B-affixes (here *di* and *ba* respectively) are the following two examples for intransitive verbs:

(12) featuring subject of series D

14	5	EC	3	0
di	*-Ø*	*ɣ*	*-Ø*	*-aq*
1sg:S	-(T)	EC	-(ASP)	-leave

'I leave'

(13) featuring subject affix of series B

8	7	EV	5	3	0
ba	*-ɣ*	*-i*	*-s*	*-Ø*	*-sal*
1sg:S	-DET	-EV	-PRS	-(ASP)	-spend night

'I spent the night'

The transitive patterns given by WERNER include a pattern with identical D-affixes in subject and object slots as in:

(14) featuring subject and object affixes of series D

14	5	3	2	0	-2
du	*-Ø*	*-Ø*	*-t*[3]	*-tiɣ*	*-in*
3:S	-(T)	-(Asp)	-1sg:O	-beat	-pl

'they are beating me up'

The second kind is represented by example (15) where the subject affix belongs in the D-series and the object affix in the B-series.

3 *-t* in position 2 is an allomorph of *-d*; compare also example (25)

(15) featuring subject affix of series D, object affix of series B:

14	8	6	5	3	0
da	*-ba*	*-t*	*-o*	*-l'*	*-oŋ*
3sg:f:S-1s:O	-PERM		-PST	-PERF	-look

'she has looked at me'

More complex and to some degree more complicated structures, the nature of which cannot be described here conclusively, come about by adding D- or B-affixes in various secondary functions that WERNER (1995: 110) calls 'versions'. The term 'version' implies that

a. Affixes of the D-series, when repeated, take over reflexive functions in the broader sense including "autobenefactive" and comitative readings. This version type is called 'subjective version' glossed as SV in the following.

b. Affixes of the B series take over orientational features connected with verbs of motion. One set of the B series, the one with *ba* for 1sg is called introverted, i.e. movement is for the most part characterized by "back-and-forth" manner, whereas the B series variant with *bu* for the 1sg is called extroverted, i.e. movement is directed away from a deictic centre. The *ba* variant will be glossed as IV, the *bu* variant as ExV in the following.

Example (16) contains a subject suffix of series D in slot 14, as well as a further pronominal element of the D-series in slot 2. The German gloss 'Sie bewegt sich.' is added in order to show the reflexive character of the construction

(16) featuring subject- and version affixes of series D:

14	5	3	EC	2	0
de	*-Ø*	*-Ø*	*-j*	*-a*	*-r'uk*
3sg:f:S-(T)		-(ASP)	-EC	-SV	-move

'she moves' cf. German 'Sie bewegt sich.'

Example (17) introduces a B-series affix of the extroverted type:

(17) featuring subject of series D and version of series B

14	9	6	5	3	0	-2
d	*-bu*	*-t*	*-s*	*-Ø*	*-aR*	*-an*
3:S	-ExV	-PERM	-PRES	-(Asp)	-run	-pl

'they are just running there'

Transitive structures are also expanded by version affixes. Example (18) shows consistent D-series marking for subject and object relations, but adds an affix in slot 9 for the extroverted orientation:

(18) featuring subject and object prefixes of series D and version affix of series B:

14	9	5	3	2	0
da	*-bu*	*-Ø*	*-n*	*-di*	*-Rus*
3sg:f:S	-ExV	-(T)	-PERF	-1sg:O	-take away

'she has led me away'

Example (19) contains a D-series affix for the subject relation, a B-series affix for the object relation and, in addition, a D-series affix as a marker of subject version, i.e. in an autobenefactive reading (... for my benefit ...).

(19) featuring subject series D, object series B and version affix of series D

14	8	5	3	2	0
d	*-o*	*-Ø*	*-l'*	*-di*	*-Ra*
1sg:S	-3sg:m:O	-(T)	-PERF	-SV	-sell

'I sold him'

The last example to be discussed here is given in (20). The transitive core of the construction, i.e. a D-affix for the subject relation and a B-affix for the object relation, is expanded by another B-affix with extroverted reading.

(20) featuring subject affix of series D, version and object affixes of series B

14	9	EV	8	5	3	0	-2
d	*-bu*	*-y*	*-aŋ*	*-Ø*	*-Ø*	*-it*	*-n*
3:S	-ExV	-EV	-3pl:O	-(T)	-(ASP)	-carry	-pl

'they are carrying them away'

5.1.2 Some typological considerations

There has been a great deal of discussion regarding the coding strategies of the pronominal affixes involved. Do Ket pronominal forms constitute active, nominative, or ergative patterns? COMRIE (1981: 264) tries to sum up matters by saying:

> "For transitive subjects (ergative relation), only the D series is used, and prefixes encoding transitive subjects always occur initially, preceding both parts of a discontinuous root. For intransitive subjects and for direct objects (absolutive relation), either the B or the D series is used, the choice being determined lexically (apparently arbitrarily); in this function, the affixes occur before a nondiscontinuous root, but between the two parts of a discontinuous root."

It should be pointed out firstly, that the indiscriminate use of the B and D series in intransitive contexts seems to be idiosyncratic and synchronically rules out any evidence for an active-inactive distinction, compare for instance WERNER's examples (1994: 107)

(21) *di -n'etet* 'I am going away'

featuring a D-series -prefix, versus

(22) *bo -γevitn* 'I am running out'

featuring a B-series prefix. The lexemes involved are semantically 'active verbs'.

What can be deduced immediately from these examples, secondly, is that ergativity (as inferred by COMRIE) would be consistent if the B-series was exclusively reserved for the subject of all intransitive verbs. On the other hand, some authors, WERNER (1994) included, use the presence of the D-series in intransitive verbs to bring together transitive and intransitive D-patterns to show that the language has nominative traits. On balance, it seems that Ket displays a mixed structure, but, as shown by WERNER (1994: 102), ŠABAEV (1986) takes the radical (and hardly understandable) view that B-affixes are exclusively reserved for the coding of object relations, whereas D-affixes are responsible for subject- and object coding. The first part of his claim can be refuted by referring to (22), again, which includes a B-series prefix with an active intransitive subject. The second part, however, can be verified by many of WERNER's examples for instance by

(23) *du -yan -s'*
3sg:S(D) -2pl:O(D) -dress
'He is dressing you'

Although we can give only a very brief summary here of the discussion on how to assign Ket to one of the three syntactic types, the whole story can be summed up in what COMRIE (1981: 264) states in his preliminary remarks on the Ket verbal system:

> "The difference between these two sets does not, however, correlate at all directly with that between subject and direct object, and in many cases it seems quite idiosyncratic which set a particular verb takes."

5.2 Incorporation

Incorporated nouns are normally placed in position 12. When there is a verbal root in position 12, nouns are placed in position 13 (cf. pattern 4). Thus, we can distinguish a minor pattern (R2 - R1) from a major pattern (R3 -R2- R1).

5.2.1 The minor pattern

The most important fact about the 'minor' pattern is that position 12, that is root 2 (R2), can also be occupied by verbal roots. Compare for instance:

(24)

	14	12	6	EV	0
at	*d*	*-don'*	*-s'*	*-i*	*-vet*
1sg	1sg:S	-knife	PRES	EV	-make

'I'm making a knife'

(25)

	14	12	EV	2	0
at	*d*	*-us'n*	*-i*	*-t*	*-tiŋ*
1sg	1sg:S	?	-EV	-SV	-turn

'I'm turning around'

(for a more detailed analysis of this form compare (64a))

The parallel behavior of verb roots and nouns in position 12 has prompted KREJNOVIČ and WERNER to assume that verb + verb constructions (as in 25) were formed after and analogous to the incorporative pattern (for a more detailed discussion, see § 6). Within this pattern we again have to make a distinction, namely between incorporation as a syntactic device and incorporation as a word-formation process.

5.2.1.1 Incorporation as a syntactic device

Unlike most languages with incorporation Ket allows incorporated nouns to be suffixed for plurality. Compare:

(26a)

d	*-don*	*-s'*	*-i*	*-vet*
1sg:S	-knife	-PRES	-EV	-make

'I am making a knife'

vs.

(26b)

d	*-don'*	*-aŋ*	*-s'*	*-i*	*-vet*
1sg:S	-knife	-pl	-PRES	-EV	-make

'I'm making knives'

WERNER (1995: 75-76) emphazises that the incorporated nouns in both cases replace pronominal object markers, to such an extent that (26a) has the analytic counterpart

(26a') *(at)* *don'* *di* *-b* *-bet*
1sg knife 1sg:S -3sg:O-make

and (26b) has the (26b') - variant

(26b') *(at)* *don'* *-aŋ* *di* *-b* *-bet*
1sg knife -pl 1sg:S -3pl:O -make

WERNER (1995: 70 and 1997: 59 ff) presents a list of verbal forms that are capable of incorporating nouns.

The most productive root-like element (as shown above) is *s'ivet* (which includes the root for 'make', and can be analysed as: *s'* (PRES) + EV + *vet* 'do'= progressive for 'be doing'):

(27) *t* *-nan'* *-s'* *-i* *-vet*
1sg:S -bread -PRES-EV -make
'I'm baking bread'

The combination of the auxiliary *vet/bet* with elements of the D- and B-series (in bold face) creates the expression for 'having'. (KREJNOVIČ 1968a: 139-142 and WERNER 1997: 217):

(28) *don'-**di**-vet* 'I have a knife'
*don'-**a**-vet* 'he has a knife'
*don'-**i(j)**-bet* 'she has a knife'

A second way to express possession is to use the first series of class B (in bold face again) in combination with *vet*:

(29) *bogdom-**ba**-ja-vet* 'I have a gun'
*bogdom-**a**-ja-vet* 'he has a gun'
*bogdom-**i**-ja-vet* 'she has a gun'

These patterns are very productive.

In contrast to the process just shown there are also cases where nouns are incorporated but the pronoun slot is also filled as in (WERNER 1995: 73):

(30) *d* *-don'* *-u* *-k* *-s'* *-i* *-vet*
1sg:S -knife -3sg:O(inan) -DET -PRES-EV -make

WERNER characterizes the pronominal element here as an indicator of an additional material relation and translates the sentence as: 'I make a knife **out of it**'.

The second circumstance in which both incorporated root and object prefix are found is where the incorporated noun is an instrument:

(31) *da -don' -ba -tet*
3sg:f:S-knife -1sg:O-hit
'she stabbed me with a knife'

5.2.1.2 Word formation by noun incorporation

Since Ket only has 50 or 60 simple verbs at its disposal (cf. WERNER 1995: 163, GAJER 1983: 62) a large number of lexical meanings have to be composed from different components. One way of combining distinct elements within an R2-R1 pattern is to join a nominal in R2 and a verb in R1 position to yield a coherent meaning. One prominent example (that is given repeatedly by WERNER) is:

(32) 14 12 9 6 5 3 0
da -tuyun' -ba -t -o -l' -ɣit
3sg:f:S-comb -1sg:O -PERM -PST -PERF-smear
lit: 'She smeared me with a comb' = 'she combed me'

This appears to be an example of instrument incorporation that has been lexicalized with the meaning 'to comb'. In an analogous manner a conventionalized noun-verb-combination is used to express English 'to forget'. DUL'ZON (1968: 203) has a detailed remark:

> "The verb expressing 'to forget' has a complex pattern that is formed by a predication in which the initial part is *en* ('memory') and the final one *ok* (*uk*) is the same as that included in *dokŋ* 'to fly'. The Ket form 'I forget' corresponds in structure to Russian 's.th. flies away from my memory' (u menja vyletaet iz pamjati)."

The relevant example is:

(33) *en -ba -s' -uk*
memory -1sg -PRES-fly
'I forget'

5.2.2 The major pattern

Most examples given for R3-R2-R1 structures include a rather fixed noun-verb-component (R3-R2) on the one hand and an auxiliarized verbal in R1 position, compare pattern 4 above.

Another case that corresponds to this kind of formation is (34), given by WERNER (1995: 156):

(34)	14	13	12	7	EV	3	2	0
	da	*-nan'*	*-bet*	*-q*	*-i*	*-n*	*-daŋ*	*-et*
	3sg:f:S	-bread	-make	-DET	-EV	-IMPF	-1pl:O	-make
	'She is making bread'							

WERNER (ibd.) takes the *et* as R1 without further commenting on it. Following DUL'ZON one could interpret this form as a variant of *bet*. Thus, we can gloss the R1 position as 'do' or 'make', parallel to position 12. As we can see below, position 12 can also be occupied by preverbial elements yielding a rather productive pattern of word formation.

In some cases a comparison between Ket complex verbs and their IE counterparts shows that the former language explicitly shows the meaning components that are 'covert' in English or German. To put it the other way round: every native speaker of English or German knows that 'to lick' / 'lecken' is an activity carried out by the tongue. It is not surprising that Ket uses a noun - verb - combination to express the meaning in question. (35) employs a noun in R3-position, a preverb in R2-position (functioning as verb, indeed) and one of the auxiliarized verbs in R1 position (DUL'ZON 1968: 371):

(35)	14	13	12	5	0
	d	*-ej*	*-t*	*-a*	*-ɣot*
	1sg:S	-tongue	-up	-PRES	-move
	lit: 'I move my tongue up, I lift the tongue' = 'I lick'				

Likewise, everybody knows that spitting means 'to emit a liquid'. Thus, Ket joins the noun for 'liquid', 'water' with a preverb and the most frequent auxiliary *bet* 'make, do' (DUL'ZON 1968: 322):

(36)	*d*	*-ul'*	*-t*	*-a*	*-i*	*-bet*
	1sg	-water	-out	-PRES	-3sg:O	-make
	lit: 'I put water out (of me)' = 'I spit'					

5.3 Local modification within the verbal complex

5.3.1 Preverbs

In examining local modification within the Ket verbal complex we encounter considerable complexity. Concentrating upon only four sources, one nevertheless is faced with several discrepancies regarding the number and status of bound local elements.

According to BOUDA (1957: 112-119) there are two groups of adverbial modifiers that both occur in position 12 within the Ket verbal template: local preverbs (that mostly derive from body part nouns) and directional preverbs. The justification for postulating two preverbial groups is, as indicated above, not only semantically motivated, but mainly based on etymological grounds as far as these can be provided.

In the first group there are 13 denominal prefixes:

I	*fa*	from a noun denoting 'breast', grammaticized to a preverb meaning 'up', (German: 'auf')
II	*ta*	from a noun denoting 'upper surface', grammaticized to a preverb meaning 'upwards'
III	*aŋte*	from a noun denoting 'back', grammaticized to a preverb meaning 'behind' and 'down'
IV	*ēs*	(no etymology), also meaning 'upwards'
V	*eta*	(no etymology), meaning 'out', 'outside'
VI	*ala*	(no etymology), meaning 'out' and 'out of'
VII	*uote*	(no etymology), meaning 'upwards'
VIII	*igda*	from a noun denoting 'backside', meaning 'behind s.th.' or 'backward'
IX	*uoske*	(no etymology), also meaning 'back'
X	*qa*	(no etymology), meaning 'into'
XI	*su*	being a root for 'house', 'home', meaning 'homeward'
XII	*ide*	cognate to *hite* 'bottom', 'underside', meaning 'to the riverbank'
XIII	*ta*	(a homonym to II), cognate to a noun *ty* 'abdomen', meaning 'down'

It is obvious that there is uncertainty in some cases with regards to etymological reconstruction.

The second group seems to contain genuine deictic elements. These so-called 'Richtungspräverbia' (directional preverbs) include

XIV	*aŋa*	'here', 'over here' (towards speaker)
XV	*ka*	(same meaning)

There is also further a miscellaneous group which features three elements with unclear or speculative etymologies:

XVI	*na*	appears in preverbal position, but has unclear semantics
XVII	*taqa*	also rather unclear, something like 'around'
XVIII	*ug*	meaning something like 'away', no etymology

DUL'ZON (1968: 496) refers back to BOUDA (1957), but his list of preverbs includes, as far as we can see, only 11 (instead of 18) forms. Identical to BOUDA's preverbs are the following:

fa (Bouda: I), *ta* (Bouda II), *e:s* (Bouda IV), *eta* (Bouda V), *igda* (Bouda VIII), *ka* (Bouda XV).

In addition there are five 'new' entries into the preverbial list for which we continue the numbering from above (without any attempt to etymologize):

XIX	*ki*	'in front of'
XX	*ni*	'down'
XXI	*e/ej*	'upwards'
XXII	*u/us'*	'downwards'
XXIII	*ik*	'towards speaker'
XXVI	*ok*	'away from speaker'

WERNER (1997: 61) presents us with nine affixes. The most prominent characteristic of his glosses is that these elements, from TALMY's point of view, do not just express only a 'path' component, but also include a 'ground' sememe. Within BOUDA's list this latter feature only applies to item no. XII, whereas DUL'ZON's description does not include any reference to 'ground'-interpretations. The most prominent point now is that the BOUDA/DUL'ZON preverbs (excluding for the moment no.XII) are only components for verbal derivation or word formation, whereas WERNER's affixes seem to modify an action or process (including the participants involved). There are, nevertheless, identical or semi-identical forms to the 18 preverbs mentioned above:

a.	*aɣa*	'from riverbank towards the wood'
b.	*ʌta*	'from the river toward the bank and wood'
c.	*igda*	'out of the wood towards the riverbank'
d.	*ɛt*	'upstream (on frozen river)'
e.	*ɛs'ka*	'upstream (by boat)'
f.	*ata*	'downstream (on frozen river)'
g.	*tiɣa*	'downstream (by boat)'
h.	*ʌl'*	'to the outside and back'
i.	*s'uya*	'back home'

It is obvious that a) is formally identical to BOUDA's no. XIV. But instead of glossing this element simply as a verbal 'satellite'(to use TALMY's term), WERNER understands it as a

bound directional modifier without attempting to give a more detailed analysis (which is given, however, in KREJNOVIČ (1968a), see below). A *-ta* component (BOUDA's no. II) seems to be included in b) and f), and c) *igda*, glossed by BOUDA (VIII) as 'back', receives a rather complicated orientational explanation.

At the bottom of it all is what KREJNOVIČ (1968a) describes for Ket in a rather similar way to the status of spatial orientation in Gilyak (KREJNOVIČ 1960). The relevant facts are also known from some Amerindian languages of the Northwestern coast region: environmental and geographical coordinates are reflected in language, whether in shape of affixes or as special verbal forms (see below). Thus, we have to take as a starting point KREJNOVIČ's basic assumptions (1968a: 171):

> "Two directions of movement are most important for the life the Kets live during winter (that is from late autumn to the beginning of spring): from the river to the depth of Taiga and [later in spring and summer, W.D.] from Taiga to the river".

Thus, KREJNOVIČ splits up *aɣa* or *aɣoe* into the noun root *a* ('woods') and a locative suffix *-ka* meaning 'towards'. So much for the linguistic reconstruction.

During winter the Kets set off for the Taiga in order to hunt. In spring they are orientated *igda*, from the Taiga back to the riverbank (for fishing). It should be stressed here that, according to KREJNOVIČ, Ket also provides two verbal roots for both movements:

ʌta	'to go from riverbank to the wood (Taiga)' and also meaning 'to go from the river to its bank'
aRat	'to go from the woods to the river'

Comparing KREJNOVIČ's and WERNER's analyses it is striking that WERNER does not take *ʌta* as a verb, but as an adverbial suffix (b) in his list, and that he adds no form similar to *aRat*.

KREJNOVIČ, on the other hand, presents examples containing *aRat* in complex structures including *-vet* (see below, §6):

(37)	*d*	*-aRar*	*-o*	*-l*	*-vet*
	1sg:S	-go from wood to river	-PST	-PERF	-do

Despite some discrepancies between WERNER and KREJNOVIČ, both authors discuss the affixes for the movement 'upstream' (WERNER's item (d) and KREJNOVIČ 1968: 175). The latter presents us again with a verbal counterpart *es'ka*, which is WERNER's item (f).

The movement 'downstream on frozen river' is given by both authors as *ata*, as a verbal affix. KREJNOVIČ again adds a verbal form *ʌl'goej*. The (g)-variant of WERNER's list, *tiɣa*, has an identical counterpart in KREJNOVIČ's text. KREJNOVIČ also adds a verbal version *tiɣat*, which seems to be a verbalized form of *tiɣa*.

The clearest and simplest cases in which these spatial affixes are used are given in (38) and (39):

(38) *d* *-at* *-l* *-aq*
1sg:S -downriver -PERF-go
'I went downriver'

(39) *d* *-ʌla* *-l* *-aq*
1sg:S -outside -PERF-go
'I went outside'

As discussed above, 'field affixes' like the ones cited modify only the verbal process or action.

5.3.2 Word formation: preverbs as parts of lexicalization

The discussion of the verbal complex in Ket would be incomplete if we were only to take into account noun incorporation and verb + verb constructions. Disregarding for the moment adjectival and numeral 'incorporation' we should emphasize that many complex verbs are built on a further pattern: 'adverbial' in position 12 (= preverb) + full verb in position ZERO.

Some clear cases of derivation are given in BOUDA's discussion of denominal preverbs, as for instance:

(40) *d* *-e:s* *-a* *-gup*
1sg:S -up -PRES-jump
'I am jumping **up**'

(41) *de* *-eta* *-l* *-aq*
3sg:f:S -out -PERF-go
'She went **out**'

(42) *d* *-ala* *-l* *-aq*
3sg:f:S-out -PERF-go
(same as 41)

As for the composition of new lexical meaning, DUL'ZON (1968: 485-562) lists several configurations that are built on the pattern 'preverb' (= R2) plus R1:

'I dress' is: (43) *ki* *-b* *-el'*
in front -1sg:S -put on

'I shave' is: (44) *d -(t)a -a -b -do*
1sg:S -away -PRES-ExV -chop

The most important point we want to make here is that the 'local-directional' preverbs in contrast to the 'local-ground' variants are constitutive parts of word formation: they are, in TALMY's terms, 'satellites' that contribute to the lexical content of many verbs. Comparing German 'aufheben' with English *'to lift'* shows that German has a satellite, a prefix, whereas its English counterpart inherently contains a 'sememe' for 'upward' without having a concrete morphological expression.

5.4 Determinatives

The elements in position 7, called determinatives, have also caused a lot of discussion. WERNER (1997: 158) assumes that the determinatives

-k- ; -t- ; -q- ; -d-; -n-; -h- and *-s-*

have their origin in demonstrative and locational elements. In addition to that he confirms KREJNOVIČ's (1968a) claim that these elements seem to have lost their meaning completely, and, synchronically, are nothing more than 'concomitant' elements of verbs. Referring back to examples (7) and (10) given above we should emphasize here that the main verbal class containing determinatives is what WERNER calls 'simple verbs', i.e. those with one root morpheme only. Some prominent examples for this class of verbs are:

(45) *bo -ɣ -avitn*
1sg:S -DET -run out
'I am running out'

(46) *t -k -is'tup*
1sg:S -DET -cover
'I am covering (s.th.)'

(47) *d -ba -ɣ -istoq*
3sg:S -1sg:O -DET -attack
'He is attacking me'

It seems that the majority of verbs listed in WERNER's survey contain the determinative *-k* and its allomorphs *-g*, *-ɣ* and *-G*.

6. Analysing verb + verb structures

In addition to the structures discussed above there are further complex structures in which R1 and R2 are **both** 'verbal'. KREJNOVIČ (1968a) dedicated a whole book to this topic. Further, we have to be aware that R1 as well as R2 could also be of 'adverbial' provenance.

6.1 Verb + verb constructions

KREJNOVIČ (1968a: 11ff) proposed the following systematization for these verbal complexes:

1) Verbs with a composed stem (containing two roots)

2) verbs with a root in initial position and a desemanticized second root (which is called a derivational form) at the end

3) verbs with a root in final position and a desemanticized second root in initial position (which, accordingly, is called derivational)

The formulae given by KREJNOVIČ are:

1') R2 + R1 where the latter is in most cases:
aq, bet, git, do, don',
q, n, s'in, tij, tet

2') R2 + grammaticalized R1-derivations as:
dij/rij, a , aq, 'da/r'a, din/dän, i, ij,
k, kŋ, ŋ, o, t , t/a, ta, to ,ti,
tn, uk, Ro, Rus/äm, Rut/ut, Rut/damin

3') grammaticalized R2-derivations as:
u/us', e, ej/ij, uk, el', al' + R1

Since all the elements which appear to be derivational are found in root positions, WERNER rejected KREJNOVIČ's interpretation: instead of calling these desemanticized roots 'derivative' he proposed the term 'semi-affix', which does not make matters much easier. The main point here is that all elements cited have been more or less auxiliarized. Basing our analysis, however, on KREJNOVIČ's exposition, we can get a better grip of the problem, especially when it is seen from a grammaticalization point of view, a term which is also explicitly used by KREJNOVIČ himself. But it is to be stressed that Krejnovič's somewhat one-sided ap-

proach (which concentrates upon verbal sources for all the items given) will be supplemented by alternative types of explanation in 6.3.

Before going into KREJNOVIČ's analyses, we must clarify which functions are served by verb + verb configurations in general, that is, we must list what serial verb constructions accomplish from a comparative point of view. In essence, this means that we will try to fit KREJNOVIČ's systematization into a serial verb construction framework. Findings of the last two decades show that cross-linguistically we can distinguish at least **four** functions for serial verb configurations:

1: Lexical compounding

An example cited often in this context is Yoruba:

(48) *ó* *mú* *ìwé* *wá*
3sg took book came
'He brought the book' (BAMGBOSE 1974: 14)

That is, 'take' + 'come' yields 'bring'.

2: Classifying an action/event/process by a 'generic' verb and naming the action itself by a further, specialized second verb, which precedes it. This type of construction is found in a large number of Papuan languages, for instance Kalam (PAWLEY 1980; FOLEY 1986: 115ff):

(49) *d* *nŋ*
take perceive
'feel'

(50) *pk* *nŋ*
hit perceive
'nudge'

It should be stressed that there is a striking parallel between Kalam and Ket in that the position of the first verb can also be filled by a noun:

(51) *wdn* *nŋ*
eye perceive 'see'

(52) *tmwd* *nŋ*
ear perceive 'hear'

Similar phenomena can be observerd in some non-Pama-Nyungan languages of Australia, as for instance in Jaminjung, cf. SCHULTZE-BERNDT (2000).

3: Grammaticalization of one verb to express some kind of aspect or 'aktionsart'. This is known from south-east Asian languages and some Niger-Congo languages, for instance Yoruba:

(53) (BAMGBOSE 1974)
ó *nsùn* *lo*
3sg sleeping go
'He is falling asleep'

In this case the verb for 'go' is auxiliarized to a marker of inceptivity.

4: 'Coverbal' functions by which semantic relations are introduced or eliminated, i.e. valency-increasing or decreasing functions are brought about by these coverbs, as for instance in Thai (BLAKE 1994: 163):

(54) *Thân* *cà* *bin* *maa* *krungthêep*
he will fly come Bangkok
'He will fly to Bangkok'

In this case the verb for 'come' (as a coverb) is auxiliarized to a preposition-like GOAL marker.

Whereas the Thai example is simply a case of valency-increasing, valency-decreasing by means of verb serialization and related phenomena is not particularly widespread. There are, nevertheless, some cases in which just a lexeme for 'give' which 'normally' introduces recipients can serve to suppress the object (antipassivization), for instance in Chechen (JAKOVLEV 1940: 75):

(55 a,b) transitive: *iecan* 'to buy s.th.'
vs.
intransitivized: *iecadalan* 'to be busy buying'

In the latter example the verb *dalan* 'give' is auxiliarized to a type of aspectual marker that demotes the object and focusses the action in process. Similar cases can be found in German dialects (Ripuarian) where constructions of the type *er gab sich ans Übersetzen* ('he started translating') express inceptive action and also demote the object-relation.[4]

4 It is more frequent, of course, for GIVE to introduce recipients; it is also known to be a source for (permissive) passives, causatives, etc.

6.2 Verb + verb constructions in Ket

Based on our general survey of 'serial verb' constructions we may now try to identify some of the functions listed in Ket.

6.2.1 Lexical compounding

Although the majority of Ket examples suggest that one of the two roots involved in a complex verbal form is desemanticized, we have one clear case, where two verbal forms yield one reading. KREJNOVIČ (1968a: 186) discusses the form

(56) *Ø -ka-s -ti -Rus*
3sg:S-take-PRES -1sg:O-take away
'3sg leads me away'

R1 (*Rus*) has the primary meaning 'to take away', R2 has an infinitive *kases*, meaning 'to take, to buy'. KREJNOVIČ combines the two meanings and (after translating them by Russian 'vzjat'' and 'unesti') gives the final meaning as 'uvesti' ('to lead away').

A kind of semantic bleaching of *-Rus* (R1) occurs in an example that is given by KREJNOVIČ immediately after the first one above:

(57) *d -usk -a -v -us*
1sg:S -open -PRES -3sg:n:O-take away
'I am opening it'

It may be that the action of opening s.th. includes a movement away from the actor, e.g. when opening a tin the protagonist will turn the object away from himself when performing the action to avoid hurting himself with the instrument used. Nevertheless, from KREJNOVIČ's point of view we should know now that the verbal form in question contains a derivation of *Rus* in R1 position.

6.2.2 Generic roots

By 'generic' we here refer to verbs with a relatively neutral nonspecialized meaning which take over the function of classifying an action or event, for instance a root 'do' can be understood as a general concept for a great number of actions.

Whereas the marginal examples in 6.2.1 seem to belong in KREJNOVIČ's group 2', the examples in question now belong to group 1'. In some cases it is possible to give full meanings to the R1, especially for:

bet	'do, make'	*aq*	'go' and 'give'
git	'wipe, smear'	*n*	'become'
don'	'tear'	*s'in*	'be'
tij	'fall'	*do*	'chop, hew'
tet	'hit'	*kot*	'lift, move'

The most clear-cut case is *tet* ('hit') which seems to be a 'classifier' for actions carried out by the hands. Thus,

(58)	*to -tet*	is literally:	'seize/grasp' + 'hit'
		meaning:	'to seize/to grasp'
(59)	*t'el -tet*	is literally:	'push' + 'hit'
		meaning:	'to push'

In addition to KREJNOVIČ's and WERNER's list, DUL'ZON (1968: 367ff) presents another marginal form, *-kot*, which literally means 'to lift, to move' (in intransitive and transitive senses). Among the few examples that can be found are:

(60) *d -ugd -a -ɣot*
1sg:S -dig -PRES -move
'I am digging'

and

(61) *d -ul' -d -a -ɣot*
1sg:S -go -ext -PRES -move
'I am going away'

Another obvious case is *kit* which literally means 'to smear', but generally takes on the meaning of anything that has to do with a 'wiping or rubbing movement':

(62) *si -t -e -git*
sweep -1sg:S -EV -smear
'I am sweeping'

This verb, of course, also can incorporate nouns, as in:

(63) *d -ul -t -Ø -a -kit*
1sg:S -water -DET -3sg:O -PRES -smear
'I moisten it'

Whereas in the aforementioned cases the 'generic' verb is placed at the end of the compound, there are some cases, corresponding to KREJNOVIČ's formula (3') (cf. § 6.1), in which it is R2 that is desemanticized and shows generic traits.

The first of the six cases quoted by KREJNOVIČ, *-u-*, is glossed as 'to split s.th.', but - due to the intense polysemy of Ket monophonemic roots - an alternative is also possible: *-u-* can also can mean: 'to put' or 'to hold'. Now, if we assume the latter variant we can interpret *u* = 'hold' as generic, by analogy with several Papuan languages which possess 'hold' as a generic verb. Thus, we can put forward our own analysis as in

(64a)	*d*	*-u*	*-s'*	*-n*	*-a*	*-don'*
	1sg:O	-hold	-PRES	-DET	-3sg:S	-drag
	'it (the bear) is dragging me'					

A third possibility of analysis emerges by referring back to a proposal by DUL'ZON (1968: 303-304). *u/us'* can also be interpreted as a preverb meaning 'downwards', 'down to an end'. Thus, we could alternatively analyze:

(64b)	*d*	*-u*	*-s'*	*-n*	*-a*	*-don'*
	1sg:O	-down	-PRES	-DET	-3sg:m:S	-drag

In other examples the verbal root in position 2 seems to resist reconstruction. Thus, KREJNOVIČ (1968a: 242-3) has to admit that in (65) R2 (*el'*) is unclear.

(65)	*d*	*- el'*	*-t*	*-a*	*-j*	*-tet*
	3sg:m:S	-???	-PERM	-3sg:m:O	-DET	-hit
	'he is hitting him'					

Referring, however, to DUL'ZON (1968: 334) we learn that *-el'-* can be glossed as 'hand'. It it obvious that inserting 'hand' in R2 position is the most plausible analysis. Finally, then, (65) should be regarded as an incorporative pattern in the narrowest sense, because of the nominal (and no longer verbal) interpretation of R2.

6.2.3 Aspect and 'aktionsart'

The most basic opposition within the domain of 'aktionsart' is 'momentaneous' vs. 'iterative'. One way to mark this opposition is to move R1 from zero-position to position 12 and to insert *bet/vet* in zero position:

(66a)	*d*	*-aŋ*	*-s'*	*-ej*	(momentaneous)
	1sg:S	-3pl:O	-PRES	-kill	'I am killing them now'

vs.

(66b) *d* *-ej* *-aŋ* *-a* *-vet* (iterative)
1sg:S -kill -3pl:O -PRES -do 'I kill them' (again and again)

We can speculate about another so-called 'derivative' morpheme in KREJNOVIČ's list: *-aq*. In his work on the Ket verb USPENSKIJ (1968: 220) assumes that *-aq*, in its primary function as a lexeme denoting 'go', was grammaticalized as a marker for the momentaneous 'aktionsart'. In this case the momentaneous *-aq* version is opposed to *-ij* for iterativity. It should be added here that a reconstruction for *-ij* cannot be given for the moment.

(67a) *d* *-ul'tav* *-d* *-Ø* *-aq*
1sg:S -release -1sg:SV -3sg:n:O -go
'I am letting it go' (momentaneous)

(67b) *d* *-ul'tav* *-Ø* *-ij*
1sg:S -release -3pl:O -ITER
'I am letting them go'

An alternative analysis (deviating from his 'verb-centered' interpretations) has been put forward by KREJNOVIČ. In his paper (1968b: 109) he traces *-aq* back to the numeral *-q* 'one'. This seems to be a plausible way of explaining the semelfactive meaning in question, and the whole thing could be arranged in conformity with the general pattern of Ket compound verbs. Thus, referring back to DUL'ZON's general statement cited in § 5 we could reduce KREJNOVIČ's interpretation to the following pattern

12	0
R2	R1
full verb	numeral

Applying this pattern to a structure as in (68) would also fit in well with KREJNOVIČ's analysis: in his paper on 'aktionsarten' (1968b) he presents another numeral-based interpretation by connecting *da/r'a* with a collective marker of the same form.

(68) *tar'ɛlKaŋ* *d* *-us'n* *-av* *-r'a*
plate 1sg:S -put -3pl:O -COLL
'I put the plates down one after another'

Departing from a verb + verb pattern, however, (as USPENSKIJ does in his analysis of *aq*) we could propose our own, verb-based, reconstruction: *da/r'a* could also be linked to the verb *da* meaning 'to distribute', 'to give'. This would tie in with KREJNOVIČ's formula 2', that is, we would have a pattern with two verbal roots.

6.2.4 Coverbal functions

Some of the elements given in KREJNOVIČ's lists for the R1 and R2 positions take on intransitivizing functions:

One prominent case is the verbal root *n* ('become'), which in R1 position serves as an intransitivizer. *eet'k*, when appearing in R2 position alone, means 'to show', but in combination with R1 ('become') the resulting compound verb is intransitive (inchoative) (KREJNOVIČ 1968 a: 123):

(69) *asl'ineŋ eet'k -a -R -a -n*
boat show -3sg:S -DET -PRET-become
'The boat appeared'

In a second example *n* joins the 'verbalizer' *t* (KREJNOVIČ 1968a: 201), which can also be interpreted as causative according to WERNER (1995: 52):

(70a) transitive: *Ø -di -ri -t*
3sg:n:S -1sg:O -hide -CAUS
'x is hiding me'

(70b) intransitive: *di -ri -t -n*
1sg:S -hide -CAUS -'become'
'I'm hiding'

Besides the *n* there is another morpheme for 'become', *a*, which also yields an inchoative reading:

(71a) transitive: *t -qav -i -t -Ø*
1sg:S -know -EV -CAUS -3sg:O
'I report on it'

(71b) intransitive: *Ø -qav -i -t -a*
3sg:n:S -know -EV -CAUS -become
'it becomes known'

The next interesting case is *dij/rij* which means literally: 'to grow'. This verbal root also seems to have taken over the function of intransitivization (KREJNOVIČ 1968a: 205).

(72a) transitive: *da -s'iraqa -d -da*
3sg:f:S -learn -1sg:O -CAUS
'she teaches me'

(72b) intransitive: *s'iraqa* *-d* *-dij*
learn -1sg:S -grow
'I learn'

(72a), containing a *da*-formative also could be linked to KREJNOVIČ'S hypothesis (1968a: 204) that in some cases *da-*, meaning 'give', could also be analysed as a causative marker - a phenomenon that, incidentally, can be observed in several languages of the world.

Another parallel development is that of *-tij/tej* ('to fall'):

(73a) transitive: *t* *-kav* *-i* *-t* *-Ø*
1sg:S -tear -EV -CAUS -3sg:O
'I'm tearing it'

(73b) intransitive: *Ø* *-kav* *-ä* *-tij*
3sg:n:S -tear -become(?) -fall
'it is torn'

It has been sufficiently shown in the literature on grammaticalization processes that besides 'be' and 'become' auxiliaries, lexemes meaning 'fall' can be sources for passivization. Thus, Ket *tij/dij* agrees well with Korean *ji-* ('fall'), Tamil *paṭu* ('fall') and some Bantu examples with similar meaning (cf. HASPELMATH 1990). In addition, some languages in a few of their idiomatic phrases display 'fall'-verbs as a means to express inchoativity and agentless expressions, for instance English 'to fall asleep' or German *trocken fallen* ('to become dry').

Also the verbal root *Rut*, literally 'to end' (KREJNOVIČ 1968a: 10) is instrumentalized for intransitivizing purposes:

(74a) transitive: *d* *-ul'tav* *-d* *-Ø* *-aq*
1sg:S -release -1s:SV -3sg:n:O -go
'I'm untying it'

(74b) intransitive: *Ø* *-ul'tav* *-ut*
3sg:n:S -release -end
'it is getting loose'

7. Some provisional results and further problems

It will not have escaped the reader's attention that we could only deal with a small fraction of the grammaticalizable and grammaticalized morphemes listed in KREJNOVIČ's formulae 1' to 3'. In fact, we have only dealt with about 25% of the 38 morphemes (and variants) listed there. But we hope, nevertheless, to have presented a framework which is capable of integrating the majority of the remaining morphemes that are either badly documented, unclear or controversial.

One interesting problem for the whole discussion so far has been raised by the general question of how the status of root morphemes can be best defined. In this connection the phenomenon called conversion plays an important role. WERNER (1997: 72 ff) proposes 10 subgroups of conversion, examples being (among others):

(75) verb: *sa:l* 'to spend the night'
and noun : *sa:l* 'overnight stay'

(76) verb: *an'ing* 'to play'
and noun: *an'ing* 'the game'

As far as the verb roots in 6.1 are concerned, we could choose the root *don'* which in KREJNOVIČ's analysis is glossed as 'to tear' (KREJNOVIČ 1968a: 102). On the other hand, some examples given above (in 5.2) featured the 'noun' *don'* 'knife'. If we now assume that a conversional relation exists between

(77) verb: *don'* 'to tear/cut'
and noun : *don'* 'the knife'

we are faced with some speculative questions. In 5.2.1.1 we quoted example (31), repeated as (78) here:

(78) *da -don' -ba -tet*
3sg:f:S -knife -1sg:O -hit
'she stabbed me with a knife'

Now, bearing in mind that *don'* also has a verbal counterpart, an interpretation and analysis such as:

(79) *da -don' -ba -tet*
3sg:f:S-cut -1sg:O-hit

- despite this unconventional approach - nevertheless fits into the Ket patterns, because we could now interpret the *tet*-component as generic, as proposed in 6.2.2. But, admittedly, this generates another problem:

(80) *don'-di-vet* (cf. (28))

must be analysed in the first place as a possessive construction: 'I have a knife'. We could now speculate as to whether the combination of R2 (*don'*) and R1 (*vet*) - maybe as part of a more complex verb fragment - could be placed within an aspectual framework such as that found in 6.2.3 where *vet* is identified as a marker of iterativity. Thus, we have to ask if a verbal form like

(81) *d-don'-aŋ-a-vet* (in analogy to *dejan'avet*, ex. 66b))

is possible in Ket (with the translation 'I cut repeatedly')[5].

In the end, these and further speculations show that it is not at all far-fetched to postulate a common origin for incorporative and verb + verb constructions as KREJNOVIČ and WERNER have done, but also that there is a number of conceivable interpretations and analyses.

ABBREVIATIONS

a = animate, ASP = aspect, B = B pronominal series, CAUS = causative, COLL = collective, D = D pronominal series, DET = determinative, EC = epenthetic consonant, EV = epenthetic vowel, ExV = extroverted (version), f = femininum, GNR = generic, IMPF = imperfective, inan = inanimate, ITER = iterative, IV = introverted (version), m = masculinum, n = neutrum, O =object, PAT = patient, PERF = perfective, PERM = permansive, pl = plural, PORT = portative, PRES = present, PRET = preterite, PROG = progressive, PST = past, QA = quotative aorist, R = root, S = subject, SEMEL = semelfactive, sg = singular, STAT = stative, SV = subject version, T = tense

REFERENCES

BAKER, M.C. 1996. *The Polysynthesis Parameter*. New York: Oxford University Press
BAMGBOSE, A. 1974. On serial verbs and their status. *Journal of West African Languages* 9: 17-48.
BELIMOV, E.I. 1991. *Ketskij sintaksis. Situacija, propozicija. predloženie*. Novosibirsk.

5 It should not be forgotten here that 'verbal' *don'* primarily (as part of KREJNOVIČ's formula 1') also can appear in R1 position. In his example (1968a:102) *d-u-n-e-b-don'* ('I tear it') *don'* (R1) combines with an adverbial R2, *u* (from *ugd'*, 'long'), yielding something like German: 'der Länge nach durchreißen' ('to tear in long strips'). This implies that R1-R2 configurations also include manner specifications that are brought about by the R2- part of the construction as pointed out by DUL'ZON.

BLAKE, B.J. 1994. *Case*. Cambridge Textbooks in Linguistics. Cambridge University Press.

BOAS, F. 1911. Introduction. In: BOAS, F. (ed.). *Handbook of American Indian Languages* I. Smithonian Institution. Bureau of American Ethnology 40. Government Priting Office, Washington, D. C.

BOUDA, K. 1957.Die Sprache der Jenisseier. Genealogische und morphologische Untersuchungen. *Anthropos* 52: 65-134.

BOUDA, K. 1968. Die Sprache der Jenisseier.II. *Orbis* 17/1: 158-186.

BOUDA, K. 1971. Die Sprache der Jenisseier.IV. *Orbis* 20: 143-156.

CASTRÉN, M.A. 1958. *Versuch einer jenissei-ostjakischen und kottischen Sprachlehre nebst Wörterverzeichnissen aus den genannten Sprachen.* St. Petersburg.

COMRIE, B. 1981. *The Languages of the Soviet Union.* Cambridge Language Surveys. Cambridge: Cambridge University Press.

DONNER, K. 1916. *Beiträge zur Frage nach dem Ursprung der Jenissei-Ostjaken. Journal de la Société Finno-Ougrienne*. 37: 1-21.

DROSSARD,W. 1998. Polysynthesis and polysynthetic languages in comparative perspective. *LP* 96 (*Linguistic and Phonetics*, Prague, B. Palek, ed.): 251-264.

DUL'ZON, A.P. 1968. *Ketskij jazyk*. Tomsk.

FISCHER, J.E. 1768. *Sibirische Geschichte*. 1. St. Petersburg.

FOLEY, W.A. 1986. *The Papuan languages of New Guinea.* Cambridge Language Surveys. Cambridge: Cambridge University Press

FOLEY, W.A. 1991. *The Yimas Language of New Guinea*. Stanford: Stanford Univerisity Press

FORTESCUE,M. 1994. Morphology, Polysynthetic. In: *The Encyclopedia of Language and Linguistics* (R.E. Asher & J.M.Y. Simpson, eds). vol 5: 2600-2602.

GAJER,R.S. 1983. Struktura i modeli protych glagolov ketskogo jazyka. *Voprosy enisejskogo i samodijskogo jazykoznanija*. Tomsk 1983: 62-61.

GEORGI, J.G. 1776-1780. *Beschreibung aller Nationen des Russischen Reichs, Lebensart, Religionen, Gebräuche, Wohnungen, Kleidungen und übrigen Merkwürdigkeiten.* St. Petersburg.

HASPELMATH,M. 1990. The grammaticalization of passive morphology. *Studies in Language* 14: 25-72.

JAKOVLEV, N.F. 1940. *Sintaksis čečenskogo literaturnogo jazyka.* Moskva/Leningrad. Akademia Nauk.

KARGER, N.K. 1934a. *Bukvar (bukvar na ketskom jazyke)*. Leningrad.

KARGER, N.K. 1934b. Ketskij jazyk. In: *Jazyki i pis'mennost' narodov severa III*:223-228. Leningrad.

KREJNOVIČ, E.A. 1960. Ugraženie prostranstvennoj orientacii v nivxskom jazyke. In: *Voprosy jazykoznanija* 9/1: 78-89. Moskva.

KREJNOVIČ, E.A. 1968a. *Glagol ketskogo jazyka*. Leningrad.

KREJNOVIČ, E.A. 1968b. Sposoby dejstvija v glagole ketskogo jazyka. *Ketskij sbornik Linguistika.* Moskau: 75-138.

MESSERSCHMIDT, D.G. 1964. *Forschungsreise durch Sibirien 1720-1727, Tagebuchaufzeichnungen Januar 1723- Mai 1724*. Berlin.

MITHUN, M. 1996. Overview of General Characteristics. In: GODDARD, I. (ed.). *Handbook of North American Indians*. vol. 17 (=Languages). Smithsonian Institution. Washington: 137-57

PALLAS, P.S. 1776. *Reise durch verschiedene Provinzen des Russischen Reiches, III.* St. Petersburg.

PAWLEY, A. 1994. Kalam. Exponents of Lexical and Semantic Primitives. In: GODDARD, C. / A. WIERZBICKA (eds.). *Semantic and Lexical Universals.* Amsterdam-Philadelphia: J. Benjamins: 387-422

ROOD, D. 1976. *Wichita Grammar.* New York: Garland Publishing.

ŠABAEV, V.G. 1986. Projavlenie čert aktivnosti i ėrgativnosti v ispol'zovanii ličnych pokazatelej Ketskogo glagola. In: *Paleoaziatskie jazyki*: 251-261. Leningrad.

SAPIR,E. 1911. *The problem of noun incorporation in American languages.* AmA 13: 250-82.

SCHULTZE-BERND, E. 2000. *Simple and Complex Verbs in Jaminjung.* MPI Series in Psycholinguistics: Nijmegen.

SWANTON, J.R. 1911. Haida. In: BOAS, F. (ed.). *Handbook of American Indian Languages.I*): 205-282.

TALMY, L. 1985. Lexicalization patterns: semantic structure in lexixal forms. In: SHOPEN, T. (ed.). *Linguistic Typology and syntactic Description.* Cambridge University Press. vol.III: 57-149.

USPENSKIJ, B.A. 1965. *Strukturnaja typologija jazykov.* Moskva.

USPENSKIJ, B.A. 1968. O systeme ketskogo glagola. In: *Ketskij sbornik Linguistika* Moskva: 196-228.

VALL, M.N. & I.A. KANAKIN. 1988. *Kategorii glagola v Ketskom jazyke.* Novosibirsk.

WERNER, H. 1989. *Novyj Ketskij alfavit.* Krasnojarsk.

WERNER, H. 1994. *Das Klassensystem in den Jenissei-Sprachen.* Veröff. der Societas Uralo-Altaica. 40. Wiesbaden: Harrassowitz.

WERNER, H. 1995. *Zur Typologie der Jenissej-Sprachen.* Veröff. der Societas Uralo-Altaica. 45. Wiesbaden: Harrassowitz.

WERNER, H. 1996. *Vergleichende Akzentologie der Jenissei-Sprachen.* Veröff. der Societas Uralo-Altaica. 46. Wiesbaden: Harrassowitz.

WERNER, H. 1997. *Die ketische Sprache.* (Tunguso-Sibirica 3). Wiesbaden: Harrassowitz.

WERNER, H. 1998. *Probleme der Wortbildung in den Jenissej-Sprachen.* Lincom Studies in Asian Linguistics. 25. München - Newcastle: Lincom Europe.

The rise and fall of polysynthesis in the Eskimo-Aleut family

MICHAEL FORTESCUE

1. Introduction

Eskimo-Aleut is arguably the most polysynthetic family in the world (*pace* BAKER 1996:19) as regards the number and variety of morphemes that can be packed into individual word-forms, and the fact that only one of its hundreds of bound suffixes can be reconstructed as an independent stem in the proto-language suggests that its polysynthesis is very ancient (see FORTESCUE 1994:2602). But how ancient is 'ancient'? The morphology of these languages was surely not always so complex. In this paper I shall sketch the route whereby Eskimo languages may have reached their present typological extremity from far less synthetic origins reflecting a middle-of-the-road agglutinative type quite familiar in northern Eurasia. A reasonably good case can be made for an actual genetic link to Uralic languages (cf. FORTESCUE 1998), but that is of secondary importance here - the point I shall be arguing is that Eskimo-Aleut derived from some proto-language not *typologically* very different from proto-Uralic[1]. Clues as to how this came about can be found via internal reconstruction and comparison with its Siberian neighbours. Also the *reduction* of polysynthesis that characterizes Aleut vis-à-vis Eskimo is revealing. That language appears to represent a rare example of the (partial) reversal of drift towards polysynthesis.

2. Suffixes in the polysynthetic complex

First, we need to compare the suffixal categories of both language families (neither has prefixes of any consequence). These are presented on Table 1. The actual categories found in modern Eskimo languages (and reconstructible for the proto-language) are taken as the starting point. Here 'i' indicates inflectional categories, 'd' ordinary derivational ones, and 'D' recur-

1 Note that Eskimo-Aleut, although polysynthetic, is of the double-marking rather than head-marking type typifying neighbouring North American Indian languages. Its Eurasian neighbours to the west (including Uralic) are dependent-marking, so it falls typologically (as well as geographically) between the two on this parameter. For further typological links between the two families, such as the prevalence of participial forms used as indicative verbs, see FORTESCUE (1998:78).

sively derivational ones (where more than one per word is possible). Slots marked with a star are filled in at least some branches of Uralic and categories in parentheses may be expressed (also) with auxiliary verbs. Categories d-g are the 'sentential' suffixes and k-q the v>v 'verbal extenders' of Eskimo (see FORTESCUE 1980, also for a list of actual exponent suffixes).

Following BERGSLAND (1997a) I take it that proto-Eskimo-Aleut was much closer to proto-Eskimo than to Aleut, which, although retaining a few archaisms lost in Eskimo, is generally characterized by significant restructuring due in large part to the collapse of the Eskimo-Aleut morphological ergative system. Modern Aleut is thus placed in a column of its own for comparative purposes. The distinction between inflectional and derivational suffixes is fairly easy to make for these languages: inflections are obligatory and form paradigms of portmanteau elements standing at the end of words, whereas derivational suffixes are not obligatory and do not form closed paradigms. What really distinguishes Eskimo from Uralic languages is not so much the particular categories involved, it is rather the number of their exponents and their recursivity (also Aleut lacks the widespread recursivity of the nominalizers and verbalizers of Eskimo). The more easterly Uralic languages have since become somewhat more synthetic (admitting further suffixal categories), but influence in the opposite direction has played a role in the European ones.

TABLE 1: *Suffixal categories*

	(Proto-)Uralic	(Proto-)Eskimo	Aleut
1. *Nominal*			
a. pure number [2]	i	i	i
b. possessor person	i	i	i
" number	i	i	i
c. case	i	i	i
d. nom. extenders (n>n)	d	d	d
e. nominal modifiers	d	D	d
f. nominalizers (v>n)	d	D	d
2. *Verbal*			
a. subject person	i	i	i
" number	i	i	i
b. object person	i*	i	i
" number	-	i	i
c. mood [3]	d	i	(i)
d. tense	(d)	d	(i)
e. epistemic modality	d)	d	(d)
f. subjective valuation	-	d	d

2 Also number of possessum.

3 Most mood markers were originally nominalizations (e.g. participles) in both families.

g.	conjunctional	-	d	-
h.	aspect	(d)	D	d
i.	degree [4]	-*	D	d
j.	manner	-	D	d
k.	root modality	-*	d	d
l.	conative	(d)	d	d
m.	desiderative	-*	d	d
n.	causative	d	d	d
o.	valency reducers	d	d	d
p.	negation[5]	-	D	i/-
q.	judging/saying	-*	d	-
r.	verbalizers (n>v)[6]	d	D	d

4 'Degree' is a matter of adjectival derivation in Uralic - but 'adjectives' in both families are nominalizations of verb stems.

5 Negation is expressed by an auxiliary verb in Uralic (as in the Finnish example below) and in most constructions by an enclitic based on a copular 'be' stem in Aleut (the *u-* in *ulax* below) - however, morphological negation of the Eskimo type is used in the present tense:

(i) FIN *he* they *ei-vät* not-PRES.3PL *ava-a* open-NONF (< **avata-k*)
'They do not open'

(ii) WG *tiki-nngil-aq*
come-NEG-3S.INDIC
'He hasn't come'

(iii) AL *awa-za-ĝ-ulax* vs. *awa-laka-x̂*
work-HAB-3S-NEG work-NEG.PRES-3S
'He never works' 'He isn't working'

6 Verbalizers include copular suffixes in Eskimo but not Aleut or Uralic (though a Ø-copular on nouns is found in some Uralic), whereas 'have' is expressed morphologically only in Eskimo-Aleut (for an example see note 19):

(i) FIN *varpunen* sparrow *on* be.3S *lintu* bird
'The sparrow is a bird'

(ii) WG *Jensi* Jens.ABS *ilinniartitsisu-u-vuq* teacher-be-3S.INDIC
'Jens is a teacher'

(iii) AL *uchiitilax̂* teacher *a-ku-q* be-PRES-1S
'I am a teacher'

Object inflections were originally only for the 3rd person (singular) and derive from corresponding nominal possessor inflections in both families (cf. JANHUNEN 1982:35 for Uralic and FORTESCUE 1995:65f. for Eskimo-Aleut). Thus compare the following (partial) West Greenlandic nominal possessed and transitive indicative verbal paradigms (stems *irniq* 'son' and *taku-* 'see' respectively):

Possessor \ Possessum:	singular	plural
1s	*irni-ra*	*irni-kka*
2s	*irnir-it*	*irni-tit*
3s	*irnir-a*	*irnir-i*
1p	*irnir-put*	*irni-vut*
2p	*irnir-si*	*irni-si*
3p	*irnir-at*	*irnir-i*

Subject \ Object:	3s	3p
1s	*taku-a-ra*	*taku-a-kka*
2s	*taku-a-t*	*taku-a-tit*
3s	*taku-a-a*	*taku-a-i*
1p	*taku-ar-put*	*taku-a-vut*
2p	*taku-ar-si*	*taku-a-si*
3p	*taku-a-at*	*taku-a-at*

The corresponding category in Aleut refers to any 'anaphoric' complement rather, not necessarily an object but also an anaphoric possessor of an object or an adjunct of a locative expression. Compare the following transitive sentences. Eskimo is again represented by West Greenlandic and the Aleut is Atkan ('REL' refers to relative case, which in Eskimo languages combines genitive and ergative subject function).

(1) *Jensi-p* *tuttu* *tuqup-paa*
WG Jens-REL caribou.ABS kill-3S/3S.INDIC
'Jens killed the caribou' [7]

(2a) *hla-s* *su-ku-ngis*
AL boy-PL take-PRES-3PL/3PL
'The boys have taken them'

7 Versus intransitive WG *niri-vugut* (eat-1PL.INDIC) 'we are eating', *niri-sar-pugut* (eat-HAB-1PL.INDIC) 'we eat', AL *awa-ku-x̂* (work-PRES-3S) 'he is working' but *awa-za-x̂* (work-HAB-3S) 'he works'.

vs. (2b) *hla-s* *karandaasi-s* *su-ku-s*
boy-PL pencil-PL take-PRES-3PL
'The boys have taken the pencils'

and with anaphoric possessor adjunct:

(3) *Piitra-m* *ada-ngis* *kidu-ku-ngis*
Peter-REL father-3PL help-PRES-3PL/3PL
'Peter is helping their father' [8]

The language I have taken to represent Uralic, Finnish, does not have a transitive 'objective' paradigm, but has the usual accusative clause type[9].

3. The relative order of suffixes

On Table 2 can be seen the relative order of the suffix categories on Table 1 for Uralic and Eskimo-Aleut in respectively the nominal and the verbal complex. Optional categories are

---

8 Versus non-anaphoric (specified) *Piitra-x̂ hla-s ada-a kidu-ku-x̂* (Peter-ABS boy-PL father-3S/PL help-PRES-3S) 'Peter is helping the father of the boys'.

9 As in *(minä) kirja-n sa-i-n* ((I) book-ACC/GEN receive-IMPERF-1S) 'I received the book', versus intransitive *laula-a* (< **-va*) (sing-PRES.3S) 'he sings/is singing'. However, observe the following forms from Samoyedic Nganasan, completely semantically parallel to the West Greenlandic above (stems *kətu* 'nail' and *kontu-* 'carry', past tense, respectively):

Possessor \ Possessum:	singular	plural
1s	*kətu-mə*	*kəða-nʲe*
2s	*kətu-rə*	*kəða-ce*
3s	*kətu-ðu*	*kəða-cü*
1p	*kətu-mu'*	*kəða-nʲü'*
2p	*kətu-ru'*	*kəða-cü'*
3p	*kətu-ðuŋ*	*kəða-cüŋ*

Subject \ Object:	3s	3p
1s	*kontu-dʲuə-mə*	*kontu-dʲuə-nʲe*
2s	*kontu-dʲuə-rə*	*kontu-dʲuə-ce*
3s	*kontu-dʲuə-ðu*	*kontu-dʲuə-cü*
1p	*kontu-dʲuə-mu'*	*kontu-dʲuə-nʲü'*
2p	*kontu-dʲuə-ru'*	*kontu-dʲuə-cü'*
3p	*kontu-dʲuə-ðuŋ*	*kontu-dʲuə-cüŋ*

All combinations of dual as well as plural possessor/possessum and subject/object are also possible, as in older forms of Eskimo.

in parentheses, conflations connected by pluses and optional choices by obliques. As can be seen, there is rather little difference between the languages (the relative order of possessor and case in proto-Uralic is debated). There was no original nominal compounding in either family. Eskimo 'sentential' suffixes appear between square brackets.

TABLE 2

	Nominal categories
Uralic:	Stem (or v. stem - 1f) - (1d) - (1e) - 1a(+1c) - (1b) (but Ugric 1a(+1b) - (1c))
Eskimo:	Stem (or v. stem - 1f) - (1d) - (1e) - 1a - (1b) - (1c)
Aleut:	Stem (or v. stem - 1f) - (1d) - (1e) - 1a - (1b)(+1c) [10]
	Verbal categories
Uralic:	Stem (or n. stem - 2r) - (2n)/(2o) - (2h) - (2c)/(2d) - (2b+)2a
Eskimo:	Stem (or n. stem - 2r) - (2h-q) - [(2d) - (2e) - (2q) - (2f/2g)] - 2c - (2b+)2a
Aleut:	Stem (or n. stem - 2r) - (2h-q) - (2e) - (2q) - (2c)/(2d) - (2b+)2a

Note that 'stem' can cover anything from a simple root to a complex derived base (multiply verbalized/nominalized) in Eskimo but only the former in Uralic and Aleut[11]. Possessor/subject/object number follows person within 1b, 2a and 2b in so far as these are historically transparent. 1a marks possessum number if 1b is present, but the distinction is conflated (ambiguous) in Aleut for 3rd person. The relative case marker precedes 1b in EA except for 3rd person possessor, which it follows. 2a follows 2b in EA for 3rd person object, other combinations are the reverse (later accretions in Eskimo that probably never developed in Aleut). The relative order of categories 2h-q in Eskimo is constrained by the (more or less) non-recursivity of categories 2k-n (the succession of suffixes in EA is in general a matter of cumulative semantic scope).

The following simple verb forms in West Greenlandic and Finnish illustrate the difference between the two families as regards the expression of tense and epistemic modali-

10 But oblique case is expressed by postpositions after ordinary nouns in Aleut.

11 Thus the following construction in West Greenlandic, where the derivational suffix *-niit-* has 'incorporated' a noun phrase consisting of possessor plus possessum (the latter inflected for person/number):

nirrivi-up	*qaa-va-niip-put*	*baaja-t*	*pingasut*
table-REL	top-3S-be.in/on-3PL.INDIC	pilsner-PL	three

There were three (pilsner) beers on the table'

For the nearest (analytic) Aleut equivalent compare the first Aleut sentence under section 5.

ty, these both being derivational (and optional) in WG, as opposed to mood, which is an obligatory inflectional category. Contrast this with Finnish, where tense and mood stand in a paradigmatic (inflectional) relationship (basically imperative vs. conditional vs. a tensed indicative form), although epistemic modality - the so-called 'potential mood' - is arguably derivational:

(4) *(minä) sano-i-n*
FIN (I) say-IMPERF-1S
'I said'

(5) *syö-köön*
eat-3S.IMP
'Let him eat!'

(6) *naura-ne-vat*
laugh-POT-PRES.3PL
'Perhaps they will laugh'

(7) *taku-ssa-agut*
WG see-FUT-1PL.INDIC
'We'll see each other'

(8) *pitsaa-gunar-puq*
be.good-probably-3S.INDIC
'It is probably good'

This difference can be explained in terms of grammaticalization: tense (past vs. non-past) has become absorbed into the verbal inflection in Uralic in a way that has not (yet) happened in Greenlandic[12]. However it has in fact been absorbed in some western forms of Eskimo and in Aleut (where present versus past paradigms must be chosen). This gradual 'left-to-right' categorial shift is still in process today (see FORTESCUE 1996).

Derivation in Finnish is fairly restricted, but covers at least the following productive types:

12 The present (from an original participle) also marks the future in Finnish, whereas the indicative in West Greenlandic marks the present or the past (future must be derivationally marked). The present in Aleut can also refer to the immediate past.

syö-tta-	'feed'	(cf. *syö-* 'eat')
kansoi-tta-	'populate'	(cf. *kansa* 'people')
sairaa-la	'hospital'	(cf. *sairas* 'ill (person)')
elä-mä	'life'	(cf. *elä-* 'live')
juokse-ntella-	'run about'	(cf. *juokse-* 'run')
kuul-u-	'be heard'	(cf. *kuule-* 'hear')
hyvä-ksy-	'approve of'	(cf. *hyvä-* 'good')

The potential of derivation is far greater in EA, as the following examples illustrate:

(9) *aju-nngit-su-liur-vigi-nnit-tuar-tu-u-nngil-aq*
WG be.good-NEG-PART-make-have.as.place.of-ANTIPAS-all.the.time-PART-be-NEG-3S.INDIC
'He is not (much of) a benefactor'

(10) *iqquus-sinnaa-junnaar-sur-aa*
be.at.peace-can-no.longer-think.that-3S/3S.INDIC
'He thought he (another) could no longer be at peace'

(11) *(tama-nut) isiginnaa-ga-ssa-nngur-tit-tar-paa*
(all-ALL) view-PAS.PART-FUT-become-cause-REP-3S/3S.INDIC
'He put him on show (in front of everyone)'

(12) *(qisus-suaq) qussar-palu-inna-li-q-aa*
(tree-big.ABS) ram.wedge.into-be.heard-only-INCH-very.much-3S/3S.INDIC
'Someone could be heard just ramming a wedge hard (into a big tree-trunk)'

(13) *chala-si-isix̂*
AL clam-dig.out-instrument.for
'clam-digging stick'

(14) *hidu-usa-ku-u*
go.out-do.with-PRES-3S/3S
'He took it out'

(15) *ting adalu-usa-naaĝi-iĝuta-masu-x̂ta-ku-x̂* [13]
I lie-do.with/for-try-again-perhaps-PERFECT-PRES-3S
'Perhaps he has tried to fool me again'

However, this is basically just a matter of 'more of the same', not of deep structural differences. The same can be said of the multiplication of nominal cases in Uralic: the starting point was a smaller set similar to those of the Eskimo languages (cf. JANHUNEN 1982:30).

4. Stages in the genesis of Eskimo-Aleut polysynthesis

Proto-Eskimo-Aleut, to judge by the number and types of suffixes reconstructible for it, appears to have been essentially as polysynthetic as West Greenlandic is today (cf. FORTESCUE et al. 1994). By what earlier stages, then, could it have developed out of a more modestly agglutinative type like that of the Uralic languages? This question appears to be tied in with the origin of ergativity in the family[14].

Ergativity arose in Eskimo-Aleut not from a passive construction that underwent a 'markedness shift' (the most common source of the phenomenon) but from the merging of two pre-existing constructions, namely a possessed passive participle and an ordinary nominative/accusative S-O-V clause structure (for definite object), which blended to produce the ergative clause glossable as 'X was the V-ed thing of Y' (cf. FORTESCUE 1995). The first construction contributed the morphology, the second the word order (and the introduction of an object constituent). This can be schematically represented as follows, using the same abbreviations as in the tables above, with hyphens between morphemes in single words (note that 2b in the resultant construction is historically identical with 1b, hence the 'equals' sign):

	(i)	Possessor NP (gen./relative case)	Possessum NP	(pass. participle)-1b
plus	(ii)	Subject NP (nom.)	Object NP (acc.)	Verb-2a
→	(iii)	Subject NP (erg./rel.)	Object NP (abs.)	Verb-1b=2b

An example of the resulting transitive clause was given in sentence example (1), with which compare the possessed NP *Jensi-p tuqu-ta-a* (J.-REL kill-PAS.PART-3S) 'Jensi's killed thing'.

---

13 This is about the limit of word-expansion in Aleut. Such Eskimo suffixes as WG *-sur(i)-* 'think that', for example, in the second WG sentence above, which can incorporate whole tensed propositions, have no counterpart in Aleut.

14 Note that the two parameters have nothing intrinsically to do with each other. In fact, NICHOLS has shown that statistically ergativity and head-marking (the latter the most salient feature of polysynthetic languages) tend *not* to go together, ergativity being associated rather with dependent-marking languages (NICHOLS 1992:113).

At a later stage 1st, 2nd and reflexive 3rd person pronoun objects were also enclitically attached to the verb in this construction (and ultimately absorbed into portmanteau inflections in Eskimo at least).

However, where the verb concerned was 'have' or 'get' (or others corresponding to semantically transitive 2r suffixes), which often occur with indefinite objects, an incorporating construction may have developed as a way of distinguishing non-definiteness of object:

	Subject NP (nom.)	Object NP (acc.)	Verb-2a
→	NP (nom.)	Incorp. object-(bound)	verb-2a

Thus: (16) *piniartuq* *(marlun-nik)* *savi-qar-puq*
hunter.ABS (two-ALL.PL) knife-have-3S.INDIC
'The hunter has a knife (two knives)'

This fell in line with the intransitive (non-ergative) construction that remained with nominative case subject - corresponding to zero-marked absolutive - which now contrasted with relative case transitive subject[15]. Only the (stem of) the head of the object phrase was incorporated, leaving any modifier of the object (assuming case agreement with the head) stranded in the instrumental case, as in parentheses in (16). This corresponded presumably to the original accusative (there is some comparative evidence drawing upon Uralic for this - see FORTESCUE 1998:116).

At some subsequent stage verb-extending and negative auxiliaries were also incorporated into the verb:

	Main verb (non-finite form)	Aux. verb-2a
→	Verb stem-deriv. suffix-2a	

Most if not all such auxiliaries were probably based on copula-type stems, which also resulted in Eskimo 2r 'be' suffixes, parallel with the 'have/get' ones above, only intransitive (and with absolutive case stranded modifiers). Compare thus WG *-sinnaa-* 'can' from nominalizer *-sinnaq* 'only' + *-u-* 'be', and negative suffix *-nngit-* from nominalizer *-niq* + *-(ng)it-* 'be without'. Such auxiliaries could have drawn their own adverbial and aspectual modifiers along into the verbal complex as 'verbal modifying' suffixes.

Then, via word-internal grammaticalization shift 'from left to right', non-recursive 'sentential' suffixes arose out of earlier suffixes of aspect, degree, and root modality, etc. In

15 Also the object of antipassivized verbs (and transitive 'agentive' verbs like *niri-* 'eat') came naturally to follow the same pattern of indicating indefinite objects with the instrumental case, as in *niqi-mik niri-pput* (meat-INSTR eat-3PL.INDIC) 'they ate meat'.

other words, certain derivational suffixes with broad word-internal scope - situated adjacent to the inflection - become reinterpreted as taking sentential scope:

Verb stem(-deriv. $suffix_1$)-deriv. $suffix_2$-2a
→ Verb stem(-deriv. $suffix_1$)-sentential suffix-2a

Compare the source of WG *-ssa-* 'objective future' in *-ja-* 'be apt to', of distant future *-jumaar-* in *-juma-* want to', and of epistemic *-sima-* 'probably' in aspectual *-(si)ma-* 'perfective state' (all of which are still more or less productive in that language). This stage was probably never reached by Aleut (i.e. it probably post-dated the split between the two branches). Almost certainly the final stage of full morphological recursivity, with the extension of the category 'stem' to include complex or phrasal constituents as derivational 'base', was never reached by Aleut[16].

5. Stages in the reduction of Aleut polysynthesis

When one compares Aleut to any Eskimo language and to the Uralic languages, one is initially struck by certain ways in which Aleut seems closer to Uralic than Eskimo does. The principal ways can be enumerated as follows:

No recursivity in morphology
Only 3rd person objects in transitive paradigms
Nominative-accusative alignment (only marginal ergativity in Aleut)
Non-derivational negation (by enclitic or aux. verb)
Auxiliaries for tense/aspect/modality (also Ø-markers)
Rigid word order
Tense and mood paradigmatic (non-compatible)
Independent copular verb

16 Nothing would seem structurally to preclude the multiple succession of nominalizers and verbalizers in Aleut, though there is no trace of it having occurred earlier. Note that this - as well as derivation based on phrasal bases - is somewhat more restricted in western forms of Eskimo than in Greenlandic. Aleut subjective valuation suffixes are in a sense recursive (can be piled up successively) but this is a matter of the building up of nuanced combinations within the same category, not of the same suffixes occurring at more than one (potentially separated) level of a complex word. Suffix *-za-/-da-* of repeated action can indeed appear twice in one word, but (like the cognate suffix in Eskimo) this is because it may fulfill two different functions, either aspectual modification of the base or habitual tense relating to the whole sentence.

These could be taken as reflecting ancient retentions in Aleut, but - as BERGSLAND came to realize - they are all except the last the result rather of later restructuring in Aleut. Thus he suggests that Aleut may have lost non-3rd person object marking on the verb in the (originally) transitive construction by dislocation of the relevant pronouns, which may only have been loosely enclitic at the proto-EA stage (BERGSLAND 1997:12). They would at all events not have been consonant with the transformed function of the transitive verb construction in Aleut (a matter of anaphoric versus non-anaphoric rather than presence versus absence of object). Following BERGSLAND's lead let me trace the domino-like stages of that restructuring.

Apocope appears to have led to the merging of the instrumental and locative (*-*mi* and *-*mə̆ng* respectively) with relative case *-m*, and to have contributed also to the reduction of the original nine-way distinction of 3rd person possessor/possessum combinations (and hence also of transitive subject/object combinations on the verb[17]). After the collapse of the core case distinctions the ergative/antipassive distinction broke down for the semantically transitive clause. This heralded a shift to a distinction between anaphoric and non-anaphoric clauses, in which original EA morphological structure was partially preserved but its function was radically changed (compare examples (1) and (2a)/(2b) under section 2 above). Thus the 'specified' (non-anaphoric) construction was opposed to the anaphoric one (with no overt object NP), which 'remained' but in the new function:

		Subject (rel.)	Object (abs.)	Verb-2b-2a
	→	Subject (abs.= nom.)	Object (abs.= nom.)	Verb-2a
as opposed to:		Subject (rel.)	(Anaph. object)	Verb-2b/2a

The second type was extended to include other anaphoric complements, as in the following, where the original 3S subject/3S object inflection now combines 3S subject and 3S anaphoric adjunct of the postposition:

(17a) *qalgadax̂ ku-gan a-ku-u*
food.ABS top-3S.LOC be-PRES-3S/3S
'There is food on it'

17 Thus *kidu-ku-u* is either 'he is helping him' or 'they are helping him', and *kidu-ku-ngis* is either 'he is helping them' or 'they are helping them'. The corresponding Eskimo 3rd person markers are, respectively, *-(ng)a* and *-(ng)i*, and *-(ng)at* and *-(ng)it*. As regards the first pair, note that the only final vowels in Aleut are now single /a/ (with very few exceptions) and double /ii/ and /uu/ (usually from */aa/, as in *-ku-u* above). Vowels in general have undergone much change in Aleut, including via assimilation at a distance. BERGSLAND (1997b:351) specifically suggests that original 3PL/3S *-*ngi* may have fallen together with 1S/1S *-ng* by apocope, thus opening the way for the generalization of 3PL/3PL *-ngis*.

vs. (17b) *stuuluĝi-m* *ku-gan* *qalgadax̂* *a-ku-x̂*
table-REL top-3S.LOC food.ABS be-PRES-3S
'There is food on the table'

Thereafter transitivity was no longer a morphologically unambiguous category, e.g. as regards the distinction between passive and active participles, and various constructional blends arose that also preserved EA structures but assigned new functions to them. Thus there arose a 'new' head-modifier construction, where the morphology comes from the ordinary possessed NP construction but now attached to the head plus modifying participle constituents of the ordinary (non-possessed) modified NP construction (as in Eskimo):

	(i)	Possessor (rel.)	Possessum-1b
plus	(ii)	Head noun (abs.)	Modifier nominal/participle (abs.)
→	(iii)	Head noun (rel.)	Modifying part.-1b

Thus: (18) *tayaĝu-m* *aygag-na-ngis*
man-REL walk-PART-3PL
'walking men' [18]

This is literally (from an EA point of view) 'a man's their walking' - compare *tayaĝu-s aygagi-ku-s* (man-PL walk-PRES-3PL) 'the men are walking', with ordinary subject number.

Then, with the fixation of word order to distinguish absolutive subjects and objects, and the cumulative marking of anaphoric 'adjuncts' on the final verb (whatever their 'internal' relationships within the complex sentence), the tendency to increased morphological analyticity accelerated, resulting in auxiliary verb constructions that correspond to most of the EA 'sentential' suffixes of the type:

	Verb-suffix-2c-2a	
→	Main verb (non-fin.)	Aux.-suffix(-2c/d)-2a

This was not literally a matter of sentential suffixes becoming 'detached' from the head verb to appear as auxiliaries, but of suffixes with these meanings (relating to the whole outer sentence) that followed anaphoric marking to the end of the sentence by becoming attached to copular stem *a-* (which, it will be recalled, had probably always remained an independent stem in Aleut). The main verb was put in the conjunctive (later also participial or intentional)

---

18 Compare WG *tuluk kalaalli-suur-tuq* (Englishman Greenlander-do.like-PART) 'a Greenlandic-speaking Englishman' and *tulu-up kalaalli-suur-nir-a* (Englishman-REL Greenlander-do.like-NOM-3S) 'the Englishman's speaking Greenlandic'.

mood. This was patterned on pre-existing EA constructions with a coordinate verb in the contemporative (= Aleut conjunctive) that preceded the main verb in coordinated and object clause constructions[19]. The auxiliary inherited any anaphoric (e.g. transitive object) marking of the main verb, irrespective of its own transitivity. Later other verbs took on auxiliary function of this kind, in particular ones indicating a more precise tense relationship. Thus:

(19) *saĝa-l* *a-x̂ta-ku-x̂*
sleep-CONJ be-apparently-PRES-3S
'He is apparently sleeping'

(20) *ukux̂ta-l* *saĝa-qa-ng*
see-CONJ sleep-PAST-1S/3S
'I saw him yesterday'

(21) *ma-angan* *aĝi-ku-ng*
do-INT.1S put-PRES-1S/3S
'I'll do it now'

Some of these analytic constructions with copular *a-/u-* are again becoming fused enclitically with the stem (perhaps even attaining suffixal status) in modern Atkan Aleut, as if the same process described 'on the way up' for EA verbal extenders is beginning to repeat itself. The Aleut enclitic negative *ulux/ulax* (based on copular *u-* 'be') probably had a similar origin as a form of the negated copula. Note that it occurs (actually or previously so) with participial, i.e. nominal forms, as opposed to the older morphological negative present tense form *-la-* (*ka-x̂*, etc.) shared with Eskimo.

Furthermore, in constructions where EA morphology and syntax were already intertwined, in particular those involving 2r suffixes with meanings such as intransitive 'have' as opposed to transitive 'have as', anomalous constructions resulted from the breakdown of transitivity distinctions:

(22a) *ukina-ĝi-ku-x̂*
knife-have-PRES-3S
'He has a knife'

---

19 As in Aleut *hatax̂-s ayu-ku-x̂* (slip-CONJ fall-PRES-3S) 'he has slipped and fallen'. Compare WG *sinik-kunar-puq* (sleep-apparently-3S.INDIC), and *qungujul-luni sinip-puq* (smile-4S.CONT sleep-3S.INDIC) 'he is smiling as he sleeps'.

vs. (22b) *alax* *ukina-x̂ta-ku-x̂*
two.ABS knife-have.as-PRES-3S
'He has two knives'

Compare the WG construction with intransitive *-qar-* (the equivalent of Aleut *-ĝi-*) in sentence (16)[20].

Nor could the intransitive 'have' suffix be used (any more) in indefinite expressions with participial bases like in Eskimo, for example WG *qallunaa-nik pulaar-tu-qar-puq* (Dane-INSTR.PL visit-PART-have-3S.INDIC) 'there are (Danish) visitors'. This was presumably because transitivity marking had been disturbed (and the possibility of a stranded instrumental case modifier no longer existed). Instead, an analytic construction with *a-* was necessitated, in the anaphoric case resulting in another odd 'blend':

(23) *uchiitilax̂* *hnux̂ta-na-s* *a-na-s*
teacher.ABS visit-PART-PL be-PAST-3PL
'(There were) some (who) visited the teacher' [21]

(24) *hnux̂ta-na-m* *a-qa-a*
visit-PART-REL be-PAST-3S/3S
'Someone visited him'

Further 'anomalous' consequences of these successive restructurings are seen in 'chained clauses' such as in (25) below. Here there is anaphoric reference belonging semantically to the first clause marked on the second - compare (26) where the reverse holds (*-ng* goes with *-i* 'his', referring to 'outer subject' *tayaĝux̂*, unmarked on the final semantically transitive verb):

---

20 The transitive morpheme corresponding to Aleut *-x̂ta-* can be seen in the following WG example:

taanna *qaja-ri-vara*
that.ABS kayak-have.as-1S/3S.INDIC
'I have/use that as a kayak'

The object *taanna* could not be replaced by *marluk* 'two' here, as in corresponding intransitive sentence (16). Compare also *qaja-ri-sa-ra* (kayak-have.as-PAS.PART-1S) 'the one I have as (= my) kayak', Aleut *ada-x̂ta-na-ng* (father-have.as-PART-1S) 'the one I have as (= my) father'.

21 Compare also Finnish *ilma-ssa on pihka-n tuoksu-a* (air-IN be.3S resin-GEN smell-PARTI) 'there is a smell of resin in the air'.

(25) *kuusxi-m* *su-l* *amaana-qa-a*
cat-REL grab-CONJ run.away-PAST-3S/3S
'The cat grabbed it and ran away'

(26) *tayaĝûx* *ilga-ku-ng* *hla-hli-i* *uku-na-q*
man.abs look.for-PART-1S/3S son-only-3S find-PAST-1S
'I looked for the man but found only his son'

Such phenomena call for an analysis in which the 'inner' syntactic level - the verb with its subject and object and complements - is opposed to a parallel 'outer' (pragmatic) one, whether this be defined in terms of 'outer subject' in the manner of BERGSLAND, or of 'Given Topic' in a Functional Grammar framework, say. Whatever one's terminological preferences, the distribution of morphemes marking anaphoric reference to the 'outer' level is something hard to reconcile with any autonomous approach to grammar. Whether treated as an agreement phenomenon or in terms of pronominal cross-reference, the structural configurations and 'nodes' that are needed to invoke the relevant 'migration' of morphemes cannot be confined to the strictly syntactic but must refer to at least one pragmatic category.

6. Conclusion

I have argued elsewhere (FORTESCUE 1992:245) that once the rubicon to polysynthesis is passed - and derivational morphology and (inflectionally marked) syntax become inextricably intertwined - the development can only be 'reversed' by radical restructuring that may result in long-lasting typological anomalies, anomalies of the sort we do indeed see in the case of Aleut. But how much of the 'anomalous' nature of Aleut is due to the reduction of polysynthesis as such and how much to the collapse of ergativity?

Aleut was undoubtedly once much more polysynthetic than it is today: witness the approximately 570 derivational suffixes counted by BERGSLAND, of which only rather few (175 at most) are productive today - the rest are largely opaque relics of an earlier much more productive stage still reflected in the Eskimo languages (BERGSLAND 1997b:105). It is conceivable that if the language were (by some miracle) to live on well into the 21st century it would become still more analytic through massive relexicalization with English loans and/or continuing lexicalization of combinations of native stems plus suffixes. But it is also possible that reabsorption of auxiliary verbs into the verb complex (as new suffixes) would compensate for this to some degree. We shall probably never know. It may be that all surviving polysynthetic languages will drift towards decreasing polysynthesis as stem vocabulary increases (via loans or lexicalized native combinations) in order to keep abreast of new 'international'

concepts[22]. However, stem/suffix balance in the lexicon is one thing, the productivity of the morphology another. What one needs to look at, amongst other things, is the degree of recursivity of that apparatus - definitely much less in Aleut than in Eskimo - and the complexity of syntax versus morphology in the overall grammar - definitely greater in Aleut than in Eskimo.

Now it may well be that the collapse of EA ergative morphology in Aleut was the direct cause of the widespread restructurings that typify that language. But why should such a development lead to typological abnormality? Languages throughout the world move back and forth between ergative and nominative/accusative alignment, for example, rather smoothly and rapidly as markedness shifts and natural drift do their work. Perhaps it is the rather unusual combination of ergativity and polysynthesis that has made the transition so radical. Certainly one can envisage the entanglement of morphology and syntax (or of derivation and inflection), which is typical of polysynthetic languages, interfering with the relatively smooth readjustment of syntax and morphology necessitated by a transition between major alignment types. The properties of derivational morphemes geared to ergative systems (e.g. the 'have' versus 'have as' suffixes of EA) will also have to undergo radical adjustment once the overall external syntax shifts alignment. Recall from section 4 above how ergativity also played a central in the *development* of polysynthesis in EA. These two most prominent features of EA have been closely intertwined in both the rise and the fall of polysynthesis within the family.

But perhaps there were also 'external' factors that speeded up the unusual developments within Aleut, for example the decimation by epidemic and geographical displacement at the hands of the Russian fur traders during the early contact period. This was probably too late to be the (main) cause, however - the earliest records of Aleut in the late 18th century already display its characteristic un-Eskimo traits. These upheavals did lead - or contribute - to the disappearance and mixture of earlier dialects, but that is another matter altogether[23]. A more likely - and long-lasting - source of influence on developments here was substratum and/or areal effects. Thus 'promiscuous number marking', whereby the plurality of *any* referent in a sentence is marked on the verb, overriding any singular marking, as in neighbouring coastal Na-Dene languages and Haida (cf.LEER 1991), could be directly involved in some of the odd number agreement phenomena in Aleut. In the following example the number of the adjunct of the subject is marked on the verb rather than that of the subject itself:

---

22 There is indeed evidence of this happening in West Greenlandic in so far as the stem lexicon has increased considerably in recent years thanks to loans from Danish, but this has not resulted in any noticeable decrease in the role suffixes play in the language. The fully productive suffixes of the language are as productive as ever, although the number of exemplars of a given suffixal category that are still fully productive may be decreasing as bilingualism increases. The piling up of especially subjective valuation suffixes is less typical of the language today than in the old oral tradition.

23 The case of Copper Island creole is special and can not be transferred to Aleut as a whole (see GOLOVKO & VAKHTIN 1990 for the sociolinguistic background).

(27) *ada-ngis* *awa-ku-s*
father-3PL work-PRES-3PL
'Their father is working' [24]

What looks at first sight like features that Aleut (as opposed to Eskimo) shares with Uralic languages may thus reflect not so much ancient retentions but the 'undoing' of some of the processes that led up to polysynthesis in the family in the first place - an internally dictated development spurred on by substratum/areal influence. The most crucial of these processes involved the emergence of ergativity, which interlocked the inflectional syntax and the derivational morphology of EA in such a way that radical restructuring was incurred by its collapse. This resulted in the typological anomalies of Aleut syntax that have still not been ironed out by normal processes of drift to this day.

ABBREVIATIONS

ABS	absolutive	INSTR	instrumental
ACC	accusative	INT	intentional
ADESS	adessive	LOC	locative
AL	Aleut	NEG	negative
ALL	allative	NONF	non-finite
ANTIPAS	antipassive	PART	participle
CONJ	conjunctive	PARTI	partitive
CONT	contemporative	PAS	passive
EA	Eskimo-Aleut	PAST	past
FIN	Finnish	PERFECT	perfective
FUT	future	PL	plural
GEN	genitive	POT	potential
HAB	habitual	PRES	present
IMP	imperative	REL	relative
IMPERF	imperfect	REP	repetitive
INCH	inchoative	S	singular
INDIC	indicative	WG	West Greenlandic

24 Compare also with auxiliary *a-* 'be' in the analytical modal sentence type:

Piitra-m *uku-qa-ngis* *a-x̂ta-ku-s*
Peter-REL find-PAST-3S/3PL be-apparently-PRES-3PL
'Peter found them, apparently'

REFERENCES

BAKER, MARK C. 1996. *The Polysynthesis Parameter.* New York/Oxford: Oxford University Press.

BERGSLAND, KNUT. 1997a. How did the Aleut language become different from the Eskimo languages? In: *Languages of the North Pacific Rim,* vol. 2, eds. Osahito Miyaoka & Minoru Oshima. Kyoto: Kyoto University Graduate School of Letters, 1-17.

BERGSLAND, KNUT. 1997b. *Aleut Grammar.* Fairbanks: Alaska Native Language Center.

FORTESCUE, MICHAEL. 1980. Affix ordering in West Greenlandic derivational processes. *International Journal of American Linguistics* 46 (4): 259-278.

FORTESCUE, MICHAEL. 1992. Morphophonemic complexity and typological stability in a polysynthetic language family. *IJAL* 58 (2): 242-248.

FORTESCUE, MICHAEL. 1994. Polysynthetic morphology. In: *The Encyclopedia of Language and Linguistics,* vol. 5, eds. R. E. Asher et al. Oxford: Pergamon Press, 2600-2602.

FORTESCUE, MICHAEL. 1995. The historical source and typological position of ergativity in Eskimo languages. *Etudes/Inuit/Studies* 19 (2): 61-75.

FORTESCUE, MICHAEL. 1996. Tense, mood and aspect grammaticalization in West Greenlandic and Chukchi. In: *La dynamique dans la langue et la culture inuit,* eds. N. Tersis & M. Therrien. Paris: Peeters, 151-175.

FORTESCUE, MICHAEL. 1998. *Language Relations across Bering Strait: Reappraising the Archaeological and Linguistic Evidence.* London & New York: Cassell Academic.

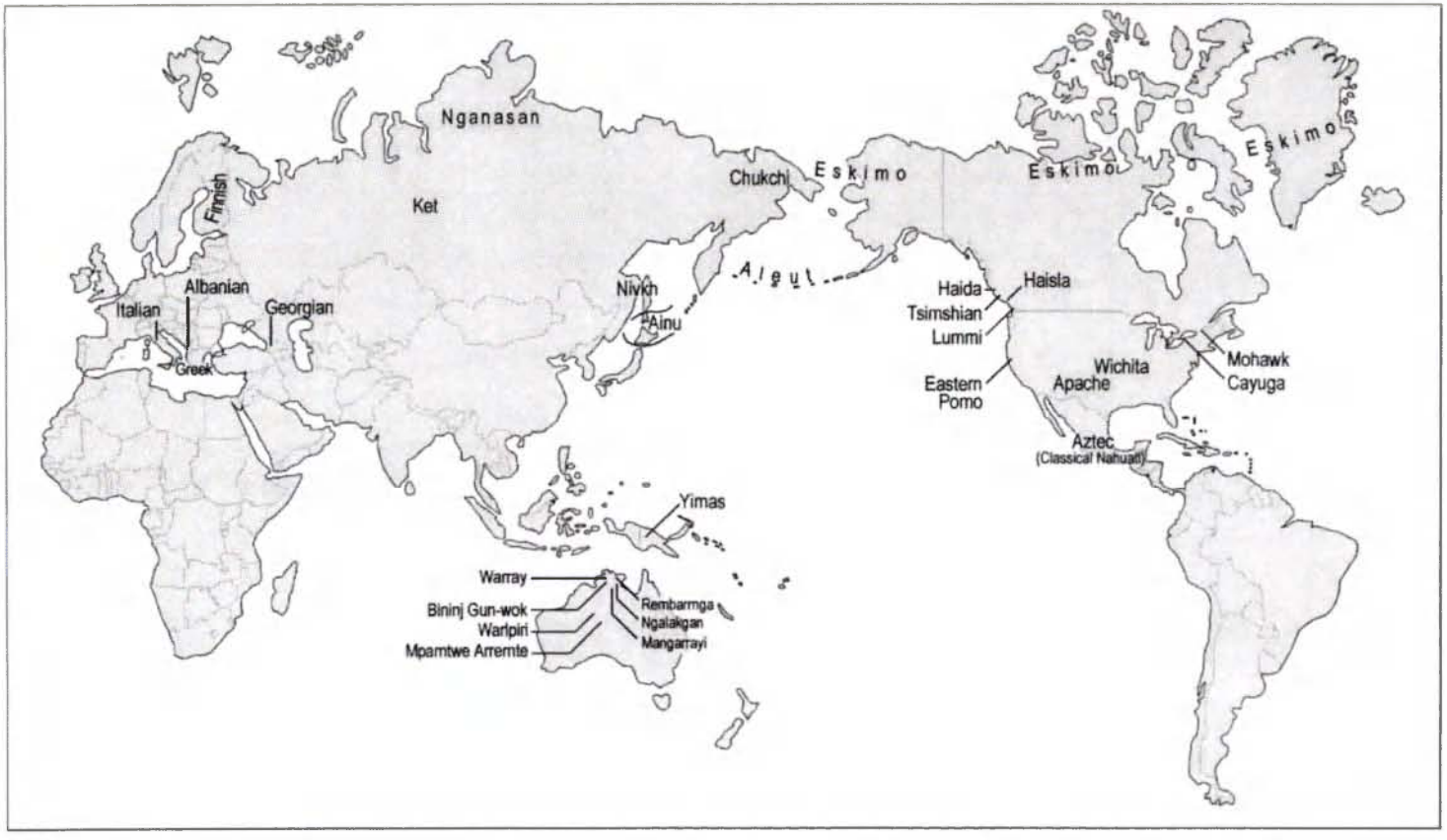
FORTESCUE, MICHAEL and LAWRENCE KAPLAN and STEVEN JACOBSON. 1994. *Comparative Eskimo Dictionary with Aleut Cognates.* Fairbanks: Alaska Native Language Center Press.

GOLOVKO, EVGENIJ and NIKOLAI VAKHTIN. 1990. Aleut in contact: the CIA enigma. *Acta Linguistics Hafniensia* 22: 97-125.

JANHUNEN, JUHA. 1982. On the structure of proto-Uralic. *Finnisch-ugrische Forschungen* 44: 23-42.

LEER, JEFF. 1991. Evidence for a northern northwest coast language area: promiscuous number marking and periphrastic possessive constructions in Haida, Eyak and Aleut. *IJAL* 57 (2): 158-193.

NICHOLS, JOHANNA. 1992. *Linguistic Diversity in Time and Space.* Chicago & London: University of Chicago Press.



Map of the languages discussed in this volume

Contributors

BRETT BAKER
Department of Linguistics, F12
University of Sydney
NSW 2006
Australia
brettb@sultry.arts.usyd.edu.au

WILFRIED BOEDER
Universität Oldenburg
Fachbereich 11: Anglistik-Sprachwissenschaft
26111 Oldenburg
Germany
wboeder@hrz1.uni-oldenburg.de

WERNER DROSSARD
Institut für Sprachwissenschaft
Universität zu Köln
Albertus-Magnus-Platz
50923 Köln
Germany
W.Drossard@uni-koeln.de

NICHOLAS EVANS
University of Melbourne
Department of Linguistics & Applied Linguistics
Parkville 3052
Australia
n.evans@linguistics.unimelb.edu.au

MICHAEL FORTESCUE
University of Copenhagen
Department of General and Applied Linguistics
Faculty of Arts
Njälägade 80
2300 København S
Denmark
mf@cphling.dk

DAGMAR JUNG
Institut für Sprachwissenschaft
Universität zu Köln
Albertus-Magnus-Platz
50923 Köln
Germany
djung@uni-koeln.de

MICHEL LAUNEY
IRD - Centre de Cayenne
BP 165 Cayenne
97323 French Guaiana
launey@cayenne.ird.fr

or

CNRS-CELIA
44, rue del'Amiral Mouchez
75014 Paris
France

JOHANNA MATTISSEN
Institut für Sprachwissenschaft
Universität zu Köln
Albertus-Magnus-Platz
50923 Köln
Germany
johanna.mattissen@uni-koeln.de

RACHEL NORDLINGER
University of Melbourne
Department of Linguistics & Applied Linguistics
Parkville 3052
Australia
r.nordlinger@linguistics.unimelb.edu.au

ADAM SAULWICK
Maningrida Arts & Culture
Via Darwin
PMB 102 Winnellie NT 0822
Australia
macsaa@maningrida.bu.aust.com

HANS-JÜRGEN SASSE
Institut für Sprachwissenschaft
Universität zu Köln
Albertus-Magnus-Platz
50923 Köln
Germany
Hj.Sasse@uni-koeln.de