

**PROCEEDINGS OF THE THIRTY-SECOND ANNUAL MEETING OF THE  
BERKELEY LINGUISTICS SOCIETY**

February 10-12, 2006

**SPECIAL SESSION**

**on**

**THE LANGUAGES AND LINGUISTICS OF OCEANIA**

**Edited by**

Zhenya Antić  
Charles B. Chang  
Clare S. Sandy  
Maziar Toosarvandani

**Berkeley Linguistics Society  
Berkeley, CA, USA**

Berkeley Linguistics Society  
University of California, Berkeley  
Department of Linguistics  
1203 Dwinelle Hall  
Berkeley, CA 94720-2650  
USA

All papers copyright © 2012 by the Berkeley Linguistics Society, Inc.

All rights reserved.

ISSN 0363-2946

LCCN 76-640143

Printed by Sheridan Books  
100 N. Staebler Road  
Ann Arbor, MI 48103

## TABLE OF CONTENTS

A note regarding the contents of this volume .....	iii
Foreword .....	iv

### SPECIAL SESSION

Oceania, the Pacific Rim, and the Theory of Linguistic Areas .....	3
BALTHASAR BICKEL and JOHANNA NICHOLS	
Australian Complex Predicates .....	17
CLAIRE BOWERN	
Composite Tone in Mian Noun-Noun Compounds .....	35
SEBASTIAN FEDDEN	
Reconciling <i>meng-</i> and NP Movement in Indonesian .....	47
CATHERINE R. FORTIN	
The Role of Animacy in Teiwa and Abui (Papuan) .....	59
MARIAN KLAMER AND FRANTIŠEK KRATOCHVÍL	
A Feature Geometry of the Tongan Possessive Paradigm .....	71
CATHERINE MACDONALD	
On the Argument Structure of Complex Predicates in Kalam, a Language of the Trans New Guinea Family .....	83
ANDREW PAWLEY	
Apical Obstruents in Pre-Proto-Tangkic and the Origins of the Non-Zero Absolutive .....	109
ERICH R. ROUND	

### A note regarding the contents of this volume

The following authors also presented papers as part of the Special Session of the conference, though their work does not appear in this volume: Frances Ajo; Keira Gebbie Ballantyne; Ian Maddieson; and Norvin Richards.

## **Foreword**

We are pleased to present the proceedings of the BLS 32 Special Session, held at UC Berkeley in February 2006. We would like to thank the contributors to this volume and all those who attended and participated in the conference. A special thanks to Chundra Cathcart of the 2011 editing team for helping to finalize this volume.

Clare S. Sandy

**SPECIAL SESSION:  
THE LANGUAGES AND LINGUISTICS OF  
OCEANIA**



# Oceania, the Pacific Rim, and the Theory of Linguistic Areas

BALTHASAR BICKEL and JOHANNA NICHOLS  
*University of Leipzig\* and University of California, Berkeley*

## 1. Introduction

A linguistic area is "a geographical region in which neighboring languages belonging to different language families show a significant set of structural properties in common, where the commonalities in structure are due to historical contact between speakers of the languages, and where the shared structural properties are not found in languages immediately outside the area (ideally where these include languages belonging to the same families as those spoken inside the area)" (Enfield 2005:190). That is, a linguistic area is defined by a group of *variables* (henceforth we use this term rather than *features*, *properties*, etc.) each of which constitutes an isogloss demarcating the area. Some linguists seek variables that form an isogloss bundle (e.g. Campbell et al. 1986, Joseph 1983, 2001); others do not (e.g. Emeneau 1956, Masica 1976), but nonetheless implicitly assume that some core part of the area should ideally emerge as located inside of all the isoglosses. Some works seek isopleths rather than isoglosses (van der Auwera 1998) and rank languages for the number of areal features they share. All of these approaches assume what we will call *categoriality* in the distribution of the defining variables: some value of a variable is present inside the area and absent outside of it (that is, in the neighboring languages outside of it).

Variable-defined areas present various problems. First, there are no criteria for deciding which are the diagnostic variables. This problem has an empirical side: the linguist needs to determine which variables are more and less frequent worldwide, which ones are most and least likely to diffuse, to be inherited; etc. It also has a statistical side. Suppose the linguist sorts through 200 variables and finds five that appear to be area-defining. Is this a significant result, or could one expect to find five out of 200 shared variables for any random set of languages and any random set of variables? The isogloss-bundled areal features standardly accepted for the Balkan and Mesoamerican language areas are selected from the entirety of the sound system, inventory of morphological forms, and basic syntactic inventory, a total set of elements that must number at least 200 and appears to be open-ended in practice. Half a dozen out of 200, or even 100, surveyed variables could easily cooccur in some set of languages by chance if they were at all frequent;

---

\* Note: 2006 affiliation. Affiliation as of publication: University of Zurich.

only if they were quite rare would it be unexpected for the set of languages to all show the entire half dozen variables. Our impression is that the classic Balkan features (to be listed below) include a few variables of sufficiently low frequency to be of diagnostic value, while the Mesoamerican ones include some that occur in one-quarter or more of the world's languages (head-marked nominal possession, non-verb-final basic word order), and one could expect five such to turn up in a survey of 200 or even 100 languages.<sup>1</sup> This issue has not had the discussion it deserves in the areal literature.

Second, a language may be a recent immigrant to an area and its speakers wholly involved in areal behavior such as bilingualism and code switching, yet the areal variables have not yet affected that language; does the linguist then draw a discontinuous isogloss quarantining the new language, disregard that language, or lower the standards for density of attestation of the criterial variables in the area? An example is Turkish spoken in Bulgaria, a core part of the Balkan linguistic area, by speakers bilingual in Bulgarian and/or Romani, both core Balkan languages. Balkanists have traditionally emphasized categorical variables found in all and only Balkan languages, with continuous isoglosses defining a coherent geographical area, and Bulgarian Turkish presents obstacles to the approach.

Third, the variables that can be identified as defining an area may be a motley set that raises few fruitful typological questions and does not fully capture the linguistic spirit of the area. An example of this is the classic Balkanisms (Joseph 1983:1, 2001:21): (i) postposed definite article, (ii) variant preposed future tense marker derived from a verb of volition, (iii) clitic doubling for objects, (iv) noun case mergers (especially displacement of genitive by dative; in the extreme situation, complete or near-complete loss of noun cases); (v) mid central vowel, (vi) lack of infinitive (finite subordinate clauses where most European languages use infinitives). It is true that identifying categorical Balkanisms is difficult because, except for Turkish, the Balkan languages are all related (as Indo-European) and much of what they have in common is inherited and shared with non-Balkan sisters. That said, the fact remains that the classic Balkanisms do not do a very complete job of defining the shared grammar that makes for the notable intertranslatability of Balkan languages.

Fourth, variables exhibiting the requisite isoglossic behavior may have to be defined as an abstraction which is in itself unlikely to be able to diffuse: an example is non-verb-final word order, a Mesoamerican areal variable identified by Campbell, Kaufman, and Smith-Stark (1986).

All in all, the variable-defined approach is unlikely to be able to define large, old, or inactive areas or areas with significant linguistic immigration very satisfactorily. This is because such areas are most likely to have diffuse boundaries, to

---

<sup>1</sup> A full statistical assessment will need to look at the worldwide frequency of the variable, the number of languages in the area, and the number of languages outside of but adjacent to the area (an area-defining feature cannot occur in any of these neighbors, though it can occur elsewhere in the world), and determine the probability of finding, say, five such variables given up to 200 attempts (or, perhaps more accurately, an open-ended number of attempts).



have internal nonconformities, to be typologically embedded in larger units, and to have confounding local divergence from areal norms.

Our approach turns the usual procedure on its head and defines variables from areas rather than vice versa. We define an area based on a theory of population and language spread and on information from other disciplines; hypothesize that it is a linguistic area; and test the hypothesis by seeking statistically non-accidental signals. We call this approach Predictive Areality Theory (PAT).

## **2. Predictive Areality Theory**

Each typological variable has its own history of and potential for change and spread, and therefore has its own distinct distribution over the world's languages. What underlies the impression of areality is that some such distributions overlap in a non-accidental way. If they overlap non-accidentally, one plausible explanation is shared history, by which we mean (any kind of) contact-induced change and/or shared inheritance (whether reconstructed and known or unreconstructible and unknowable). Such an explanation is a PAT holding for the specific regional overlap of the observed distributions. For a PAT to work, it must be grounded in what we know about population history from archaeology, genetics, ecology, geography, economics, demography, etc. Under this approach, then, areality is not a property of languages (e.g. 'in the Balkan Area' vs. 'not in the Balkan Area') but only a property of variables and sets of variables. In other words, areality is not, as under classical approaches, a typological observation. On the contrary, it is a theoretical predictor variable predicting observable typological distributions. The more the theory's predictions are statistically supported in such a series of predicted variables, the more robust the theory is.

Regional overlap can be explained by a PAT only if we can demonstrate that the overlap does not result from (a) universal preferences (e.g. VP ~ PP order, or noun incorporation and head marking), (b) reconstructible shared genealogy, or (c) chance. We can use regular statistical inferencing to determine the probability of (c), but we need to control for (a) and (b). We control for (a) and (b) using standard typological methods: for (a) by rejecting typological variables as independent areal signals if they are known to be associated universally; for (b), i.e. for known genealogical relatedness, by constructing genealogically-balanced samples instead of random samples. The consequence of this sampling decision is that we cannot apply standard sampling theory and need to rely on randomization-based statistical methods. (See (Janssen et al. 2006) for further discussion.)

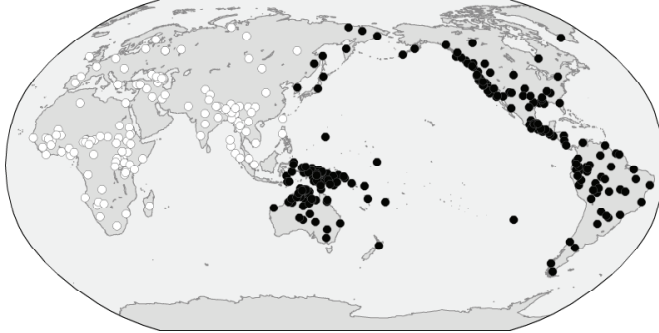
## **3. The Pacific Rim as a Linguistic Area**

In the 15 years since the first maps of numeral classifiers, head marking, and *n - m* personal pronouns were displayed to show a striking coast-hugging distribution all around the Pacific Rim (PR), a number of additional otherwise infrequent variables have been shown to have notably high concentrations in the Pacific-facing parts of the world. Yet the distributions of the variables that mark this putative area are manifestly not categorical or congruent. The area spans several

continents and lacks the compactness and centeredness of well-known smaller areas. Therefore, instead of attempting to trace area-defining isoglosses, we first define the area geographically and then ask whether any variables are significantly more (or less) frequent in the area than outside of it, and whether there are enough such to legitimately define an area. The rationale for grouping the entire Pacific Rim together as a single area includes human genetic and archaeological data indicating that the entire region was initially settled by migrations from ancient mainland Southeast Asia, continued to receive new colonizations from there up to and including the Austronesian expansion, and functioned as a contact and migration zone the whole time (Nichols 1997a, b, 2000, 2002).

We define the PR area as follows: Pacific-facing coast up to the lower slope of the far side of the major coast range (e.g. Andes, Sierras and Cascades, eastern Himalayas) or up to a coastal scarp (as in northern Australia). The Pacific Rim area is the more strictly coastal part of a larger area which we call the Circum-Pacific (CP) area. This comprises all of the Americas, Oceania (including Australia and New Guinea), and the mainland Asian Pacific Rim as just defined. That is, the CP area is the entire region anciently settled from coastal Southeast Asia and including the coastal Asian migration route. However, we exclude Southeast Asia (which we define as mainland Southeast Asia plus island Southeast Asia up to the Wallace line, i.e. including western Indonesia and the Philippines) from the CP area because it has considerably stronger historical and prehistorical ties to mainland Asia (Matisoff 1991, Enfield 2005) than to the other regions around the Pacific. We therefore expect Southeast Asia to pattern more often with Eurasia than with the CP. Drawing the boundary at the Wallace Line may appear arbitrary, but this is a natural breakpoint in our samples. Map 1 shows the definition of the CP area on a genealogically-balanced sample of languages.<sup>2</sup>

*Map 1. Definition of the Circum-Pacific area (black dots) in our sample*



There are five issues about this area (and similarly large areas) that now arise:  
(a) Variance. Languages with PR or CP features everywhere coexist with languages lacking them. Classical definitions of areality (Masica 1976, Campbell

---

<sup>2</sup> The underlying table with genealogy and geography coding is available on our project website: <http://www.spw.uzh.ch/autotyp/>. All other codings discussed below are also deposited there.

et al. 1986, Joseph 2001; survey: Enfield 2005:190) assume near-100% consistency in variables across an area, but in reality within-area variance in otherwise good areal features is common. A clear example of such a variable is multiple possessive classes (more than one "inalienable" class of nouns; Nichols and Bickel 2005, POSSCL in the Appendix below). In fact, in the PR and CP areas, variance is expected and likely to have been an ancient and stable characteristic because the territory is almost entirely residual zone in the terms of Nichols 1992, and because the expansion of languages bearing PR features involved movement into already inhabited lands so that languages with PR features did not displace others but intermingled with them. Given this, we maintain that our areality prediction is confirmed by any statistically significant difference in frequencies inside vs. outside the area – regardless of variance inside the area.<sup>3</sup>

(b) Leakage. In certain places, PR variables "escape" into the nearby (and not-so-nearby) interior: syntactic noun incorporation (Houser and Toosarvandani 2006) in North America; ergativity [COMALN5], inclusive/exclusive pronouns (ExInDist, Bickel and Nichols 2005b) and reduplicated plurals in Australia; many variables in South America (where "PR" is a misnomer as there is almost never a discernible coastal cluster of PR variables). Under a PAT approach, this is expected because it has clear historical motivations. Wherever a spread zone abuts the PR zone (North America, Australia, inner Eurasia), "escaped" features are likely to spread far. Thus, for example, the spread of domestication from Mesoamerica impelled PR features eastward via the Caribbean coast. In our statistical survey below, we use the larger CP area as a predictor in order to capture at least the leakages on the American side.

(c) Greater variance and general diffuseness of PR variables in Oceania. A number of PR variables form notably denser clusters in the Americas than in Oceania, raising questions about the unity of the area and its specific history. Examples include high inflectional synthesis of the verb (Bickel and Nichols 2005a, SYN) and *n-m* personal pronouns (Nichols and Peterson 1996, 2005; NICNMP2). Rather than a problem, under a PAT approach this is again an expected phenomenon: Oceania has been inhabited longer than the Americas and domestication occurred earlier there than in the Americas (Denham et al. 2003), so the land was already linguistically and demographically saturated when the PR expansion began. In saturated conditions, new linguistic features had less impact and took root less readily.

(d) A troubling historical question: How could PR variables persist so long in an area when there are many cases of their loss within historically reconstructed language families that are younger than the PR? Rather than a shortcoming we see this as a defining property of diagnostic areal features: they are more persistent in areas than in families. This must be because their retention can be favored by

---

<sup>3</sup> Still, it might be useful to distinguish these general kinds of areal signals from signals that show strong within-area homogeneity (as measured for example by chi-square deviations from expected distributions within the area).

areal pressure, and because in linguistic areas they are prone to be transmitted not only by inheritance but also by substratal retention and diffusion.

#### 4. Survey

We tested our predictions about CP areality against the dataset available in the World Atlas of Linguistic Structures (WALS; Haspelmath et al. 2005), amended by our own richer datasets for the variables that we contributed ourselves to the Atlas. The WALS dataset is not (and is not meant to be) a genealogically balanced sample. Therefore we constructed an all-purpose sample for WALS, called ‘WALSG’, with one representative per genus (as that is defined in the Atlas). When there was a choice we opted for the language that is coded for the largest number of coded variables. For our own chapters, we used our standard genealogical sample in AUTOTYP, called ‘GEN’. WALSG contains 193 languages, GEN 316. Using GIS software, we coded each language in both samples as belonging or not to the CP area. We used the larger CP area rather than its PR subpart because of the issue of leakage discussed above.

On an all-purpose sample, variables end up with many missing values. Of all variables available in WALS (or our versions of them) we selected those that have at least 150 (i.e. about 75%) non-empty values. This yields 75 variables.

The values of a typological variable can generally be lumped or split in various ways. For example, the variable of case alignment in (Comrie 2005) distinguishes marked from unmarked nominative/accusative alignment, while for different purposes one could treat them as the same and put them in opposition to several other alignments. In technical terms, these are all different ontologies derived from a single variable. In universals research we generally know which ontology is of interest to the prediction (e.g. accusative vs. other non-neutral alignment for predictions about which alignment type is preferred in agreement as opposed to case systems), but in areal typology we cannot know a priori which ontology will show areal overlap in its distribution. Re-ontologizing, or recoding, is of course only possible for multinomial variables and not all possible recodes are linguistically meaningful. With these constraints in mind, we recoded 23 of the 75 variables, with the number of recoded variants of each variable ranging from 2 to 6 (mode = 2). This yielded a total set of 100 variables. Note that some recodes increase again the number of missing values, but now these are logically necessary and not sampling gaps: for example, a binary recode of subtypes of accusative marking will have missing values only in languages that do not have accusative case alignment, but this is a fact of life and not a sampling problem.

We then tested our areality prediction against the 100 variables. That is, we surveyed not a hand-picked number of variables and not an open-ended set, but all variables available in testably high frequencies in both databases under genealogical sampling. For each variable, we tested whether there was a statistically significant difference between its frequencies in the Circum-Pacific and the rest of the world (i.e. Africa and non-Pacific Eurasia). For binary typological variables we used a 2x2 (typological variable x CP) Fisher Exact Test; for multinomial and

scalar variables we ran randomization-based chi-square and one-way anova tests, respectively, as described in (Janssen et al. 2006). We report the results in the Appendix, ranked by p-values.

## **5. Results**

When interpreting the results, we need to control for the fact that some variables might be universally correlated. We have not tested all possible universal correlations among the 100 variables, but the following word-order variables are well-known to correlate: DRYOBV0 ~ DRYGEN0 ~ DRYSOV0 ~ DRYSBV0 ~ DRYADP0 ~ DRYCOQ0 ~ DRYPQP01 ~ DRYPQP02 ~ VFIN ~ VFIN2 ~ VINIT ~ VINIT2; CORSEX01 ~ SIEGEN2 and SIEAPV2 ~ SIEVPA02 ~ POLYAGR are respectively the same or very similar variables coded by different researchers (see Appendix for what these labels stand for). What other correlations exist is an open question, one that needs extensive analysis. For now, we assume that 86 of the variables tested are distributionally independent of each other.

Running the same test on various recodings of the same variable increases the risk of familywise error of rejecting true null hypotheses. We controlled for this by applying Holm corrections to the  $p$ -values of each set of mutual recodings of a single variable (e.g., we corrected the  $p$ -values of all our 6 recodings of DRYSOV, Dryer's (2005) S-O-V order variable).

At a conventional .05 rejection level, we find that about 40% of the 86 variables that we assume to be independent show significant frequency difference between the CP area and the rest of the world. About 30% do so at a .01 level.<sup>4</sup>

## **6. Conclusion**

This has been an exercise in applying Predictive Areal Theory to a deep, old, and very large area which a priori presents many problems for areal analysis. We defined the PR and CP areas geographically, basing the definition and the geographical extent on what is known about human migrations and the settlement of the Pacific and the New World, then assembled a list of all variables which had enough data in an general-purpose database (WALS) and tested whether frequencies of variables in the area are significantly different from those outside the area. The outcome was that (depending on one's significance criterion) 30-40% of the variables yielded significance, and we regard each of these as a likely areal feature. This success rate is high enough to convince us that we have detected multiple symptoms of genuine areality. Note that the datasets were controlled for genealogical bias by an all-purpose sample, and this often meant that the actual dataset had to be shrunk, reducing the power of the statistical tests. It is possible that a sampling procedure that leads to larger samples would reveal more significant associations.

Our understanding is that the PR formed as coastally adapted people, and their

---

<sup>4</sup> Space limits make it impossible to include maps of the variables, but a sense of their actual distribution can be gained from the maps in WALS.

languages and cultures, spread out of Southeast Asia beginning late in the last glaciation and continuing into recent centuries with the Austronesian spread and the Chukchi spread to the Bering Strait. They spread coastally, as is shown by the striking coastal distributions of variables such as V-S order and multiple possessive classes. We tested for CP rather than more strictly for PR areality because leakage is such a pervasive problem as to obscure the linguistic boundary between the two (though not the geographical boundary, which we defined in advance).

All theories of areality take account of cultural, historical, and ecological factors as well as linguistic structure, but PAT differs in its crucial respects – defining areas geographically, no assumption of categoriality in variable distributions, testing all available variables for areality – because it was developed for work on large, old areas for which categoriality and neat isoglosses cannot be expected. Much work remains to be done, including development of statistical tools to define the minimum success rate that can be judged non-chance and to disentangle the PR from the CP. Even without these tools, however, the CP area has emerged as a clear linguistic area established by many independent variables.<sup>5</sup>

## References

- Bickel, Balthasar, and Nichols, Johanna. 2005a. Inflectional Synthesis of the Verb. In M. Haspelmath, M. Dryer, B. Comrie & D. Gil, eds., *The World Atlas of Language Structures*, 94-97.. Oxford: Oxford University Press.
- Bickel, Balthasar, and Nichols, Johanna. 2005b. Inclusive/exclusive as Person vs. Number Categories Worldwide. In Elena Filimonova, ed., *Clusivity*, 47-70. Amsterdam: Benjamins.
- Bickel, Balthasar, and Nichols, Johanna. 2002ff. *The Autotyp Research Program*. <http://www.uni-leipzig.de/~autotyp/>
- Campbell, L., T. Kaufman, and T.C. Smith-Stark. 1986. Mesoamerica as a Linguistic Area. *Language* 62:530-570.
- Denham, T. P. et al. 2003. Origins of Agriculture at Kuk Swamp in the Highlands of New Guinea. *Science* 301:189-193.
- Dryer, M. S. 2005. Order of Subject, Object, and Verb. In M. Haspelmath, M. Dryer, B. Comrie & D. Gil, eds., *The World Atlas of Language Structures*, 330-34.
- Emeneau, Murray B. 1956. India as a Linguistic Area. *Language* 32:3-16.
- Enfield, Nicholas J. 2005. Areal Linguistics and Mainland Southeast Asia. *Annual Review of Anthropology* 34:181-206.
- Haspelmath, M., M. Dryer, B. Comrie, and D. Gil, eds. 2005. *World Atlas of Language Structures*. Oxford: Oxford University Press.
- Houser, Michael, and Maziar Toosarvandani. 2006. A Non-Syntactic Template for Syntactic Noun Incorporation. LSA Annual Meeting, Albuquerque.

---

<sup>5</sup> We thank Sven Siegmund and Anja Gampe for their help with the recoding of the WALs data.

*Oceania, the Pacific Rim, and Linguistic Areas*

- Janssen, Dirk P., Balthasar Bickel, and Fernando Zúñiga. 2006. Randomization tests in language typology. *Linguistic Typology* 10. 419-440.
- Joseph, Brian. 2001. Is Balkan Comparative Syntax Possible? In M. L. Rivero and A. Ralli, eds., *Comparative Syntax of Balkan Languages*. Oxford: Oxford University Press.
- Joseph, Brian. 1983. *The Synchrony and Diachrony of the Balkan Infinitive*. Cambridge: Cambridge University Press.
- Masica, Colin P. 1976. *Defining a Linguistic Area: South Asia*. Chicago: University of Chicago Press.
- Matisoff, James A. 1991. Sino-Tibetan Linguistics: Present State and Future Prospects. *Annual Review of Anthropology*, 469-504.
- Nichols, Johanna. 2003. Genetic and Typological Diversification of Language. In B. Joseph and R. Janda, eds., *Handbook of Historical Linguistics*, 283-310. London: Blackwell.
- Nichols, Johanna. 2002. The First American Languages. *Memoirs of the California Academy of Sciences* 27:273-293.
- Nichols, Johanna. 2000. Estimating Dates of Early American Colonization Events. In C. Renfrew, A. McMahon and L. Trask, eds., *Time Depth in Historical Linguistics, volume 2*, 643-663. Cambridge: McDonald Institute for Archaeological Research.
- Nichols, Johanna. 1997a. Sprung from Two Common Sources: Sahul as a Linguistic Area. In P. McConvell and N. Evans, eds., *Archaeology and Linguistics: Aboriginal Australia in Global Perspective*, 135-168. Melbourne: Oxford University Press.
- Nichols, Johanna. 1997b. Modeling Ancient Population Structures and Movement in Linguistics. *Annual Review of Anthropology* 26:359-384.
- Nichols, Johanna. 1992. *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press.
- Nichols, Johanna. 1994. Ergativity and Linguistic Geography. *Australian Journal of Linguistics* 13:39-89.
- Nichols, Johanna, and Bickel, Balthasar. 2005. Possessive Classification. In M. Haspelmath, M. Dryer, B. Comrie & D. Gil, eds., *The World Atlas of Language Structures*, 242-245.
- Nichols, Johanna, and Peterson, David A. 1996. The Amerind Personal Pronouns. *Language* 72:336-371.
- Nichols, Johanna, and Peterson, David A. 2005. Personal Pronouns: M-T and N-M Patterns. In M. Haspelmath, M. Dryer, B. Comrie & D. Gil, eds., *The World Atlas of Language Structures*, 544-551.
- van der Auwera, Johan. 1998. Revisiting the Balkan and Mesoamerican Linguistic Areas. *Language Sciences* 20:259-27.

Balthasar Bickel and Johanna Nichols

Balthasar Bickel  
Seminar für Allgemeine Sprachwissenschaft  
Universität Zürich  
Plattenstrasse 54  
CH-8032 Zürich

[balthasar.bickel@uzh.ch](mailto:balthasar.bickel@uzh.ch)

Johanna Nichols  
Department of Slavic Languages and Literatures, #2979  
University of California  
Berkeley, CA 94720-2979

[johanna@berkeley.edu](mailto:johanna@berkeley.edu)



Oceania, the Pacific Rim, and Linguistic Areas

Appendix: evidence for the CP area, ranked by corrected p-values

Variable	Values	Rough explanation (see WALS chapters for details)	Recode	WALS Chapter	Sample	N	CP evidence: corrected <i>p</i>	uncorrected
MADGAP	5	Missing common C		5	WALS.G	175	9.97E-15	
MADVOI2	2	Voicing	1-2/3/4	4	WALS.G	175	1.01E-14	5.07E-15
AUWEP12	2	Epistemic modality verbal vs. affixal	1-2/3	75	WALS.G	150	4.34E-11	
POL.YAGR	2	Obligatory agreement with both A and P		22	GEN	276	4.08E-09	
POSSCL	2	Inflectional possessive classes		59	GEN	238	4.87E-08	
MADVOW	3	Size of vowel inventory		2	WALS.G	175	2.96E-06	
DRYPOS0	3	Possessive prefix vs. suffix vs. both	1-2-3	57	WALS.G	94	3.00E-06	1.50E-06
MADLAT2	2	Laterals		8	WALS.G	175	8.22E-06	4.11E-06
SIEAPV2	2	Agreement with both A and P	1-2/3/4/5	104	WALS.G	180	3.68E-05	1.84E-05
COMNUM5	5	Counting systems	1-2/3-4-5-6	131	WALS.G	122	5.07E-05	
MADVOI0	2	Voicing in plosives vs. fricatives vs. both	2-3-4	4	WALS.G	108	7.56E-05	7.56E-05
MADDCON	5	Number of consonants		1	WALS.G	173	1.00E-04	
SYN	scale	Inflectional synthesis degree (w/o roles)		22	GEN	202	1.00E-04	
BAKADP2	2	Adpositions	1-2/3/4	48	WALS.G	179	1.06E-04	3.54E-05
SIEPAS	2	Passive		107	WALS.G	178	2.78E-04	
NICMTP2	2	m/-pronouns	1-2/3	136	GEN	185	4.68E-04	
DRYGEN0	2	GenN order	1-2	86	WALS.G	163	8.56E-04	
MIEASY	7	Asymmetry types in NEG		114	WALS.G	170	9.69E-04	
MADLAT0	4	Lateral series	2-3-4-5	8	WALS.G	141	1.05E-03	1.05E-03
COMALN5	5	Case alignment of nouns (ACC collapsed)	1-2/3-4-5-6	98	WALS.G	164	1.88E-03	
SIEALI0	5	Alignment in agreement	1-2-3-4-5-6	100	WALS.G	140	2.07E-03	
ExIndist	2	Incl/Excl-Distinction		*	GEN	289	2.14E-03	
HAAEVD2	2	Evidentials	1-2/3	78	WALS.G	170	2.92E-03	1.62E-03
SIEZER2	2	S agreement	1-2/3/4/5/6	103	WALS.G	180	3.24E-03	3.42E-03
SIEVPA01	2	Agreement	0-2/3/4/5	102	WALS.G	180	3.86E-03	1.93E-03
SIEVPA02	5	A and/or P agreement	1-2-3-4-5	102	WALS.G	140	3.86E-03	
CORASS01	2	Semantic vs. semantic and formal gender	2-3	32	WALS.G	53	4.87E-03	
MADTON02	2	Tone	1-2/3	13	WALS.G	169	5.72E-03	2.86E-03
DRYPRE0	3	Affix position trend	2/3-4-5-6	26	WALS.G	145	6.49E-03	
DRYSOV0	6	S,V,O orders	1-2-3-4-5-6	81	WALS.G	145	7.50E-03	1.25E-03
HAAEVC0	5	Evidential marking types	2-3-4-5-6	78	WALS.G	78	7.73E-03	
DRYPOS2	2	Possessive affixes	1/2-3	57	WALS.G	151	9.97E-03	9.97E-03

Balthasar Bickel and Johanna Nichols

ExaPers	2	incl/excl as person	*	GEN	289	0.01025	
MADFRV2	2	Front rounded V	11	WALSG	174	0.01100	5.50E-03
VINIT2	2	V-initial or free order	81	WALSG	175	0.01320	2.64E-03
DRYSBV0	2	SV vs VS order	82	WALSG	172	0.01374	
DRYRC00	5	Type of pronominal subject expression	101	WALSG	152	0.01406	
LoePOSSU2	2	Double-Marking possessor and object	23	GEN	248	0.01790	
DRYADP0	2	Adposition: post vs. pre vs. in	1-2-3	WALSG	164	0.02322	
SEJZER0	2	Nonzero vs. zero in 3sAGR	85	WALSG	164	0.02322	
ANDANG2	2	Velar nasal present	103	WALSG	135	0.02349	0.02349
IGNUM0	2	Case	9	WALSG	168	0.02527	0.01263
WOFREE	2	Free word order	49	WALSG	172	0.02657	0.01328
DOBOP1	2	Inflectional Optatives	81	WALSG	175	0.03316	0.00829
BAECXY01	3	Case syncretism degree	73	WALSG	157	0.03577	
DRYOBV0	2	OV vs VO	2-3-4	WALSG	64	0.03918	0.01959
CORSEX01	2	Gender	1-2	WALSG	169	0.04322	
AUWHOR	4	Type of hortative system	1-2/3	WALSG	147	0.04749	0.01583
SIEGEN2	2	Gender	72	WALSG	153	0.05141	
CORSEX	3	no gender vs. sex-biased vs. other	44	WALSG	178	0.06509	0.03255
VFIN2	2	V-final or free order	31	WALSG	147	0.07560	0.03780
CORNUM	2	Number of genders	81	WALSG	175	0.07932	0.02644
DRYDEM	6	Demonstrative initial vs. final	30	WALSG	143	0.08000	
DRYDEM0	2	Number of genders	88	WALSG	174	0.09089	0.04545
NICNMP2	2	n/n-pronouns	1/2-2/4	WALSG	163	0.09089	0.07309
DRYCOQ0	2	W/H initial	1-2/3	GEN	185	0.10616	
MAuDist	2	Minimal/augmented system	93	WALSG	140	0.11897	
SIEGEN0	5	Gender across person and number categ.	*	GEN	289	0.13819	
IGNUM	scale	Number of cases	44	WALSG	55	0.14064	0.14064
DRYNPL	9	Coding type of plural	49	WALSG	172	0.14340	0.14340
COMALP0	5	Pronoun alignment (ACC collapsed)	33	WALSG	164	0.16180	
BAKADP02	3	Adposition agreement	99	WALSG	147	0.16335	
MADPRS0	6	Presence of uncommon consonants	48	WALSG	152	0.16556	0.08278
PREROLE	2	Some agreement prefixed	19	WALSG	28	0.23562	
MADTON01	2	Simple vs. complex tone	22	GEN	160	0.24503	
DANPLU04	3	Types of expressing plural an pronouns	13	WALSG	56	0.25715	0.25715
BAEPSY01	2	Subject agreement syncretism	35	WALSG	166	0.28073	0.05615
AUWIMP2	2	Morphological imperative	29	WALSG	124	0.28713	0.14356
LoeU2	2	Double-Marking object	70	WALSG	170	0.37041	
			25	GEN	245	0.40627	

Oceania, the Pacific Rim, and Linguistic Areas

SONNON2	2	Nonperiphrastic causatives	1-2/3/4	111	WALS:SG	158	0.40891	
DRYNUM0	2	NumN vs. NNum	1-2	89	WALS:G	161	0.43265	
MADSYL	3	Complexity of syllables		12	WALS:G	167	0.45860	
DRYPQP01	2	Position of Q-particle	1-2-3-4-5	92	WALS:G	88	0.46069	0.30851
DRYPQP02	4	Position of Q-particle ('early' collapsed)	1/3-2-4-5	92	WALS:G	88	0.46069	0.23034
MADUVU0	3	Uvular C series	2-3-4	6	WALS:G	35	0.48796	0.24398
MADUVU2	2	Uvular C	1-2/3/4	6	WALS:G	175	0.48796	0.25946
BAKADP01	2	Agreement on adpositions	2-3/4	48	WALS:G	152	0.49592	0.49592
VINIT	2	V-initial	3/4-1/2/5/6/7	81	WALS:G	175	0.53143	0.26571
DRYTA02	2	Tense/aspect inflection	1/2/3/4-5	69	WALS:G	176	0.53531	
MADGLO0	2	Glottalized C	1-2/3/4-5	7	WALS:G	175	0.59847	
BAEPSY02	2	Subject agreement	1-2/3/4/5/6/7/8	29	WALS:G	171	0.60374	0.60374
LoePOSS2	2	Double-Marking possessor	1-2/3	24	GEN	244	0.61984	
DRYCAS	9	Morphological type of case		51	WALS:G	165	0.62349	
HANAS	2	Nasal vowels		10	WALS:G	149	0.69216	
CORSEX02	2	Sex-based vs. non-sex-based gender	2-3	31	WALS:G	53	0.74044	0.74044
SEAPV0	2	A before P vs. P before A in agreement	2-3	104	WALS:G	72	0.79550	0.79550
AUWPRH21	2	Dedicated prohibitive	1-2/3/4	71	WALS:G	154	0.85358	0.42679
AUWPRH22	2	Prohibitive as imperative	1/2-3/4	71	WALS:G	154	0.85358	0.51540
DRYADJ0	2	AdjN vs Nadj	1-2	87	WALS:G	161	0.87035	
VFIN	2	V-final order	1/6-2/3/4/5/7	81	WALS:G	175	0.87767	0.87767
MIESYM	3	Symmetric vs. asymm. vs. mixed negation		113	WALS:G	170	0.91513	
DANPLU01	7	Type of plural coding on subject pronouns	3-4-5-6-7-8	35	WALS:G	166	0.98075	0.24519
DRYPOQ2	2	Interrogative - declarative distinction	1/2/3/4/5/6-7	116	WALS:G	155	1.00000	
ANDANG0	2	Velar nasal banned from initial position	1-2	9	WALS:G	79	1.00000	1.00000
BAECASY02	2	Case syncretism	2/3-4	28	WALS:G	64	1.00000	1.00000
DANPLU02	2	Subject pronouns (present or not)	1-2/3/4/5/6/7/8	35	WALS:G	175	1.00000	0.50910
DANPLU03	2	Person-number vs. person stems	4/5/6-7/8	35	WALS:G	146	1.00000	0.66323
DANPLU05	3	Person and number coexistence in pron.	3/4-5/7-6/8	35	WALS:G	166	1.00000	0.90139
MADFRV0	3	Type of front rounded vowel	2-3-4	11	WALS:G	9	1.00000	
DRYNEG2	2	Single vs. double negation	1/2/3/4-5-6	112	WALS:G	166	1.00000	1.00000

**Explanations:**

'Recode': the definition of how values in the WALS database were recoded. The values are shown by the numerical labels they have in WALS and '/' means that values were collapsed whereas '-' means they were kept distinct. Values that were excluded are those that are not listed here.

'N': the number of languages with a non-missing value for the variable in the sample.

\* Bickel and Nichols (2005b), corresponding to WALS Chapters 39 and 40 by Michael Cysouw



# Australian Complex Predicates

CLAIRE BOWERN

*Yale University*

## 1. Introduction

The aim of this paper is to report on the investigation of the heterogeneity of complex predicates among certain Northern Australian languages. In particular, I describe preliminary results of research into complex predicates involving a preverb (or coverb) and an inflecting light verb.<sup>1</sup>

I follow Butt and Geuder (2001, 325) in considering light verb constructions as a type of complex predicate which consists of a main lexical verb in combination with a lexically defective verb (of course not all complex predicates are V V constructions, and not all V V constructions are complex predicates). I assume Butt's features of complex predicates (extracted from Butt and Geuder 2001, 323-327; see also Butt 1995, 2). The definition provided by Alsina *et al.* (1997, 1) is similar: each component of the complex predicate contributes to the predicate information normally associated with a head. Thus complex predicates are 'complex' because they consist of two (or more) constituents which do the work of a single verb; in other words, the functions of the predicate are spread across multiple constituents.

---

<sup>1</sup> My work on Bardi was funded by AIATSIS grants G2001/6505 and G2003/6761. Many thanks to †Nancy Isaac, Bessie Ejai and Jessie Sampi, who provided most of the Bardi data, for their friendship and patience in teaching me their language. The participants in Rice University's graduate syntax seminar (Spring 2006) and Beth Levin provided useful feedback. Some of the languages under discussion also show verb serialization but the scope of this paper is limited to light verb constructions, and so "complex predicate" here should be taken to refer to complex predicates of the preverb + light verb type.

Many constructions fall under this definition, including serial verb constructions, restructuring predicates, and various “light verb” constructions. The most usual construction in Northern Australian languages comprises an uninflecting (or largely uninflecting) preverb and one of a small number of inflecting verbs. All the languages in question are head marking and the “inflecting” verbs function as predicates in their own right.<sup>2</sup>

In §2 I give a more detailed overview of the light verb constructions. The aim of §3 and §4 is to quantify the extent of diversity among these constructions, in order to see whether the different languages show similar syntactic behavior in their light verb constructions (and if not, how they differ, and whether it is possible to trace the source of the difference).

## 2. Australian Complex Predicates

Light verb complex predicates are found in a belt of languages across the North of Australia, from the Dampier Peninsula in the West to Arnhem Land at the eastern end of the Northern Territory. This area includes multiple genetic families, and indeed at this stage there seems to be little correlation between genetic affiliation and the presence or absence of complex predicates. A map is given in Figure 1 below.

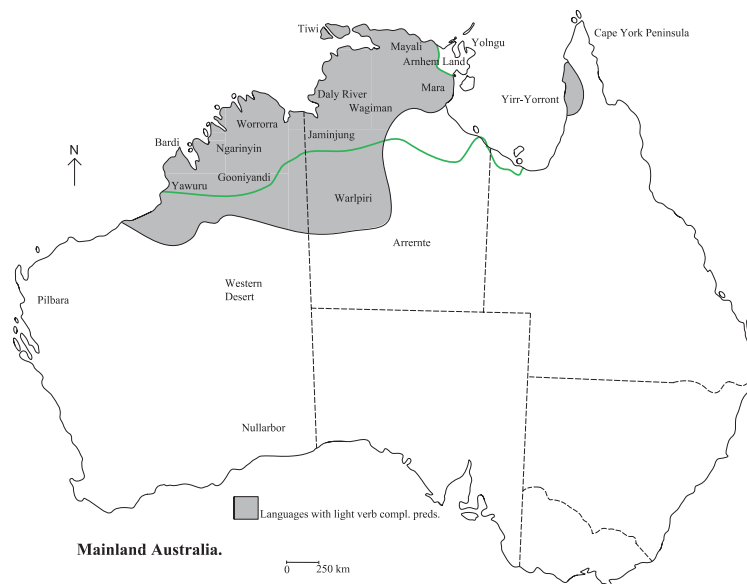


Figure 1: Map of Australia showing distribution of complex predicates

<sup>2</sup> There is an exception to this statement: in a few languages the light verbs and preverbs have unverbated and form a single phonological word. This has occurred, for example, in Gooniyandi (McGregor 1990) and Nunggubuyu (Heath 1984).

The type of complex predicate under consideration is described here with reference to Bardi, a Nyulnyulan (non-Pama-Nyungan) language spoken by about 25 people on Australia's north-west coast. Complex predicates in Bardi (and other Nyulnyulan languages) comprise an inflecting *preverb* which immediately precedes an inflecting verb root.<sup>3</sup> The inflecting verb hosts agreement and tense/aspect morphology, as seen in the previous section. In Bardi the preverb cannot take any verbal or other affixes, although many (but not all) preverbs can be reduplicated. The examples in (1) provide some illustration. In (1a), the preverb *garr* combines with the inflecting verb *-boo-* to form a complex predicate meaning 'rub'. The predicate is transitive. Example (1b) shows another complex predicate, this time intransitive. The preverb *roowil* combines with the inflecting verb *-inya-* to form a predicate meaning 'walk'. Note that although glosses have been given for *roowil* and *garr*, neither item exists independently of the complex predicate construction.<sup>4</sup>

- (1) a. *garr nganamboogal* 'I rubbed him.'  
       Preverb: *garr* 'rub'  
       Inflecting verb: *-boo-* 'hit'  
       Entire predicate: 'to rub (something) to stop the pain'.  
   b. *roowil innyagal* 'He was walking.'  
       Preverb: *roowil* 'walk'  
       Inflecting verb: *-nya-* 'pick up, catch'  
       Entire predicate: 'to walk'

Preverbs are an open word class. Loan verbs, for example, are borrowed as preverbs and assigned an inflecting verb based on the semantics of the action denoted by the preverb. The Kriol verb *boojoom* 'push' (from English 'push'im'<sup>5</sup>), for example, is borrowed as a preverb into Bardi and takes the inflecting verb *-ma-* 'put', along with many other verbs that imply an action involving 'transfer'. These preverbs show the same properties of classification that the non-borrowed Bardi lexicon does (see Bower (2004, ch. 9) for more information).

Superficially, most of the languages of Northern Australia would appear to have the same type of complex predicate construction. Compare (2) (from Bardi) with the example in (3), from Wagiman. The two languages have in common a restricted set of inflecting light verbs and open class of preverbs. They are also spoken in different parts of the country and are not demonstrably related.

<sup>3</sup> Constituent order is constant. The only material which may intervene between a preverb and an inflecting verb is a sentential clitic; these cliticize to the end of the first phonological word in the phrase.

<sup>4</sup> A transitive verb is here defined as one which may take two nominal arguments, one of which appears in ergative case. The other usually appears in the absolutive. Intransitive predicates take single absolutive argument.

<sup>5</sup> In Kriol (the English-based creole widely spoken in north-west Australia, third person singular object clitic pronouns have been reanalyzed as markers of transitive verbs. Bardi borrows Kriol transitive verbs with the transitive marker intact.

- (2) *Lalba innyagal* *ginyinggi bardag moonga-ngan*  
 split 3SG-TR-‘catch’-REC.PST this tree sugarbag-ALLATIVE  
*manyan gorna jamooyoon-oong.*  
 INFIN-take-CONT well axe-INSTR  
 “He split the log into pieces with an axe to get the bush honey easily.” (Bardi)
- (3) *Ngonggo-gin wayi-tjjalbu-yi weeh ngan-ya-ny // lari-leying.*  
 2SG.GEN small-ERG vomit.PFV 3SG/1SG-do-PPFV arm-LOC  
 “Your baby vomited on my arm.” (Wagiman)

There have been several previous descriptions of complex predicates in several Northern Australian languages, however the construction is seldom regarded as a complex predicate (or the inflecting verb equated with a light verb construction). They have been described as auxiliary + infinitive constructions, for example (Chadwick 1975), or as “verb + particle” constructions (Rumsey 1982, Merlan 1994). It is only fairly recently that the term “complex predicate” has entered the literature (cf. Schultze-Berndt 2000, Wilson 1999, Bower 2004).

Moreover, previous comparative approaches to Northern Australian complex predicates have tended to focus on three main areas. The first is the relative order of the preverb, the light verb, and any agreement marking. For example, Dixon (2002) devotes a number of pages to the different possibilities found in various Australian languages. The second area has been dubbed the ‘tightness of nexus’ (e.g. Schultze-Berndt 2000, 536ff); that is, how rigid the constituency of the preverb and light verb is. Finally, there have been several discussions of the semantics of the light verb (Schultze-Berndt 2000, McGregor 2002) and its contribution to the predicate as a whole.

In this paper I instead focus on a broader range of syntactic behavior, including the determination of argument structure in the predicate and the ability of preverbs to appear without accompanying light verbs.

### 3. Similarities

Let us consider some of the properties which all these languages share. Firstly, all the constructions under consideration have two components – a morphologically complex verb which can also stand alone as a verbal predicate, and another item, here termed the ‘preverb’. Examples were seen in (2) and (3) above.

The inflecting (“light”) verbs which can appear in these constructions are similar to those found elsewhere in the world. The number of light verbs in a given language varies, but the inventory typically includes verbs which translate as ‘do’, ‘put’, ‘sit’, ‘hit’, ‘go’ and ‘take’. The set of the most common Bardi and Wagiman light verbs are given in Table 1.

While the light verbs used in any given language are roughly predictable, there is a great deal of diversity in the ratio of light verbs to simple predicates. For example, the Western Nyulnyulan languages each have about 10 common light verbs,



Table 1: Bardi and Wagiman light verbs

Bardi	gloss	Wagiman	gloss
-boo-	hit	-bu-	hit
-ga-	carry	-ga-	take
-ma-	put	-ge-	put
		-ma-	get
-ni-	sit	-ni-	be
-ar-	spear	-ra-	throw
-joo- ~ -di-	say, do	-ya-	say/do/become
-jiidi-	go	-ya-	go

another 20 or so rare light verbs, and about 200 other simple predicates (Bower 2004, McGregor 2002). At the other end of the scale are languages such as Malak-Malak (Birk 1976) with six simple predicates, all of which also participate in light verb constructions.

In all these languages, preverbs form an open, heterogeneous word class (cf. Schultze-Berndt 2001). That is, they are a large class in dictionaries, they are frequently the vehicle for borrowing verbs into the language. See (4) for an example from Bardi.

- (4) *Wajim inamana.*  
 wash 'im 3SG-TR-'put'-REM.PST  
 "He/She washed it."

Preverbs appear to be a heterogeneous class in many languages. In Bardi, for example, they have diverse etymologies and may be recruited from other word classes. (5) gives some examples of Bardi preverbs which also belong to other word classes.

- (5) a. Nouns  
*giringg* 'a cough'; *giringg -ar-* 'to cough';  
*anggoorr* 'tears'; *anggoorr -ma-* 'to mourn for someone'
- b. Adjectives  
*ngaada* 'short'; *ngaada -joogooloo-* 'to break in half';  
*rambin* 'heavy'; *rambin -joo-* 'feel heavy'
- c. Adverbs  
*angan* 'closeby'; *angan -ganyi-* 'to come up close';  
*bard* 'away'; *bard -ga-* 'take across'
- d. Loans from other languages  
*boojoom* 'push 'im' (Kriol); *boojoom -ma-* 'to push off (a boat)';  
*warrgam* 'work 'im' (Kriol); *warrgam -joo-* 'to work';

Unfortunately we do not yet have detailed etymological data and reconstructions which would allow us to be more certain about the history of preverbs in most of the languages of Northern Australia.

#### 4. Differences

Let us now turn to a few of the areas in syntax where Northern Australian languages differ. For ease of presentation, and due to space restrictions, I present only a comparison between Bardi and Wagiman here, although other languages will also be mentioned where relevant. Note also that Bardi and Wagiman are not meant to present two ends of a spectrum, merely two possibilities amongst many. All data for Bardi are from my field notes or those of Gedda Aklif; data for Wagiman are from Wilson (1999).

##### 4.1. Bardi Preverb conjunction and stacking

Previous studies have described differences in Australian languages in terms of the ability of preverbs and their light verbs to be separated by intervening material (that is, in terms of the tightness of nexus between the preverb and the light verb). A related question is whether preverbs are a phrasal category in a given language, and whether they can be conjoined under a single light verb.

Australian languages differ in the extent to which they allow preverbs to be stacked with a single light verb, or conjoined. For example, in Bardi conjunction of preverbs is almost impossible, and always dispreferred even when grammatical. Each preverb normally combines with a single light verb. The example in (6) is very unusual; the preferred construction is given in (7).

- (6) *Bilirl agal giringg nganarij* *bardi.*  
yawn and cough 1-TR-spear-MID.PERF yesterday  
'I yawned and coughed all day yesterday.' (CB/FN: NI.3/47)
- (7) *Bilirl nganarij* *bardi agal giringgiring*  
yawn 1-TR-spear-MID.PERF yesterday and cough  
*nganarij.*  
1-TR-spear-MID.PERF  
'I yawned and coughed all day yesterday.'

There is one case, however, where multiple preverbs are allowed with a single light verb in Bardi. That is where the first preverb marks the trajectory of the action (the most common of these is *bard* "off (away from speaker)"). Two examples are given in (8).

- (8) a. *bard arr -joo-*  
off go do/say  
'go off'

- b. *bard roowil -nya-*  
 off walk catch  
 ‘walk off’

Let us now compare the data for Bardi to some parallel examples in Wagiman. In Wagiman it is much more common to have multiple preverbs. (9) is an example. (My understanding from Wilson (1999) is that such phrases are not common in his corpus, but they were produced spontaneously. This is in contrast to Bardi, where they never occur spontaneously apart from sentences such as (8), and the examples in (6) were only constructed by a speaker after a great deal of thought and some prompting.)

- (9) *Gurruwitj-yi nangh berrh la-ng.*  
 motorcar-ERG knock down.PFV throw.PFV 3SG.throw-PPFV  
 “A car knocked them down.”
- (10) *Gabarn-ma wek-ga ga-ra-n.*  
 quickly-ASP swallow-ASP 3SG-throw-PRES  
 “He swallows it quickly.” (Wilson 1999, 71)

Warlpiri is another language where preverb stacking is possible, although the possibilities are more limited, and depend on the semantic class of the preverb.<sup>6</sup>

- (11) *kanginy- purda- nya -nyi*  
 misperceiving- hear- perceive -CLS  
 ERG fail to hear ABS properly (Nash 1982, 175)

#### 4.2. Preverbs as heads of phrases

In the previous section we saw that the situations in Bardi which allowed more than one preverb per phrase were very limited. Furthermore, in Bardi it is almost impossible for a preverb to appear without an associated light verb. That is, preverbs in Bardi are not the sole heads of phrases, no matter what other head-like properties they exhibit. Two examples are given in (12) and (13). It should be noted that in the first, *jirrma* is probably not a preverb at all, but a noun homophonous with a preverb. The preverb status of (13) is more secure.

- (12) *Roowil innyij jirrma-nyarr.*  
 walk 3-TR-catch-MID.PERF singing-COMIT.  
 ‘He walked away singing.’ (Aklif 1990-1994, E0/11)

<sup>6</sup> For details about the limitations, see Nash (1982).

- (13) *Banyjoord gorna [bangalonngan oogool] irrjimbin=jamb*  
 poison root good reef holes-ALL/PURP scatter 3PL-die-CONT=THUS  
*aarli.*  
 fish.  
 ‘Banyjoord poison root is used for scattering in crevices of reefs so that the fish die.  
 (Aklif 1999)

Most preverbs do not follow this pattern, however, and instead an infinitival form of the light verb appears (that is, there is a nominalized verb phrase instead of a single preverb). (14) gives an example. In (14a), *manyan* is the infinitive of the light verb *-(i)nya-* ‘catch’; (14b) is an equivalent sentence, without the infinitive. It is ungrammatical.

- (14) a. *Roowil-ngan manyan gorn=amb.*  
 walk-ALL GER-catch-CONT good=THUS  
 ‘It’s good to walk.’ (CB/FN: NI. 11/26)
- b. \**Roowil-ngan gorn=amb.*  
 walk-ALL good=THUS

Such behavior is again a difference between Bardi and many other Northern Australian languages. Other members of the same language family allow more freedom than Bardi (e.g. Yawuru; cf. Hosokawa 1991). Two examples are given here from Wagiman. In (15) we see a list of actions; note that only the first has a light verb. In (16) we have an example of an imperative.

- (15) *Ngi-ya-nggi woerrkge-ma maman // garatjjin dorroh-dorroh // denh-na*  
 1PL-go-PAST work-ASP good grass pull out-REDUP cut-ASP  
*wirin.*  
 tree.

‘We worked well, pulling out grass, cutting trees.’

- (16) *Gurrh-ma welin!*  
 dig-ASP hole  
 ‘Dig a hole!’ (Wilson 1999, 73, 81)

Such examples are even more frequent in some other Northern Australian languages, for example in Jaminjung (Schultze-Berndt 2001) and Wardaman (Merlan 1994). In some languages, nearly half the verbal clauses in texts appear without an accompanying light verb.<sup>7</sup>

<sup>7</sup> It should be noted that the ability of apparently non-finite forms (preverbs) to appear without finite verbs is an interesting piece of evidence against the idea that light verbs are argument licensers; that is, that the arguments of the preverb are transferred to an item capable of assigning case and licensing them within the clause (e.g. Grimshaw and Mester 1988). I leave this point for further discussion, since I assume a unification analysis here and the point is immaterial to this paper.

### 4.3. Preverb inflection

Although preverbs in these Northern Australian languages are typically characterized as ‘uninflecting’, most languages allow limited inflection on the preverbs. Bardi is unusual in allowing only reduplication; there is no other productive preverb inflection in this language (and the unproductive inflection is so rare that the meaning of the morpheme cannot be identified with any certainty). An example of reduplication is given in (17) below. Reduplication in the preverb marks iterativity or distributivity.

- (17) a. *Garrjagarrja anama!*  
 sharpen 2.IMP-TR-put-FUT  
 ‘Sharpen it!’ (NI: CB/20.6:54)
- b. *Bawinbawin ingirrinnyagalirr goorlil.*  
 cut up-REDUP 3PL.PST-‘catch’-REC.PST=3PL.OBJ turtle.  
 “They were cutting up the turtles.”

When we compare the examples with the data for Wagiman, we again see a difference. In (18), for example, the preverb takes the negative imperative suffix (and there is no accompanying light verb) – this is the most common way to mark negative imperatives in the language. In example (19) we have an allomorph of the aspectual morpheme *-ma*, while in (20) the preverb is nominalized and exhibits the ablative case marker *-gunda*.<sup>8</sup>

- (18) *Lurt-wehen danganyin!*  
 give-NEGIMP tucker  
 “Don’t give (them) tucker!” (Wilson 1999, 58)
- (19) *Baningh-nga mu-yama?*  
 do what-ASP 2PL-do.FUT  
 “What are you lot going to do?” (Wilson 1999, 50)
- (20) *Dilh-dil-may-gunda nga-nyar-ma-n lari.*  
 write-REDUP-NOM-ABL 1SG-be tired-VERB-PREB arm  
 “my arm is tired from writing.” (Wilson 1999, 85)

A similar array of inflectional possibilities is to be found in Wardaman (Merlan 1994). In contrast, in Warlpiri, the only inflection allowed on preverbs is marking for trajectory.<sup>9</sup>

<sup>8</sup> According to Wilson (1999) not all speakers of Wagiman use the nominalizer in such constructions.

<sup>9</sup> Note, incidentally, that trajectory is the only category which forms doubled preverbs in Bardi; see example (8) above.

#### 4.4. Preverbs, light verbs, and valency

Finally, let us consider some of the issues involved in the specification of argument structure in Northern Australian complex predicates. As is typical in complex predicates of this type, the transitivity of the clause is determined by several different parts of the predicate. In Bardi, for example, the verbal morphology, the preverb and the inflecting verb all influence the transitivity of the predicate. (21) gives an example of the reflexive/reciprocal circumfix on the inflecting verb and its effects on argument structure; (21a) is transitive and contains the first person object marker =(jarr)ngay; (21b) on the other hand is intransitive and exhibits a single participant.<sup>10</sup>

- (21) a. *Barn injoo galjarrngay.*  
 tell 3-PST-do/say-IMPERF-1 MIN.DO  
 ‘He told me to do something.’
- b. *Barn ingim.inyjigal.*  
 tell 3-PST-REFL<sub>1</sub>-‘do/say’-REFL<sub>2</sub>-IMPERF  
 ‘He thought about it.’

The choice of light verb also influences the transitivity of the predicate. In (22) and (23), for example, we see a pair of clauses which differ only in the choice of light verb. The first contains a single participant, while the second has two. In (23), *oola* ‘water’ is the subject of the clause, while in (22b), it is the direct object (Bardi is morphologically ergative; in both examples *oola* is in the absolutive case).

- (22) a. *Boorroolboorrool oonkara oola.*  
 boil-redup 3MIN.FUT-spear-FUT water  
 ‘The water will boil.’
- b. *Boorroolboorrool oonkama oola.*  
 boil-redup 3MIN.FUT-put-FUT water  
 ‘He/She’ll boil the water.’
- (23) a. *goojaj -joo-* ‘to feel weak’  
 b. *goojaj -ma-* ‘to make someone weak’

Finally, it is worth noting that the valency of the light verb itself is not straightforward. While monovalent light verbs always produce monovalent complex predicates, the bivalent light verbs do not always result in bivalent complex predicates. As seen from (24) below, a bivalent verb such as *-ma-* ‘put’ can appear with both

<sup>10</sup> Transitivity in Nyulnyulan languages is defined by the presence or absence of direct object agreement markers, by the cases in which free nominals appear (ergative and absolutive for transitive clauses) and by the number of possible free nominals which appear in the clause.

transitive and intransitive complex predicates.<sup>11</sup> Further examples are given in (25) with another light verb, *-(i)nya-* ‘get, pick up’.

- (24) *-ma-* ‘to put’ (2 obligatory arguments)
- a. *jibard -ma-* ‘to sneak up’ (1 argument)
  - b. *niya -ma-* ‘to rest’ (1 argument)
  - c. *oona -ma-* ‘to defecate’ (1 argument)
  - d. *wajim -ma-* ‘to wash something’ (2 arguments)
  - e. *garboo -ma-* ‘to dig around something’ (2 arguments).
- (25) *-(i)nya-* ‘to catch, to pick up’ (2 obligatory arguments)
- a. *ngalar -(i)nya-* ‘to have one’s eyes open’ (1 argument)
  - b. *marrmarr -(i)nya-* ‘to flash’ (1 argument)
  - c. *galgooriny -(i)nya-* ‘to swim breaststroke’ (1 argument)
  - d. *roowil -(i)nya-* ‘to walk’ (1 argument)
  - e. *joony -(i)nya-* ‘suck something’ (2 arguments)
  - f. *bawinbawin -(i)nya-* ‘cut up something’ (2 arguments)

The relationship between the valency of the light verb and the valency of the preverb is a source of major differences between Northern Australian languages. In some languages there is strict correspondence between the valency of the light verb and the number of arguments in the overall predicate. In Warlpiri, for example, the case frame of the clause and the number of arguments are directly inherited from the light verb; the preverb has no effect on this aspect of argument structure. We see from (26), for example, that the case frame for the complex predicate is the same as that of the inflecting verb.

- (26) a. *kanginypa-nya-nyi* ‘ERG fail to see ABS’ (*nya-* ‘see’ - ERG/ABS)  
 b. *kanginypa-karri-mi* ‘ABS fail to perceive DAT’ (*karri-* ‘stand’ ABS/(DAT))

In a few other languages, we see argument structure oddities of the same type that Bardi shows, with some differences. In Wagiman, for example, some monovalent light verbs may appear in transitive predicates:

- (27) *Galh-ma ngi-ya-nggi-ngana garradin.*  
 climb-asp 1pl-go-past-incl hill.  
 ‘we climbed the hill’ (Wilson 1999, 161)

<sup>11</sup> Transitive complex predicates are those where two free nominal arguments may appear, and where there are two agreement markers on the verb. The majority of such verbs have ergative/absolutive case marking and subject and direct object agreement marking.

Wagiman does not have complex predicates in which the light verb appears to be missing an argument; all examples are cases where the preverb introduces an extra argument (that is, where a monovalent light verb participates in a transitive complex predicate).

The only other language with the same type of ‘subtractive’ argument structure as Bardi (where the transitivity of the clause is not predictable from preverb or light verb alone) is Jaminjung Schultze-Berndt (2000). She discusses ‘dummy-undergoers’ of some transitive verbs, where an extra argument does not appear although it is cross-referenced in the agreement morphology.

- (28) *Ngayin=malang bul gani-ma bunyag.*  
 meat.animal=GIVEN emerge 3sg:3sg-HIT.PST 3dl.OBL  
 ‘The animal came out to/for the two.’ (Schultze-Berndt 2000, 181)

In this sentence the prefix chunk is transitive, marking a third person singular subject acting on a third person singular object. The preverb *bul* ‘emerge’, however, licenses only one argument. (The oblique pronoun *bunyag* is not cross-referenced by the verb.) Schultze-Berndt explains this as a morphology/syntax mismatch (that is, that the morphological structure of the verb requires an object morpheme to be present even if it has no exponent in the syntax). The same solution is possible for Bardi too, although it requires a theory of morphology and syntax in which this is possible.<sup>12</sup>

#### 4.5. Summary

Thus we have seen that Northern Australian languages (particularly Bardi and Wagiman) differ extensively in the syntax of their complex predication. In Bardi, but not necessarily in other Australian languages, preverbs form a close-knit unit with the light verb and are strictly preverbal. There are limited possibilities for conjunction, stacking, and appearance without a light verb. That is, preverbs differ in the extent to which they head phrasal categories. Preverbs can assign  $\theta$ -roles in Bardi and Wagiman, but the possibilities are more limited in other languages. Preverbs affect the transitivity of the predicate and case marking of core arguments in some languages, whereas in others (such as Warlpiri) case marking is determined entirely by the light verb.

Table 2 summarizes the behavior of Bardi and Wagiman.

#### 5. Discussion

From §4.5 we see that there are a number of differences between languages like Bardi versus languages like Wagiman. We have also seen that the differences are not directly correlated with the tightness of nexus between the preverb and the light verb, or the relationship between simple and complex predicates.

<sup>12</sup> See, amongst others, Samek-Lodovici (2003) for solutions in terms of argument suppression.



Table 2: Bardi and Wagiman compared

	Bardi	Wagiman
Definable word class?	Yes	Yes
LV contributes to argument structure?	Yes	Yes
CV contributes to argument structure?	Yes	Yes
Preverb stacking?	No	Yes
Light verb ellipsis?	No	Yes
Preverb ellipsis?	No	Yes
Aspect on preverbs?	No	Yes
Case on preverbs?	No	Yes (+ nom'l)
Valency of preverb = valency of predicate?	No	Yes

### 5.1. Analysis

What could be the underlying cause of the difference between complex predicates in Bardi and those in Wagiman? One way to analyze this would be to extend the mechanism used by Wilson (1999) for Wagiman, Butt (1995) for Urdu, and Alsina (1997). In these works, the preverb and light verb fuse by each providing pieces of Lexical Conceptual structure or LCS (in terms of Jackendoff 1990 and related works). The following example is taken from Wilson (1999, 154-155). Under the proposal, the light verb's LCS is defective, in the sense that it lacks the full information required to be a legitimate predicate.

- (29) a. ( $\uparrow$  PRED) = 'put< ( $\uparrow$  SUBJ)( $\uparrow$  OBJ)( $\uparrow$  OBL<sub>loc</sub>)>'  
 b. ( $\uparrow$  LCS) = [<sub>Event</sub> CAUSE([<sub>Thing</sub> ]<sub>A</sub>, [<sub>Event</sub> GO([<sub>Thing</sub> ]<sub>A</sub>, [<sub>Path</sub> TO([<sub>Place</sub> ]<sub>A</sub>)])])]

In (29b) the <sub>A</sub> annotations mean that the material corresponds to an argument in the a-structure representation of the verb. The a-structures and the arguments in the LCS are linked together by a mapping of arguments to functions (e.g. CAUSE, GO, etc).

The complex predicate is composed by unifying the LCS of the light verb with the LCS of the preverb by a process of 'predicate fusion' (Wilson 1999, 136ff). The LCS of the preverb merges the LCS of the light verb wherever it can do so without violating semantic wellformedness.<sup>13</sup>

- (30)
- |                           |  |  |   |
|---------------------------|--|--|---|
| <i>guk</i> 'sleep'        |  | [BE <sub>Id</sub> ([ <sub>Th</sub> ] <sub>A</sub> , [AT <sub>Id</sub> ( <i>asleep</i> )]) ]  | ] |
|                           |  | ↓  | ↓ |
| -ge- 'put <sub>LT</sub> ' | [CAUSE([ <sub>Th</sub> ] <sub>A</sub> , [BECOME( | [BE <sub>Id</sub> ([ <sub>Th</sub> ] <sub>A</sub> , [pl — ] ) ])                             | ] |
|                           | ↓  |  | ↓ |
| <i>guk-ge-</i>            | [CAUSE([ <sub>Th</sub> ] <sub>A</sub> , [BECOME( | [BE <sub>Id</sub> ([ <sub>Th</sub> ] <sub>A</sub> , [AT <sub>Id</sub> ( <i>asleep</i> )]) ]) | ] |

<sup>13</sup> There are implications in allowing the merger or fusion of LCSs within LFG, as in classical LFG PRED features are the only features which do not unify. we do not pursue these issues here.

(30) illustrates the merger of the LCS of the preverb *guk* ‘sleep’ with the light verb *-ge-* ‘put’. The light verb argument structure is missing part of its predicate (the attribute of the argument of BE; this is filled by the attribute of [AT ()] in the preverb. An informal definition of predicate fusion, as illustrated above, is given below:

(31) Predicate fusion (informal) (Wilson 1999, 166)

The LCS of a cover can fuse into any position of the LCS of an inflecting verb where it is able to unify.

Now, as noted by Wilson (1999, 154), the light verb provides information about the event structure of the complex predicate, while information about the argument structure comes jointly from the light verb and the preverb in some cases. As also noted above in §4.2, both the light verb (in its non-light function) and the preverb are predicational objects.

Now, in the above examples, the light verb LCS is missing certain pieces of information, but the preverb LCS is complete. Thus in Wagiman, there are many preverbs which should be complete predicates in their own right (and indeed, as we saw above, Wagiman preverbs do indeed head predicates in their own right). Now, what if we supposed that in languages with the behavior of Bardi, preverbs do not have a full LCS, but rather are missing some piece of information that licences them as full predicates? Bardi light verbs provide information about the event structure of the predicate (for example, whether it is a state, an activity, or if it involves a causer). Now, suppose that the preverb in Bardi were missing that information: preverb-light verb unification would still take place, resulting in a grammatical predicate. However, the preverb would not be able to appear without a light verb, except in the case where the deficiency in LCS structure could be repaired in another way (for example, in cases where argument structure need not be specified, such as when the preverb is in a nominal position).

The main disadvantage of such an analysis is the lack of constraints on what a partial LCS might (or might not) contain. There is nothing in the literature on predicate fusion (or the type used, for example, by Wilson 1999) to limit defective LCS structures to light verbs, or to the argument structure of light verbs. Since two LCSS in a predicate may unify in any configuration that results in a grammatical structure, it is not at all clear that it would be possible to constrain defective argument structures except by stipulation. On the other hand, defective LCSS are already constrained in that the missing information must appear somewhere in the LCS of the predicate.

## 5.2. Complex predicates as an areal feature

Finally, let us briefly consider the status of complex predicates as an areal feature in Northern Australia (cf. Dixon 2002, 188 and Schultze-Berndt (2000, 532-535)). It has been claimed that complex predicates are highly borrowable and diffusible

across linguistic boundaries – note for example that the area in Figure 1 above crosses several genetic boundaries, including that separating Pama-Nyungan and non-Pama-Nyungan languages. However, there has been no detailed work yet on the relative borrowability of preverbs as a category as opposed to, for example, nouns. That is, we have no detailed information about the processes by which complex predicates arise in different languages and what is the relative role of borrowing lexical items (such as preverbs or inflecting verbs) versus the diffusion of a particular syntactic construction. Even if we assume that the construction has spread primarily due to diffusion, we still do not know precisely what has diffused. Is it the category of preverbs? Or is it a particular consideration of preverbs and light verb? Or is it something else again? Detailed studies of the histories of individual languages and reconstruction of language families are needed before this question can be answered, and as mentioned above we lack detailed historical information for many languages and language families in Australia.

The different syntax of languages such as Bardi when contrasted with languages such as Wagiman may imply that we are not in fact dealing with the category which has diffused in the recent past, contra Dixon (2002). On the other hand, if it really is a single parameter such as a defective argument structure in preverbs, which accounts for the differences we see in these different languages, perhaps a change in type would not require a long time to be completed. However, it is not at all clear what would trigger such a change in the first place, or whether such a syntactic change is attested elsewhere in the world.<sup>14</sup>

## **6. Conclusions**

In conclusion, therefore, the diversity we see amongst northern Australian complex predicates is not at all confined to the three categories usually mentioned in the literature. We see that the syntax of complex predicates is rather different from language to language, and the differences range from the phrasal status of the preverb to the argument structure determinants of different parts of the predicate. Furthermore, we see that the differences cluster and could be described in terms of an analysis using a modified form of LFG. However, there are many languages of northern Australia for which we do not have sufficient data for us to draw further conclusions.

---

<sup>14</sup> One place where such a change may be attested is in the Turkic languages of Central Asia, however much more work is necessary in order to determine whether the behavior of complex predicates in central Asian Turkic languages is the result of a historical change and if so, what syntactic changes were involved.

## References

- Aklif, Gedda, 1990-1994. Bardi field notes. Unpublished notes. References are to Book/Page number.
- Aklif, Gedda, 1999. *Ardiyooloon Bardi ngaanka: One Arm Point Bardi dictionary*. Halls Creek, Western Australia: Kimberley Language Resource Centre.
- Alsina, Alex, 1997. Causatives in Bantu and Romance. In Alsina *et al.* (1997), 203–245.
- Alsina, Alex, Bresnan, Joan, and Sells, Peter, eds., 1997. *Complex predicates*. Stanford, California: CSLI Publications.
- Birk, D., 1976. *The MalakMalak language (Daly River)*. Canberra: Pacific Linguistics.
- Bowern, Claire, 2004. *Bardi verb morphology in historical perspective*. PhD dissertation, Harvard University, Cambridge, Massachusetts.
- Butt, Miriam, 1995. *The structure of complex predicates in Urdu*. Stanford, Calif.: CSLI Publications.
- Butt, Miriam and Geuder, Wilhelm, 2001. On the (semi)lexical status of light verbs. In Norbert Corver and Henk van Riemsdijk, eds., *Semi-lexical categories : the function of content words and the content of function words*, 323–369. Mouton.
- Chadwick, Neil, 1975. *A descriptive study of the Djingili language*. Canberra: AIAS.
- Dixon, R.M.W., 2002. *Australian languages: their nature and development*. Cambridge: Cambridge University Press.
- Grimshaw, Jane and Mester, Armin, 1988. Light verbs and  $\theta$ -marking. *Linguistic Inquiry* 19(2):205–232.
- Heath, Jeffrey, 1984. *A functional grammar of Nunggubuyu*. Canberra: Australian Institute of Aboriginal Studies.
- Hosokawa, Komei, 1991. *Meaning in Yawuru: A semantically oriented description of an indigenous language of the Kimberley region, Western Australia*. Canberra: to be published by Pacific Linguistics. [author's Ph.D. Thesis, Australian National University].
- Jackendoff, Ray, 1990. *Semantic structures*. Cambridge, Mass: MIT Press.
- McGregor, William B., 1990. *A functional grammar of Gooniyandi*. Amsterdam: John Benjamins.
- McGregor, William B., 2002. *Verb classification in Australian languages, Empirical Approaches to Linguistic Typology*, volume 25. Berlin: Mouton de Gruyter.
- Merlan, Francesca, 1994. *Wardaman*. Berlin: Mouton.
- Nash, David, 1982. Warlpiri verb roots and preverbs. *Work Papers of SIL-AAB*, volume A6, 165–216. Darwin: Summer Institute of Linguistics.
- Rumsey, Alan, 1982. *An intra-sentence grammar of Ungarinjin, north-western Australia*. B86, Canberra: Pacific Linguistics.
- Samek-Lodovici, Vieri, 2003. The internal structure of arguments and its role in complex predicate formation. *Natural Language and Linguistic Theory* 21:835–

- 881.
- Schultze-Berndt, Eva, 2000. *Simple and complex verbs in Jaminjung. A study in event categorisation*. PhD thesis, Katholieke Universiteit Nijmegen, Nijmegen, the Netherlands.
- Schultze-Berndt, Eva, 2001. Preverbs as an open word class in northern Australian languages. Preverb Workshop, Nijmegen, 19th January, 2001.
- Wilson, Stephen, 1999. *Coverbs and complex predicates in Wagiman*. CSLI Publications, Cambridge University Press.

Claire Bower  
Department of Linguistics  
Yale University  
370 Temple St, Room 204  
PO Box 208366  
New Haven, CT 06520-8366

[claire.bowern@yale.edu](mailto:claire.bowern@yale.edu)



# Composite Tone in Mian Noun-Noun Compounds

SEBASTIAN FEDDEN  
*University of Melbourne*<sup>1</sup>

## 1. Introduction

This paper deals with composite tone in Mian noun-noun compounds. The term “composite tone” is used here for a combination of the two tonal melodies for which the respective noun roots are lexically specified. The autosegmental analysis proposed in this paper assumes that composite tone melodies are assigned to the whole word after compounding and identifies word stress as the decisive prosodic feature which determines tone assignment.

Mian is a Papuan language of the Ok family (Healey 1964, Wurm 1982), spoken in Telefomin District of Sandaun Province in Papua New Guinea by approximately 1,400 people. In terms of larger-order affiliation, the Ok languages belong to the Trans New Guinea family (Wurm 1982, Pawley 2006).

Mian has 13 consonantal phonemes /b, g, t, k, m, n, ŋ, f, s, h, l, w, j/ and ten vowel and diphthong phonemes /a, a<sup>ʕ</sup> (i.e., pharyngealized a), e, i, ɔ, u; ai, ei, au, ou/. Voiced stops are slightly pre-nasalized. Morphosyntactically, the language is head-marking with almost no nominal morphology, but complex, mainly suffixal verb morphology. Serial verb and clause chaining constructions are very common. SOV is the unmarked word order.<sup>2</sup>

Mian is a word-tone language, i.e., a lexical item is specified for one out of a set of five tonemes, which is assigned to the word as a whole. Tone is mainly used for lexical distinctions; it only plays a minor role in the marking of grammatical categories. The toneme inventory will be discussed in detail below.

In many noun-noun compounds, root tones exist peacefully side by side, e.g., if both roots are specified for low tone (throughout this paper capital letters are used to indicate underlying tone and accents are used to show pitch in phonetic representations):

---

<sup>1</sup> Note: 2006 affiliation. Affiliation as of publication: University of Surrey.

<sup>2</sup> All data was recorded by me during nine months of field work in New Guinea.

(1)	Root 1:	<sup>L</sup> wan/	[wàn]	‘bird’
	Root 2:	<sup>L</sup> am/	[àm]	‘house’
	Compound:		[wà.nàm]	‘bird house’

Similarly, if the second root is specified for a contour tone, but the first root has a level tone:

(2)	Root 1:	<sup>L</sup> wan/	[wàn]	‘bird’
	Root 2:	<sup>LH</sup> a <sup>s</sup> n/	[ǎ <sup>s</sup> :n]	‘feather’
	Compound:		[wà.nǎ <sup>s</sup> :n]	‘bird feather’

However, in compounds, in which the first root is specified for a contour tone, root tones form a composite tone pattern.

(3)	Root 1:	<sup>LH</sup> gaba <sup>s</sup> m/	[ <sup>l</sup> gè.βǎ <sup>s</sup> :m]	‘head’
	Root 2:	<sup>L</sup> tɔl/	[t <sup>h</sup> ɔ̃l]	‘power’
	Compound:	<sup>LH</sup> gaba <sup>s</sup> m/+ <sup>L</sup> tɔl/	[ <sup>l</sup> gè.βǎ <sup>s</sup> :m.t <sup>h</sup> ɔ̃l]	‘brain’

The tonal melody of the compound in (3) seems to preserve the linear order of root tones. Nothing is added and nothing is deleted, but tones do not align with the morphemes to which they originally belonged; the high pitch has moved to the second stem /tɔl/. Before I discuss the tonology of Mian compounds in more detail, I will give a short account of tone behavior in word-tone languages.

## 2. Word-Tone

In a typology of tone systems based on the domain of tonal contrast, Donohue (1997:373) distinguishes word-tone systems, which use the whole word as the relevant tone assignment domain, found e.g. in Mende (a Mande language from West Africa) and Shanghai (Sino-Tibetan), from syllable-tone systems, in which each syllable is allowed to bear a distinctive tone independent of the other syllables in the word, as for example in Mandarin, Cantonese (both Sino-Tibetan), Vietnamese (Mon-Khmer), Igbo (a Kwa language from West Africa), and Chuave (Papuan, Chimbu Province).

In a word-tone language, a limited number of lexically specified tone patterns account for the tonal surface specification of entire mono- and polysyllabic words. Mian has five underlying tonal melodies (which is a typical toneme inventory for a Papuan word-tone language): high (H), low (L), rising (LH), peaking (LHL), and falling (HL). L and LH are very common. H is somewhat less common and very rare in polysyllabic words. LHL is very rare on monosyllabic and disyllabic words. Finally, HL is very rare and does not appear on monosyllabic words at all. Table 1 gives an overview of distinctive Mian tonemes and their phonetic realization on mono- and disyllabic words.



*Composite Tone in Mian Noun-Noun Compounds*

Table 1: Tonal minimal pairs

<b>Monosyllables</b>			
<b>Toneme</b>	<b>Phonemic</b>	<b>Phonetic</b>	<b>Meaning</b>
H	/ <sup>H</sup> an/	[áːn]	‘arrow; bow’
L	/ <sup>L</sup> am/	[àːm]	‘house’
LH	/ <sup>LH</sup> a <sup>˘</sup> m/	[ǎːm]	‘pandanus species’
LHL	/ <sup>LHL</sup> a <sup>˘</sup> m/	[àːm̃]	‘older sister’
H	/ <sup>H</sup> mən/	[méːn]	‘child’
LH	/ <sup>LH</sup> mən/	[mǎːn]	‘string bag’
<b>Disyllables</b>			
<b>Toneme</b>	<b>Phonemic</b>	<b>Phonetic</b>	<b>Meaning</b>
H	/ <sup>H</sup> səku/	[séːk <sup>h</sup> úː]	‘bush knife’
L	/ <sup>L</sup> afet/	[àːfɛt <sup>h</sup> ]	‘different’
LH	/ <sup>LH</sup> afet/	[ǎːfɛt <sup>h</sup> ]	‘cleared of taboo’
L	/ <sup>L</sup> ibal/	[iːβàː]	‘dust’
LHL	/ <sup>LHL</sup> ibal/	[iːβàː]	‘paper wasp’
LH	/ <sup>LH</sup> usan/	[ùːsǎːn]	‘vomit (n.)’
HL	/ <sup>HL</sup> usan/	[úːsân]	‘tail’

Table 2 shows the number of nouns occurring with a certain tonal specification out of a total of 200 monomorphemic native Mian nouns which I have used as data for the tone analysis.

Table 2: Toneme frequency

	<b>Monosyllables</b>	<b>Disyllables</b>
H	21	2
L	40	43
LH	52	37
LHL	1	3
HL	n/a	1

Tone assignment in monomorphemic Mian words can be modelled with the following autosegmental formalism (cf. Goldsmith 1990):

- a) Only vowels are tone-bearing units (TBUs).
- b) The penultimate tone in a melody is assigned to the TBU in the stressed syllable.
- c) The last tone is assigned to the word-final TBU.
- d) Remaining tones are mapped to TBUs on a one-to-one basis from left to right.
- e) Leftover TBUs are assigned tone by spreading.

There are two constraints on tone placement: First, the tonal melody a (lexical) word is specified for can maximally consist of 3 tones (MAX3) and second, contour tones are forbidden on non-final syllables (NOCONTOUR).

Stress and tone are distinct phenomena. The domain of stress in Mian is the word. Stress regularly falls on the final syllable in nouns, but can fall on the first syllable in disyllabic nouns if this syllable contains a pharyngealized /a/. As specified under (b) above, initial tone assignment depends on stress.

Stress behavior in Mian is an example of the complex interplay between segmental specification and prosody. A certain feature of a segment attracts stress which in turn attracts tone. A related phenomenon can be found in Warembori (probably Papuan but with strong Austronesian influence, north coast of Irian Jaya). Warembori (Donohue 1999:8-9) is not tonal but has two sets of nasal and voiced stops (a “normal” and a “heavy” set), which are pronounced the same but which have different effects on their segmental environment and on prosody; namely, a syllable with a consonant from the heavy series attracts stress.

Pharyngealization as an articulatory feature can be described as a gesture of the tongue root involving active retraction; the pharyngeal passage is narrower and the larynx slightly raised (Ladefoged and Maddieson 1996:306). The pharyngealized /a/ in Mian is distinct from the non-pharyngealized /a/ in that it has a noticeably lower fundamental frequency. Moreover, it is strongly associated with L or LH tonemes, has a lower third formant, and is noticeably longer in the same position and under the same tone. The pharyngealized /a/ as a phoneme is not quite as easy to establish than any of the other Mian segmental phonemes because the degree of pharyngealization varies greatly between speakers and there is only a handful of genuine minimal pairs (i.e. pairs which only differ in pharyngealization, not in tone or segmental make-up):

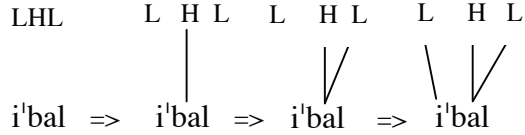
- (4) /a/     [à]     ‘feces’  
       /a<sup>ɕ</sup>/    [à<sup>ɕ</sup>·l]    ‘skin’

The derivations in (5) and (6) below illustrate how underlying tones are assigned to mono-morphemic words by the formalism:

- (5) /<sup>h</sup>i'bal/ [i.βàl] ‘dust’
- |          |          |       |
|----------|----------|-------|
| L        | L        | L     |
|          |          | /     |
| i'bal => | i'bal => | i'bal |

*Composite Tone in Mian Noun-Noun Compounds*

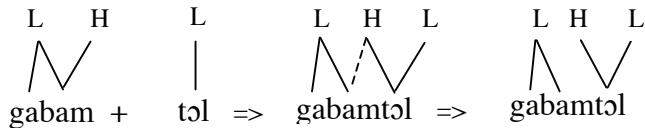
- (6)  $/^{LHL}i'bal/$  [*i.'βâl*] ‘paper wasp’



**3. Mian Composite Tone**

As shown in (3) above, in some compounds the tonal melodies of both roots are preserved in linear order and form a composite tone pattern which is created by a process that looks like “tone shunting,” set off by the language-specific constraint NOCONTOUR which outlaws tonal contours on non-final syllables. This process can be shown autosegmentally as follows (assuming for the moment that tone association takes place before compounding):

- (7)  $/^{LH}gaba^{\text{H}}m/+^{L}t\text{øl}/$  [*ᵒgə.βa<sup>ᵛ</sup>m.t<sup>h</sup>øl*] ‘brain’



What the metaphor of “tone shunting” is supposed to capture is that tones (the train cars) are moved from one root (track 1) to another root (track 2) in the compound (the train station). In the process, tones are neither copied nor deleted, nor is their linear order altered. Changing the sequence of train cars on a given track might be possible, but cars do not duplicate, nor do they disappear when shunted. Other compounded nouns that behave identically are:

- (8)  $/^{LH}una\eta+^Lam/$  [*ùnà.ηâm*] ‘women’s house’  
 $/^{LH}mi\text{l}+^Lb\text{l}\text{ô}\eta/$  [*mìl.blôη*] ‘bean pod’  
 $/^{LH}f\text{ùt}+^Lb\text{l}\text{ô}\eta/$  [*fùt.blôη*] ‘cigarette pack’

Now, could this process of tonal change also be captured by a tone sandhi rule which spells out how the tonal specification of one root is changed due to the tonal specification of the other root in the complex word and vice versa? Sandhi rules are a common way of accounting for tonal change in compounds; see for example Newman and Petterson (1990) for Kairi (Papuan, Gulf Province) and Franklin (1971) for Kewa (Papuan, Enga Province). A possible sandhi solution would look like (9):

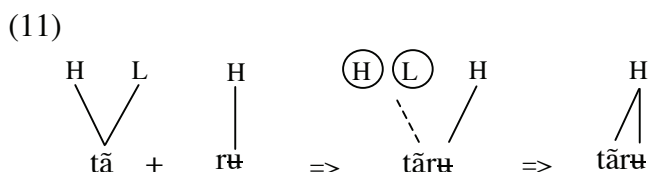
- (9)  $LH+L \rightarrow L+HL$

Although such a solution is possible, it is not desirable for two reasons: (a) both root tones have to be altered when what happens is only the H somehow being moved to the right, and (b) sandhi rules are a much more powerful mechanism than is actually needed here because they commonly change tone identity and sequence, e.g. in processes of tone assimilation.

In the literature on compound tone behavior one finds another type, namely “tone deletion and spreading.” For example, in the Papuan language Skou (Donohue 2003) the tones of one root are deleted and the melody of the other root is spread over the whole compound:

- (10) Root 1: /<sup>H</sup>Ltã/ ‘blade’  
 Root 2: /<sup>H</sup>rɯ/ ‘handle’  
 Compound: /<sup>H</sup>Ltã+<sup>H</sup>rɯ/ ‘blade handle’

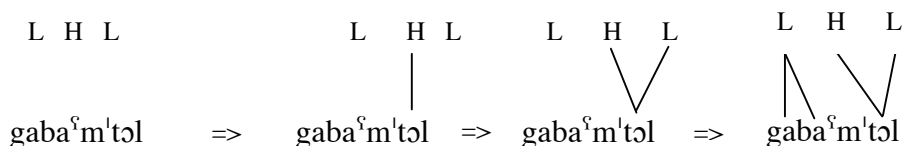
The derivation given by Donohue (2003:340) can be seen in (11):



The result is something like a truncated composite tone where part of the melody is simply deleted. Logically, the result is that Skou compounds always have a melody that conforms to one of the basic tonal melodies for mono-morphemic Skou words. A similar situation can be found in Kuman (Papuan, Chimbu Province) described by Hardie (2003).

Looking at the Mian case again, it is conspicuous that—as in Skou—tone in compounds does not seem to respect morpheme boundaries. The composite tone spreads over the whole word and consists of a basic tonal melody which can also be found on mono-morphemic words, namely LHL (the peaking toneme); cf. /<sup>LHL</sup>ibal/ ‘paper wasp’ in example (6). If this composite melody is assigned to the whole compound directly, that is after compounding has taken place, the formalism specified above comes up with the correct result because stress placement in the compound makes sure that tone is assigned correctly to the compound as whole, as in (12):

- (12) /<sup>LHL</sup>gaba<sup>̣</sup>m<sup>̣</sup>tɔl/ [ᵑgə̀.βá<sup>̣</sup>ṃ.tʰɔ̀l] ‘brain’



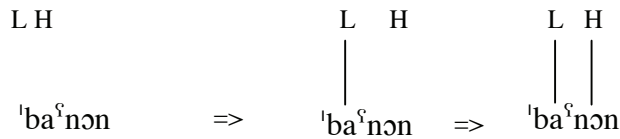
*Composite Tone in Mian Noun-Noun Compounds*

A similar solution can be found, when the second member in the compound bears a high tone, for example:

- (13)    Root 1:        /<sup>LH</sup>ba<sup>˧</sup>n/        [m<sup>˧</sup>bǎ<sup>˧</sup>:n]        ‘jaw’  
           Root 2:        /<sup>H</sup>ɔn/            [ɔ˧n]                ‘bone’  
           Compound:        [m<sup>˧</sup>bǎ<sup>˧</sup>.nɔ˧n]        ‘jaw bone’

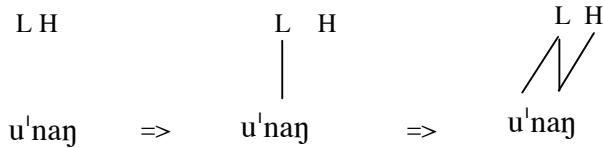
The basic melody LH is assigned to the compounded word. Again, initial tone assignment is determined by stress.

- (14)    /<sup>LH</sup>'ba<sup>˧</sup>nɔn/ [m<sup>˧</sup>bǎ<sup>˧</sup>.nɔ˧n] ‘jaw bone’

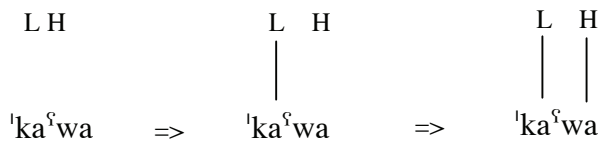


This is exactly parallel to the facts one finds in mono-morphemic words. Stress determines tone assignment. If the melody LH is assigned to the Mian word for ‘woman’, the result is [u˧nǎ˧ŋ], but if this melody is assigned to the word for ‘steel axe’, the outcome is [k<sup>h</sup>a<sup>˧</sup>wá˧] and not [k<sup>h</sup>a<sup>˧</sup>wǎ˧], as can be seen from the derivations in (15) and (16), respectively:

- (15)    /<sup>LH</sup>u˧naŋ / [u˧.nǎ˧ŋ] ‘woman’



- (16)    /<sup>LH</sup>'ka<sup>˧</sup>wa/ [k<sup>h</sup>a<sup>˧</sup>.wá˧] ‘steel axe’

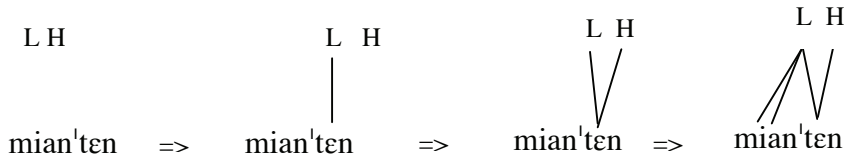


So far the examples given suggest that situations in which tones do not respect morpheme boundaries are restricted to compounds in which the first root is specified for a contour tone. This, however, is not the case.

- |      |           |  |   |               |
|------|-----------|--|---|---------------|
| (17) | Root 1:   | $/^L\text{mian}/$                      | $[\text{m}\grave{\text{i}}.\grave{\text{j}}\grave{\text{a}}\text{n}]$                                     | ‘Mian’        |
|      | Root 2:   | $/^H\text{t}\acute{\text{e}}\text{n}/$ | $[\text{t}^h\acute{\text{e}}\text{ˈn}]$   | ‘people’      |
|      | Compound: |  | $[\text{m}\grave{\text{i}}.\grave{\text{j}}\grave{\text{a}}\text{n}.\text{t}^h\acute{\text{e}}\text{ˈn}]$ | ‘Mian people’ |

In this case, root tone does not respect morpheme boundaries though the NOCONTOUR constraint is not violated. However, the resulting compound tone is not  $*[\text{m}\grave{\text{i}}.\grave{\text{j}}\grave{\text{a}}\text{n}.\text{t}^h\acute{\text{e}}\text{ˈn}]$  ‘Mian people’. Rather, there is a rising contour on the final syllable, which can be explained straightforwardly if it is assumed that a composite tone LH, which is put together from the two root tones, is assigned to the whole compound, as in (18):

- (18)  $/^LH\text{mian}^h\text{t}\acute{\text{e}}\text{n}/$   $[\text{m}\grave{\text{i}}.\grave{\text{j}}\grave{\text{a}}\text{n}.\text{t}^h\acute{\text{e}}\text{ˈn}]$  ‘Mian people’



So there are two facts which seem to be quite strong evidence that in Mian the domain of lexically specified tone is the word, whether it is mono- or polymorphemic: (1) composite tone does not refer to morpheme boundaries within the compound, and (2) composite tone assignment is dependent on stress and stress is a feature of words in Mian.

To conclude, the domain of Mian composite tone is the word. Compounds are words and therefore should be treated as words. They belong into the lexicon together with their tonal specification.

#### 4. MAX3-Constraint Violations

So far composite tone consists of maximally three tones and thus conforms to the MAX3 constraint. However, what happens if the composite tone melody violates this constraint? In other words, what if the composite melody consists of more than three tones?

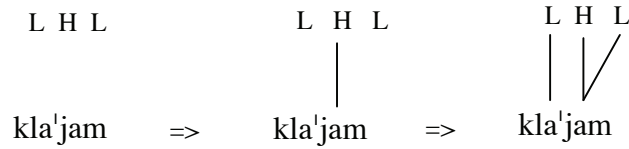
Before looking at the more complicated aspects of Mian compound tonology, a word about tone pattern frequency is necessary: while the exact degree of productivity of compounding is hard to estimate at the present stage, it is safe to say that Mian has a sizable number of both noun-noun and verb-verb compounds (the latter not being the focus of this paper) and a limited number of compounds involving other parts of speech. However, there are fundamental differences in tone pattern frequency. All patterns in which the composite tone has three tonal elements or less are more or less common, whereas the ones with more than three are very rare, which is to be expected because they violate the MAX3 constraint. So far there are three different cases:

*Composite Tone in Mian Noun-Noun Compounds*

- a) LHL+L            /<sup>LHL</sup>kla+<sup>L</sup>ajam/            ‘very good’ (very+good)<sup>3</sup>
- b) LH+LH           /<sup>LH</sup>fu+<sup>LH</sup>a<sup>ɿ</sup>n/            ‘tobacco leaf; letter’ (tobacco+leaf)
- c) LHL+LH        /<sup>LHL</sup>ibal+<sup>LH</sup>a<sup>ɿ</sup>n/ ‘paper wasp wing’ (paper wasp+leaf)

In the following, each of these cases is examined individually. The derivation for the LHL+L case is given in (19):

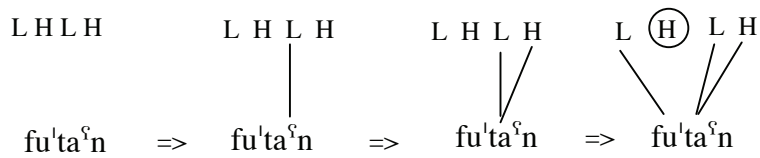
(19) /<sup>LHL</sup>klajam/ [kra.ˈjâm] ‘very good’



The compounded word has the basic melody LHL, which can be assigned straightforwardly to the word as a whole. However, the composite melody would have been LHL+L, but only LHL surfaces. This shows an important restriction on composite tone in Mian. In any tone sequence identical tones are fused. This tonal behavior is in accord with the Obligatory Contour Principle (OCP), originally proposed by Leben (1973), which rules out two identical autosegments following each other. To comply with the OCP, identical adjacent autosegments (tones, in this case), which end up next to each other due to morphological processes have to fuse into one before they are associated with elements on the segmental tier.

If two roots, each with a LH melody, are compounded, the composite melody is LHLH. However, this melody does not surface when the word is pronounced in isolation:

(20) /<sup>LHLH</sup>fu'ta<sup>ɿ</sup>n/ [fù.tʰá<sup>ɿ</sup>n] ‘tobacco leaf; letter’

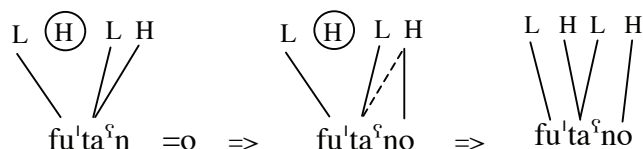


The first H remains floating and thus phonetically unrealized because there is no TBU left in the compound to which it could attach. The result is a complex word, which again conforms to the tonal restrictions of mono-morphemic words. However, when an inherently toneless article cliticizes to this word, the underlying composite melody LHLH surfaces again:

---

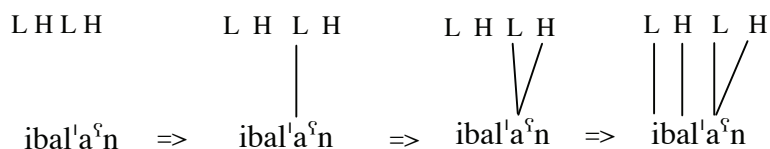
<sup>3</sup> This is not a noun-noun compound but nonetheless an interesting case because of its tonal phonology.

(21)



Obviously, the language takes great pains to preserve root tone in a composite melody LHLH even to the point of violating otherwise strong constraints, namely NOCONTOUR, which specifies that contours are only allowed on TBUs in final syllables. MAX3 appears also to be violated, but note that here the domain is a different one; we are not dealing with lexical words anymore but with prosodic words. Tonal behavior in the larger domain shows that the whole composite melody is preserved in the compound melody and that the compound has to be specified for the full composite melody LHLH. It might not be coincidence that it is the pharyngealized /a/ which carries the “illegal” second tone under these exceptional circumstances because it has the phonetic capability to do so due to its being longer than any other vowel phoneme.

In the third case LHL+LH, neighboring Ls are first fused to conform to the OCP and then the resulting composite melody LHLH is assigned to the compound. There is no violation of NOCONTOUR because the compound has enough TBUs to accommodate the whole composite melody.

(22)  $/^{LHLH}iba'l'a^{\circ}n/$  [i.βá.'lá<sup>o</sup>.n] ‘paper wasp wing’

Although it seems that MAX3 has been violated, there is nothing in the derivation in (22) that suggests a violation of any constraint. The four tones in the composite melody can be assigned to the three TBUs in the compound without problem. Obviously, the constraint is formulated too strictly. Rather, there should be a MAX constraint which specifies that the number of tones in a melody may at least be one larger than the number of TBUs in the word.

## 5. Summary

Table 3 below gives an overview of root tone combinations in Mian noun-noun compounds. The table shows that the tonal melody in attested Mian compounds is in each case a composite melody consisting of the two root tones with any neighboring identical tones fused. The table contains two root tone combinations which have not been covered in detail in the course of this paper, namely H+L



*Composite Tone in Mian Noun-Noun Compounds*

and H+H, but resulting composite tone in both cases conforms to the predictions of the proposed analysis. (As the toneme HL is only attested in a single monomorphemic word, namely /<sup>H</sup>usan/ [ú'sân] ‘tail’, which has not yet been found as an element in a compound, it has not been included in the table. Black frames indicate that the pattern is so far unattested. Rare patterns are marked grey.)

Table 3: Root tone combinations in noun-noun compounds

Root 2 \ Root 1	L	H	LH	LHL
L	L [wà.'nàm] ‘bird house’	LH [mì.'jàn.'thě'n] ‘Mian people’	LH [wà.nǎ <sup>s</sup> 'n] ‘feather’	
H	HL [án.'wâl] ‘bow string’	H [bá.'nón] ‘arm bone’		
LH	LHL [ <sup>o</sup> gə.βǎ <sup>s</sup> 'm.t <sup>h</sup> ôl] ‘brain’	LH [ <sup>l</sup> m bǎ <sup>s</sup> 'nó'n] ‘jaw bone’	L(H)LH [fù.'t <sup>h</sup> ǎ <sup>s</sup> 'n] ‘letter’	
LHL	LHL [krà.'jâm] ‘very good’		LHLH [ì.βál.'ǎ <sup>s</sup> 'n] ‘paper wasp wing’	

## 6. Conclusion and Further Research

This paper tried to show that the domain for tone assignment in Mian is the word, not the morpheme. Tonal melodies ignore morpheme boundaries in compounds. The result is composite tone which is assigned to mono- and polymorphemic words on the basis of word stress. The main difference between Mian and a language like Skou—in my analysis—is that in Skou, compound tone melodies always have to conform to tonemes which can also be found on monomorphemic words, whereas Mian does not have this restriction. Mian is an interesting case in that it wants root tones in some compounds to be preserved even if this leads to the violation of otherwise important constraints. However, basically Mian composite tone is word-tone.

This preliminary analysis of composite tone behavior in Mian noun-noun compounds is only a first step towards a fuller understanding of Mian tonology in complex words and across word boundaries. Further research will concentrate heavily on (a) verb tone, especially tone in verb-verb compounds, (b) the extent to which MAX-violating composite tone actually occurs in the language, and (c) the question as to whether the assumption of word stress as the decisive prosodic feature which determines initial tone assignment is tenable.

**References**

- Donohue, Mark. 1997. Tone Systems in New Guinea. *Linguistic Typology* 1(3):374-386.
- Donohue, Mark. 1999. *Warembori*. Languages of the World/Materials 341. Muenchen: Lincom Europa.
- Donohue, Mark. 2003. The Tonal System of Skou, New Guinea. In K. Shigeki, ed., *Proceedings of the Symposium Cross-linguistic Studies of Tonal Phenomena: Historical Development, Phonetics of Tone, and Descriptive Studies*. Tokyo University of Foreign studies: Research Institute for Language and Cultures of Asia and Africa, 329-365.
- Franklin, Karl. 1971. *A Grammar of Kewa, New Guinea*. Pacific Linguistics, C-16. Canberra: Australian National University.
- Goldsmith, John. 1990. *Autosegmental and Metrical Phonology*. Oxford: Blackwell.
- Hardie, Peter. 2003. Is Kuman Tonal?. MA sub-thesis. Canberra: Australian National University.
- Healey, Alan. 1964. A Survey of the Ok Family of Languages; Reconstructing Proto-Ok. Part of Ph.D. diss., Canberra: Australian National University.
- Ladefoged, Peter, and Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Oxford University Press.
- Leben, William. 1973. Suprasegmental Phonology. Ph.D. diss., MIT.
- Newman, John, and Robert Petterson. 1990. The Tones of Kairi. *Oceanic Linguistics* 29(1):49-76.
- Pawley, Andrew. 2006. The Chequered Career of the Trans New Guinea hypothesis: Recent Research and its Implications. In A. Pawley, R. Attenborough, J. Golson and R. Hide, eds., *Papuan Pasts: Cultural, Linguistic and Biological Histories of Papuan-speaking Peoples*, 67-108, Canberra: Pacific Linguistics.
- Wurm, Stephen. (1982): *Papuan Languages of Oceania*. Tuebingen: Narr.

Sebastian Fedden  
Surrey Morphology Group  
School of English and Languages  
University of Surrey  
Guildford GU2 7XH  
United Kingdom

s.fedden@surrey.ac.uk

# Reconciling *meng-* and NP Movement in Indonesian<sup>1</sup>

CATHERINE R. FORTIN  
*University of Michigan*

## 0. Introduction

In this paper, I present a new analysis of the Indonesian ‘voice marker prefix’ *meng-*: I argue that it is not a voice marker or a prefix at all, but instead an indefinite object clitic, unspecified for person and number, which has the effect of ‘antipassivizing’<sup>2</sup> the verb hosting it. To support my claim, I review restrictions on the distribution of *meng-*, which although previously noted (e.g. Saddy 1991, Cole and Hermon 1998), remained unexplained under previous analyses of *meng-*. I further suggest an incorporation analysis of *meng-* (Baker 1988) that accounts for these distributional restrictions.

## 1. ‘Voice’ in Indonesian

As an Austronesian language, Indonesian has traditionally been analyzed as having a system of voice-related morphology, albeit one somewhat less elaborate than that of Malagasy or Tagalog. Hence, *meng-* is typically considered to be an active voice marker (1)-(2),<sup>3</sup> by analogy with the passive voice marker *di-* (3).

- (1) Ali menulis/tulis surat ini.  
Ali *meng*-write/write letter this  
‘Ali wrote this letter.’

---

<sup>1</sup> I am indebted to my language consultants Kathy Triyana, Jingga Morry, and Nancy Surachman. For extremely helpful discussion about the ideas presented here, I am grateful to Acrisio Pires, Daniel Seely, Sam Epstein, and audiences at the Michigan Linguistic Society’s 2005 meeting (at Michigan State University) and at BLS 32, especially Mark Donohue. Any errors are my sole responsibility.

<sup>2</sup> Aldridge (2006) also suggests that Indonesian verbs marked with *meng-* are a remnant of ergativity in Indonesian, although the details of her analysis differ from the one presented here.

<sup>3</sup> The details and notation of this assumption vary. For Sneddon (1996) and Voskuil (2000), *meng-* simply marks active voice; for, e.g., Chung (1976), *meng-* marks transitivity or agentivity; for Postman (2002), *meng-* bears a [+active] feature; for Englebretson (2003), *meng-* indicates ‘agent trigger’. Under some analyses, *meng-* is not merely a marker, but a morpheme that has the function of Case-marking the direct object (e.g. Guilfoyle, Hung, and Travis 1992; Son and Cole 2004).

- (2) Ali mengira Hasan mengharap Dani mencintai Andi.  
 Ali *meng*-think Hasan *meng*-expect Dani *meng*-love Andi  
 ‘Ali thinks that Hasan expects that Dani loves Andi.’
- (3) Surat ini ditulis/\*tulis (oleh) Ali.  
 letter this PASS-write/\*write (by) Ali  
 ‘This letter was written by Ali.’

Descriptively speaking, *meng*- and *di*- behave similarly. They appear only on semantically transitive verbs (i.e. verbs that assign theta roles to both an external and internal argument) and never on semantically intransitive verbs (i.e. verbs that only have one theta role to assign to an argument), e.g. (4).<sup>4</sup>

- (4) Ali sedang tidur/\*menidur/\*ditidur.  
 Ali PROG sleep/\**meng*-sleep/\*PASS-sleep  
 ‘Ali is sleeping.’

The presence of *meng*- correlates with an agent-to-subject mapping (1), and the presence of *di*- correlates with a patient-to-subject mapping (3). However, *meng*- and *di*- differ in one (arguably crucial) respect: *meng*- is optional (1), while *di*- is not. Although not all passive clauses contain *di*-, the word order of a passive without *di*- (a ‘bare passive’) has a different word order (5) than the *di*-passive (3).

- (5) Surat ini Ali tulis/\*ditulis.  
 letter DEM Ali write/\*PASS-write  
 ‘This letter was written by Ali.’

Furthermore, *meng*- has several restrictions on its distribution that are difficult to explain under the assumption that *meng*- is simply a voice marker. As previously noted by, e.g., Saddy (1991) and Cole and Hermon (1998), *meng*- cannot appear in a clause within which a complement to the verb has been extracted. Descriptively, this has been characterized as movement of a *wh*-NP over a verb marked with *meng*- forcing the ‘deletion’ of *meng*- (Cole and Hermon 1998). In sum, *meng*- is unable to appear in its usual environment (on semantically transitive verbs) in case of overt successive-cyclic NP movement ‘over’ the verb.

Contexts in which *meng*- is prohibited include bare passives (6).

---

<sup>4</sup> There is a homophonous derivational morpheme, *meng*-, which is used in the formation of certain denominal and deadjectival verbs (i). However, the derivational *meng*- behaves differently and has a different distribution from the *meng*- that is the focus of the current paper, so it will be set aside here.

- (i) Adiknya sedang menangis/\*tangis.  
 child-3SG PROG *meng*-cry/\*cry  
 ‘His/her child is crying.’

*Reconciling meng- and NP Movement in Indonesian*

- (6) Buku ini saya (\*mem) baca.  
 book DEM 1SG (\*meng-)read  
 ‘I read that book.’

Second, *meng-* is also prohibited in object relative clauses (7a). Note that *meng-* appears in subject relative clauses (7b), as there has been no operator movement ‘over’ the verb within the relative clause.

- (7) a. Buku yang Ali (\*mem) baca itu menarik.  
 book that Ali (\*meng-)read DEM interesting  
 ‘The book that Ali read is interesting.’
- b. Laki-laki yang membaca buku itu adik saya.  
 man that meng-read book DEM sibling 1SG  
 ‘The man that’s reading a book is my younger brother.’

Third, *meng-* is prohibited in *wh*-questions within which a *wh*-NP has moved overtly over the verb, whether the *wh*-NP is the embedded clause object (8) or the embedded clause subject (9). Note that in (9), as the *wh*-NP has moved only ‘over’ the matrix verb *harapkan* ‘hope’, *meng-* can appear on the embedded verb *beli* ‘buy’.

- (8) Apa<sub>i</sub> yang kamu (\*meng)harapkan [<sub>CP</sub> t<sub>i</sub> Ali akan (\*mem)beli t<sub>i</sub>?  
 what that 2SG (\*meng-)hope Ali FUT (\*meng-)buy  
 ‘What do you hope Ali will buy?’
- (9) Siapa<sub>i</sub> yang kamu (\*meng)harapkan [<sub>CP</sub> t<sub>i</sub> t<sub>i</sub> akan membeli mobil ini?  
 who that 2SG (meng-)hope FUT meng-buy car DEM  
 ‘Who do you hope will buy this car?’

In addition to *wh*-movement to the matrix [Spec, CP], Indonesian also allows *wh*-in-situ and partial *wh*-movement to an intermediate [Spec, CP] (10). In partial movement contexts, *meng-* can appear on verbs above the landing site of the *wh*-NP. In *wh*-in-situ contexts, *meng-* can appear on any verb in the sentence, indicating that it is only overt NP movement that is incompatible with the presence of *meng-* on a verb (as argued by Cole and Hermon 1998).

- (10) Kamu mengharapkan [<sub>CP</sub> apa<sub>i</sub> yang Ali akan (\*mem)beli t<sub>i</sub>?  
 2SG meng-hope what that Ali FUT (\*meng-)buy  
 ‘What do you hope Ali will buy?’

The appearance of *meng-* on a verb is not incompatible with other types of *wh*-movement over it. Movement of a *wh*-adverbial (11) over a verb does not prohibit its bearing *meng-*, nor does the movement of a *wh*-PP (12). (12) also illustrates

that the incompatibility of *meng-* with *wh*-movement is not the result of an argument/adjunct distinction, but is instead the result of an NP/other kinds of XP distinction (as noted by Cole and Hermon 1998).

- (11) Mengapa<sub>i</sub> kamu mengharapkan [<sub>CP</sub> t<sub>i</sub> Ali akan membeli mobil t<sub>i</sub>?  
 why 2SG *meng*-hope Ali FUT *meng*-buy car  
 ‘Why do you hope Ali will buy a car?’
- (12) Kepada siapa<sub>i</sub> kamu mengharapkan [<sub>CP</sub> t<sub>i</sub> Ali akan memberi mobilnya t<sub>i</sub>?  
 to who 2SG *meng*-hope Ali FUT *meng*-give car-3SG  
 ‘To whom do you hope Ali will give his car?’

In sum, previous analyses of *meng-* as a ‘voice’, ‘agentivity’, or ‘transitivity’ marker leave several important questions unanswered. Why must such a marker be incompatible with *wh*-NP movement ‘over’ the verb hosting it, but not be troubled by movement of a *wh*-PP or a *wh*-adverbial?

In addition, if, as I have argued, *meng-* is not a voice marker, what kind of verbal prefix could it be? It is seemingly not a derivational morpheme, as the presence of *meng-* in an active clause contributes nothing to the meaning of the clause or the verb that bears it, as evidenced by its optionality in these contexts. It is likewise seemingly not an inflectional morpheme: *meng-* does not reflect agreement, tense, aspect, mood, person, number, or any of the categories typically associated with inflection (and which are accounted for within the structure of the clause motivated in Chomsky 1995, the syntactic framework adopted here). Finally, the behavior of *meng-* is unique in the set of Indonesian verbal prefixes (the others are *di-*; *ter-*, which is the involuntary/stative/abilitative marker; and *bar-*, which marks certain semantically intransitive verbs); only *meng-* is optional.

In the remainder of this paper, I will argue for a novel analysis of *meng-* that capitalizes upon its somewhat restrictive distribution. I suggest that the appropriate way to characterize the absence of *meng-* in such contexts is not that *meng-* is ‘deleted’ as the result of successive-cyclic NP movement over it, but instead that *meng-* has the effect of **blocking** successive-cyclic NP movement over it.<sup>5</sup> I argue that the ability of *meng-* to block NP movement can be accounted for under an analysis of *meng-* as a theta-marked object clitic pronoun.

## 2. *meng-* Is a Clitic Pronoun

### 2.1. When *meng-* Is Mandatory

In the previous section, I described contexts where *meng-* is prohibited from appearing in its customary environment and noted that *meng-* is otherwise optional in the context in which it is licensed: clauses containing a semantically

---

<sup>5</sup> An additional problem with characterizing the absence of *meng-* in these contexts as ‘deletion’ arises: within the Minimalist Program, it is not clear what it means to say that part of a word must delete, or how such a deletion could be implemented, given standard theories of ellipsis.

transitive verb with an agent-to-subject mapping. However, there are certain semantically transitive verbs that do not require the internal object to be overtly expressed (13a); in such contexts, *meng-* becomes mandatory (13b).

- (13) a. Ali sedang membaca (buku itu).  
Ali PROG *meng*-read (book DEM)  
'Ali is reading (that book).'
- b. \* Ali sedang baca.  
'Ali is reading.'

The obligatoriness of *meng-* only in contexts where no object NP is overtly expressed will immediately follow if *meng-* itself is the object NP.

## 2.2 The Nature of *meng-*

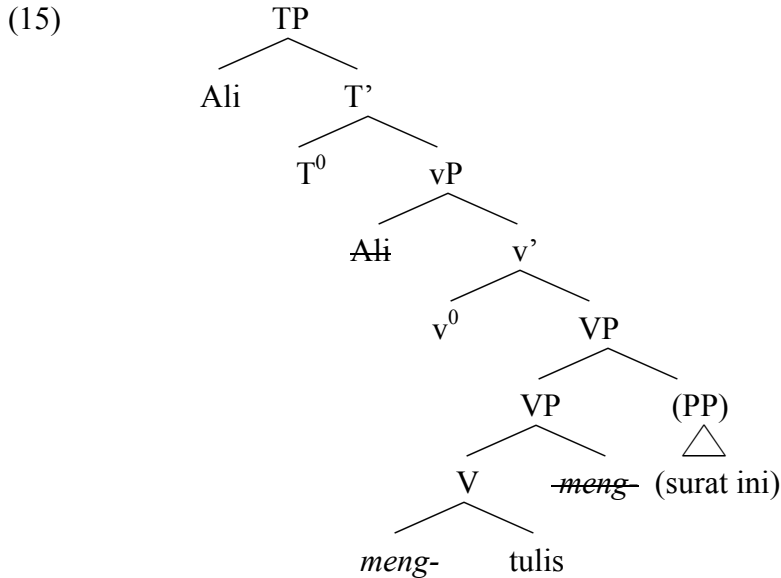
My proposal is that *meng-* does not merely signal that the clause is an active voice clause: *meng-* further has the effect of 'antipassivizing' the verb that hosts it. This is because *meng-* is not simply a piece of verbal morphology, but an indefinite object clitic pronoun unmarked for person and number. In a clause containing *meng-* (14), *meng-* originates in the theta position of the internal argument and moves to adjoin to the verb<sup>6</sup> (see, e.g., Baker 1988), as in (15). As in a true antipassive construction, in such clauses an overtly specified object NP (in this case, *surat ini* 'the letter') is optional.

- (14) Ali menulis (surat ini).  
Ali *meng*-write (letter DEM)  
'Ali wrote (the letter).'

---

<sup>6</sup> Contra, e.g., Guilfoyle, Hung, and Travis (1992), I assume that V-to-T movement does not take place overtly in Indonesian. Guilfoyle, Hung, and Travis make the assumption that it does for theory-internal reasons. Although the matter needs to be investigated more thoroughly, it seems to be true that V does not raise in the overt syntax, as it does not raise above adverbs that are generally assumed to mark the edge of the VP, such as *sering* 'often', as in (i). Nothing in my proposal hangs on this assumption, however.

- (i) a. Dia tidak sering menulis surat.  
3SG NEG often *meng*-write letter  
'He/she doesn't often write letters.'
- b. \* Dia tidak menulis sering surat.



In a clause without *meng-* (16), the object NP merges into the theta position of the internal argument. That *meng-* is mandatory when the clause contains no overtly expressed object NP follows: *meng-* in this case is required by the Theta Criterion, so that the verb's internal theta role is assigned. If this proposal is on the right track, there are no 'optionally transitive' verbs in Indonesian; in every case, the verb's internal theta role is assigned to some element, either *meng-* (if it is present) or a lexical NP.

- (16) Ali tulis surat ini.  
 Ali write letter DEM  
 'Ali wrote the letter.'

This analysis of *meng-* accounts for the prohibition of *meng-* on semantically intransitive verbs. If the verb has no internal theta role to assign to *meng-*, a violation of the Theta Criterion will result. Conversely, if the semantically intransitive verb has only an internal theta role to assign, the possibility of it being assigned to *meng-* is ruled out by its function: as in a true antipassive construction, the antipassivizing morpheme can never represent the only argument of a verb (as noted by, e.g., Baker 1988).

If both *meng-* and an overtly expressed object NP appear in the clause (as in (1)), only *meng-* originates in the theta position of the internal argument. The overtly expressed object NP is then a VP adjunct coreferential with *meng-*, such that the reference of *meng-* is determined by the adjunct NP (see, e.g., Baker 1988). Again, as in a true antipassive construction, if no object NP is overtly expressed, the object is interpreted as being indefinite, unknown, or unspecified. I assume that Case on the adjunct NP is checked by a null preposition, akin to the null preposition that checks the Case of the agent NP in *di*-passives (17). (In *by*-



phrases of *di*-passives, an overt preposition, *oleh* ‘by’, is optional.) As Indonesian displays no morphological case distinctions, oblique or otherwise, it is unsurprising that an ‘object’ NP in a *meng*-ful clause bears no morphological case marking to distinguish it from a true object NP in a *meng*-less clause.

- (17) Piring itu sudah dicuci oleh/∅ Pak Ali.  
 dish DEM already *di*-wash by Mister Ali  
 ‘These dishes were already washed by Pak Ali.’

‘Passive’ *by*-phrases containing the null preposition are prohibited from *wh*-extracting in the same way that the adjunct NP in a clause containing *meng-* is (18b), although *wh*-in-situ is available (18a). As noted above, it is possible to extract a *wh*-PP, as in (18c). There is something unique, then, with respect to the null prepositions found in passive *by*-phrases and in adjunct NPs in clauses containing *meng-*.<sup>7</sup>

- (18) a. Ali dipukul (oleh) siapa?  
 Ali PASS-hit (by) who?  
 b. \* Siapa Ali dipukul (oleh)?  
 who Ali PASS-hit (by)  
 c. Oleh siapa Ali dipukul?  
 by who Ali PASS-hit  
 ‘Who hit Ali?’

### 2.3. Why *meng-* and NP Movement Are Incompatible

If, as I have argued, *meng-* is the object NP in a *meng*-ful clause, and overt ‘object’ NPs in such clauses are in actuality adjuncts, we can easily explain why the presence of *meng-* in a clause disallows the possibility of extraction from the (apparent) complement, as in bare passives, object relative clauses, and *wh*-questions.

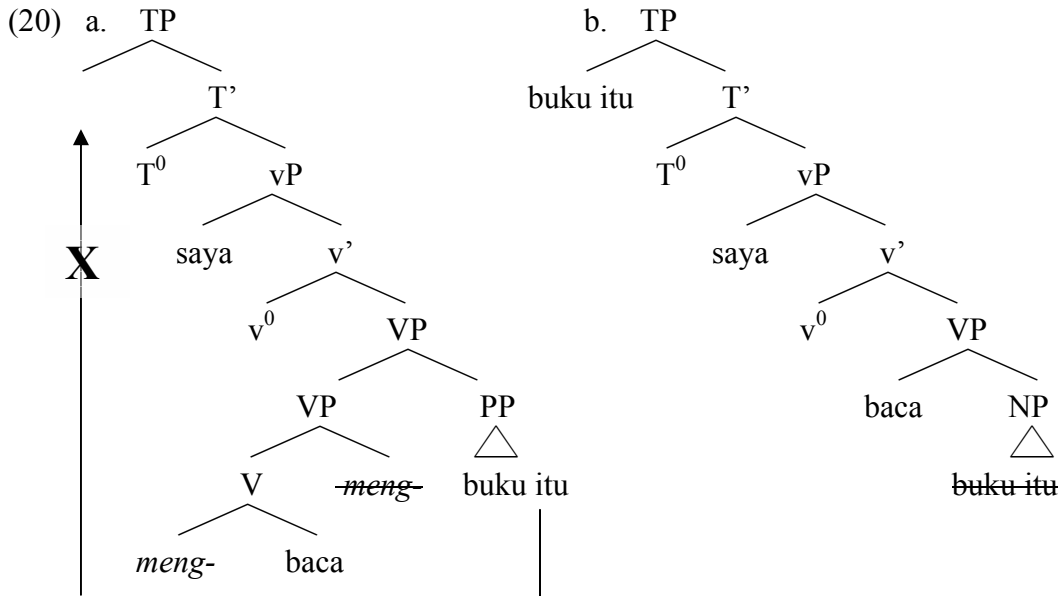
As noted above, in bare passives (19), *meng-* is prohibited.

<sup>7</sup> Mark Donohue (p.c.) has suggested that the ban on movement of the null preposition-headed passive *by*-phrase could be the result of the null preposition ‘incorporating’ into the verb. In Indonesian, *wh*-movement which strands the preposition is not possible (18b); the preposition must be pied-piped along with the *wh*-element (18c). Additionally, there is a strict adjacency requirement between the verb and a null preposition-headed *by*-phrase, although there is no such requirement for an *oleh*-headed *by*-phrase (i). For a more detailed analysis, please see Fortin (in preparation).

- (i) Ali dipecat kemarin \*(oleh) bosnya.  
 Ali PASS-fire yesterday (by) bos-3SG  
 ‘Ali was fired yesterday by his boss.’

- (19) Buku itu saya (\*mem) baca.  
 book DEM 1SG (\**meng-*)read  
 ‘I read the book.’

If *meng-* is present in the clause, *meng-* is first merged into the theta position of the object, while *buku itu* ‘the book’ is a VP adjunct (20a). Adjuncts are prohibited from raising to an A-position by the Chain Condition (Chomsky 1986). As an illustration, I will assume a canonical passive clause structure for Indonesian bare passives, similar in relevant respects to that proposed for Indonesian by Guilfoyle, Hung, and Travis (1992). Independently, Son and Cole (2004) have argued that Indonesian T<sup>0</sup> has an EPP feature which must be satisfied by an NP in its specifier position in the overt syntax; for this reason, I assume that *buku itu* is indeed in [Spec, T<sup>0</sup>] in (20b).<sup>8</sup>



Finally, in *wh*-questions where a *wh*-NP has moved overtly over a verb, *meng-* on that verb is prohibited. This is true of both monoclausal *wh*-questions (21b) and multiclausal *wh*-questions where a *wh*-NP originating in an embedded clause raises to a position in a higher clause (22b). In both kinds of questions, *wh*-in-situ does not prevent *meng-* from appearing (21a), (22a).

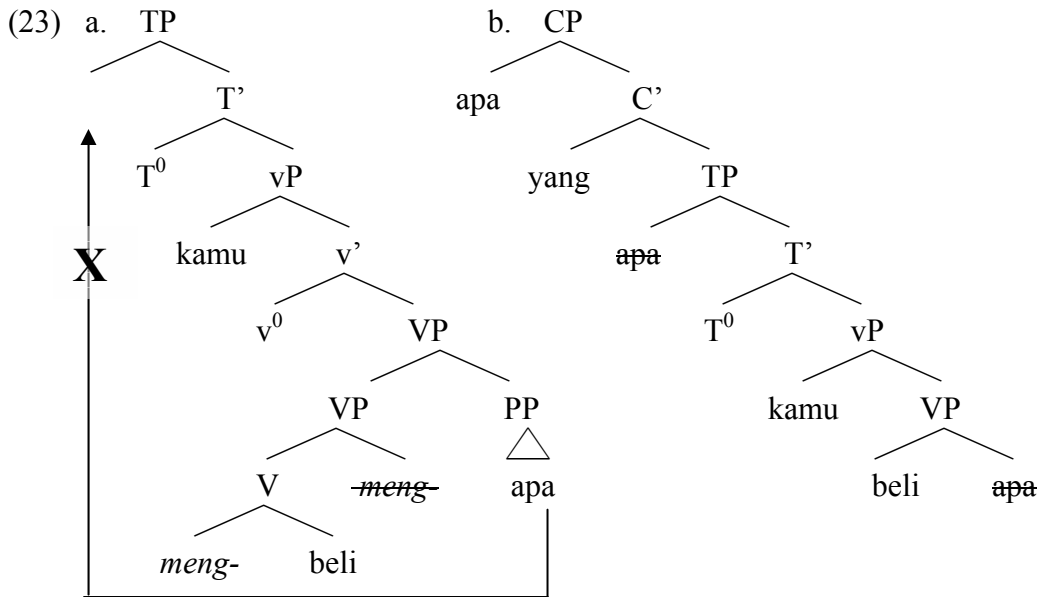
- (21) a. Kamu membeli apa?  
 2SG (*meng-*)buy what  
 ‘What did you buy?’

<sup>8</sup> For a more complete analysis of Indonesian clause structure, please see Fortin (in preparation).

Reconciling *meng-* and NP Movement in Indonesian

- b.  $Apa_i$  yang kamu (\*mem)beli  $t_i$ ?  
 what that 2SG (*meng-*)buy  
 ‘What did you buy?’
- (22) a. Kamu  $meng_i$ harapkan [<sub>CP</sub> Ali akan membeli apa]<sub>i</sub>?  
 2SG *meng-*hope Ali FUT *meng-*buy what  
 ‘What do you hope that Ali will buy?’
- b.  $Apa_j$  yang kamu (\* $meng_i$ )harapkan [<sub>CP</sub> Ali akan (\*mem<sub>j</sub>)beli  $t_j$ ]<sub>i</sub>?  
 what that 2SG (\**meng-*)hope Ali FUT (\**meng-*)buy  
 ‘What do you hope Ali will buy?’

In both monoclausal and multiclausal *wh*-questions, *meng-* first merges into the theta position of the internal argument of the verb. In the monoclausal question, the *wh*-word *apa* ‘what’ is a VP adjunct that is coreferent with *meng-* (23a); only if *meng-* is not present will *apa* be merged into an A-position, as the complement to the verb (23b). In the multiclausal question, the entire embedded CP *Ali akan beli apa* ‘Ali will buy what’ is an adjunct to the matrix VP (not shown). In both cases, extraction of *apa* from the adjunct to the matrix TP position (again, to satisfy the matrix  $T^0$ ’s EPP feature) is prohibited because it induces an island violation.



Cole and Hermon (1998) illustrate that extraction from an adjunct yields island effects in Singaporean Malay, and this is likewise true for the variety of Indonesian described in the present paper. As (24b) shows, extraction of *apa* from an adjunct (the clause headed by *karena* ‘because’) is prohibited, although leaving

*apa* in situ does not result in an island violation (24a).

- (24) a. Ali dipecat            karena dia beli apa?  
           Ali PASS-fire        because 3SG bought what  
           ‘Ali was fired because he bought what?’
- b. \* Apa        yang Ali dipecat        karena dia beli t<sub>i</sub>?  
           what        that Ali PASS-fire        because 3SG buy  
           ‘What was Ali fired because he bought?’

For similar reasons, *meng-* is prohibited in object relative clauses (25).

- (25) Buku yang Ali (\*mem)Baca itu menarik.  
       book that Ali (*meng-*)read this interesting  
       ‘The book that Ali read is interesting.’

For purposes of illustration, I will assume here a standard analysis of relative clauses for Indonesian (but see fn 8). If *meng-* is present in the relative clause, it is again first merged into the theta position of the object (for space limitations, this is not shown graphically). If this is the case, the relative clause operator cannot be merged into this position, but is instead an adjunct to the VP. The operator cannot extract from the adjunct to [Spec, CP] of the relative clause, which prevents the proper binding relationship between the operator and the variable from being established; as noted above, extraction from an adjunct in Indonesian gives rise to island effects and is thus prohibited. If *meng-* is not present in the relative, the operator is merged into the theta position and is able to extract to [Spec, CP], establishing the binding relationship between the operator and the variable.

### 3. Conclusion

In this paper, I have argued that Indonesian *meng-* is not simply a marker which signals active voice, agentivity, or transitivity, as is generally assumed (for references, see fn 3). Instead, I have argued that *meng-* is an indefinite object clitic pronoun that is unspecified for person and number. As *meng-* is first merged into the theta position of the verb’s internal argument, *meng-* has the effect of ‘antipassivizing’ the verb that hosts it. If there is an overt ‘object’ NP in the clause, the reference of *meng-* is determined by this NP, with which *meng-* is co-indexed. If *meng-* is present, however, an overt ‘object’ is not necessary; in this case, the object is interpreted as being indefinite, unknown, or unspecified, as in a ‘true’ antipassive.

Since *meng-* is merged into the theta position for the internal argument, an overt ‘object’ NP in the clause is in actuality a VP adjunct. The analysis of *meng-* suggested in the present paper uniquely accounts for the restrictions on the distribution of *meng-* in certain contexts: namely, *meng-* is prohibited from appearing in any clause within which the object NP has moved, such as bare

passives, object relative clauses, and *wh*-questions within which a *wh*-NP has moved ‘over’ the verb. This restriction on the appearance of *meng-* is not easily explained in an analysis where *meng-* simply marks agentivity or transitivity. In the present analysis, however, this restriction is easily accounted for: if *meng-* is present in the clause, any ‘object’ NP is actually an adjunct, and NP movement from out of the adjunct is prohibited. In the case of bare passives, NP movement from the adjunct to [Spec, T], an A-position, is ruled out by the Chain Condition. In the case of *wh*-movement and object relative clauses, island effects prohibit movement of (or from within) the adjunct *wh*-NP or the operator.

This analysis further accounts for the fact that *meng-* is optional in all environments in which it is licensed except for one: in case no object NP is overtly expressed. If this analysis is on the right track, it implies that no Indonesian verbs are ‘optionally’ transitive: in all cases, a transitive verb assigns its internal theta role to some element, either *meng-* or a lexical NP. If neither *meng-* nor a lexical NP are present to receive the internal theta role, a Theta Criterion violation results. Likewise, the fact that the presence of *meng-* reflects a semantic, not syntactic, transitivity is accounted for: *meng-* only appears on verbs that have an internal theta role to assign to it. As in a ‘true’ antipassive, *meng-* cannot represent the only argument of an unaccusative verb, which accounts for the fact that unaccusative verbs cannot host *meng-*.

## References

- Aldridge, Edith. 2006. Phase-Based Account of Extraction in Indonesian. Paper presented at ISMIL 10, University of Delaware, Newark.
- Baker, Mark. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.
- Chomsky, Noam. 1986. *Barriers*. Cambridge: MIT Press.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT Press.
- Chung, Sandra. 1976. On the Subject of Two Passives in Indonesian. In Charles Li, ed., *Subject and Topic*, 57-99. New York: Academic Press.
- Cole, P., and G. Hermon. 1998. The Typology of *Wh*-Movement: *Wh*-Questions in Malay. *Syntax* 1:221-258.
- Englebretson, Robert. 2003. *Searching for Structure: The Problem of Completion in Colloquial Indonesian Conversation*. Amsterdam: John Benjamins.
- Fortin, Catherine. In preparation. Indonesian Ellipsis: A Descriptive and Phase-Based Minimalist Analysis. Ph.D. diss., University of Michigan.
- Guilfoyle, E., H. Hung, and L. Travis. 1992. Spec of IP and Spec of VP: Two Subjects in Austronesian Languages. *Natural Language and Linguistic Theory* 10:375-414.
- Postman, Whitney Anne. 2002. Thematic Role Assignment in Indonesian: A Case Study of Agrammatic Aphasia. Ph.D. diss., Cornell University.

Catherine R. Fortin

- Saddy, Douglas. 1991. *Wh-Scope Mechanisms in Bahasa Indonesia*. *More Papers on Wh-Movement: MIT Working Papers* 15:183-218.
- Son, M., and P. Cole. 2004. Interface between Event Semantics and Syntax with Special Reference to the Verbal Suffix *-kan* in Standard Indonesian. Paper presented at GURT 2004, Georgetown University, Washington, DC.
- Sneddon, James Neil. 1996. *Indonesian: A Comprehensive Grammar*. London: Routledge.
- Voskuil, Jan E. 2000. Indonesian Voice and A-Bar Movement. In I. Paul, V. Phillips, and L. Travis, eds., *Formal Issues in Austronesian Linguistics*, 195-213. Dordrecht: Kluwer.

Catherine R. Fortin  
Department of Linguistics  
University of Michigan  
1190 Undergraduate Science Building  
204 Washtenaw Avenue  
Ann Arbor, MI 48109-2215

fortinc@umich.edu

## The Role of Animacy in Teiwa and Abui (Papuan)

MARIAN KLAMER and FRANTIŠEK KRATOCHVÍL  
*Leiden University*

### 0. Introduction

In this paper we discuss the role of animacy in the grammar of Teiwa and Abui, two Trans New Guinea languages spoken on two small islands just north of Timor in Eastern Indonesia. Teiwa has approximately 6,000 speakers on Pantar island, and Abui has approximately 16,000 speakers on Alor island. The paper is based on primary data collected on site in 2003-2005 (cf. Klammer, in prep.; Kratochvíl, in prep.).

While it is common to find pronominal affixes on verbs in Papuan languages, Teiwa and Abui represent the small minority of Papuan languages that only affix transitive objects (P), not subjects (A) (Foley 1986:102-104, Foley 2000:377). The alignment system of Teiwa is nominative-accusative, while Abui has an agentive alignment system.<sup>1</sup> In both languages, A is expressed as an independent NP in preverbal position. Ps always occur in between A and V, either as a lexical NP or as a verbal prefix, or as both. In this paper, we focus on the prefixation of P and show how it is influenced by the animacy value of the referent of P.

Animate entities are a salient conceptual category. In first language acquisition, the animacy of entities plays an important role: animate entities are the first to be noticed by young infants, the first objects to be individualized, and nouns referring to animate entities make up the majority of the first lexical items that children acquire (cf. Gentner 1982, Gentner and Boroditsky 2001).

Apart from playing a role in the acquisition of words, animacy also determines aspects of the grammatical system of languages. In the verbal domain, it may determine pronominal agreement or cross-reference patterns, as in those languages where animate objects are marked on the verb, while inanimate objects are not so marked. This results in grammatical systems that are grammatically ‘asymmetrical’ (Ortmann 1998), also referred to as ‘differential object marking’ (Bossong 1991, Aissen 2003). Examples of languages with asymmetrical systems

---

<sup>1</sup> Abui marks Undergoers (P and S of stative intransitives) with prefixes, in contrast to Actors (A and S of intransitive motion verbs), which are marked with free pronouns. An auxiliary verb construction may also be used to encode A (see Kratochvíl, in prep., for more information).

are Swahili, Palauan (Austronesian), Plains Cree (Algonquian) (Ortmann 1998:71-73), and the Papuan languages Usan (Reesink 1987:108-109) and Nggem (Etherington 2002). In this paper, we suggest that Teiwa and Abui are grammatically asymmetrical in a similar way.

In the nominal domain, animacy is one of the most important underlying principles for morphological gender and noun class systems. For example, animacy plays a role in the morphological gender systems of languages like Latin or Czech, and many Papuan languages divide nouns into subclasses of animates vs. inanimates and/or humans vs. non-humans (cf. Foley 2000:371-372 for examples). The languages discussed in this paper have neither morphological gender nor noun classes, but animacy plays an important role in the domain of nominal possession. Teiwa and Abui make a formal distinction between possessors that can be separated from the possessee and those that cannot (alienable vs. inalienable possession). The former are obligatorily marked; the latter are optional. Alienable possessors occur with common nouns, while kinship terms ('father', 'son') and body part nouns ('arm', 'leg') are inalienably possessed. We will suggest that the different marking of alienable vs. inalienable possessor goes back to its animacy. While common nouns may be used with or without a possessor (animate or inanimate), kinship and body part nouns only exist in relation to a possessor that is animate—a father is always a father to someone else; a leg always belongs to an animate entity. In other words, alienable possessors may be animate or not, but inalienable possessors are always animate, and this distinction is expressed in Teiwa and Abui by using different possessor marking patterns for each type.

In Teiwa and Abui, Ps and possessors are marked with homophonous prefixes. A more general aim of this paper is thus to introduce some languages that encode verbal objects and nominal possessors with prefixes from the same paradigm. This formal similarity contrasts with the more commonly observed one, where As and possessors are marked in the same way (e.g. Bittner and Hale 1996:60). The formal similarity between objects and possessors in Teiwa and Abui is not a coincidence, but depends on the semantic properties of these arguments as being animate or not. We will see that animacy overrides the two factors commonly considered to be the determiners of agreement. Neither the grammatical relation of P nor its thematic role determines its encoding on the verb; only its animacy value does.<sup>2</sup>

In section 1, we discuss the role of animacy in the marking of Teiwa objects and possessors, followed by a similar discussion of Abui in section 2. In section 3, we summarize the data and discuss the implications. Note that in describing the role of animacy, we focus on third-person pronominals since animacy is a relevant category for third-person referents only, first- and second-person referents being intrinsically animate.

---

<sup>2</sup> In section 2, we will see that in Abui P marking there are additional factors at work, but we will focus on the role of animacy.



**1. Animate Objects and Possessors in Teiwa**

In (1), the third-person singular pronominal forms to mark Ps and possessors in Teiwa are given.

(1) Teiwa 3 <sup>rd</sup> singular pronominals for P and Possessor			
<i>P</i>		<i>Possessor</i>	
Inanimate	Animate	Inalienable	Alienable
<i>ga'an</i>	<i>g(a)</i> - <sup>3</sup>	obligatory <i>g(a)</i> -	optional <i>g(a)</i> -

While all Teiwa Ps can be expressed as independent pronouns, only animate Ps may be prefixed. In (2a), the referent of *ga-* can only be interpreted as an animate entity, while the referent of the pronoun *ga'an* in (2b) is typically interpreted as inanimate (it allows an animate reading when it is used with emphatic stress).

- (2) a. A     *ga*-regan.  
       3SG  3-ask  
       ‘He asks him.’
- b. A     *ga'an*     regan.  
       3SG  he/her/it  ask  
       ‘He asks it.’ (or: ‘He asks HIM.’)

Animate referents also include animals. The concept of animacy is taken literally. In (3a), the definite, human object of ‘to bury’ is expressed with a pronoun, since a dead person is inanimate. This object cannot be marked with a prefix, as shown in (3b), which was rejected by consultants as having the “very strange” meaning that a living person was to be buried.

- (3) a. Na   ta   ma   *ga'an*   taraxa'.  
       1SG ASP come he/her/it bury  
       ‘I come to bury him.’
- b. \*Na   ta   ma   *ga-*     taraxa'.  
       1SG ASP come 3-       bury  
       not good for: ‘I come to bury him.’

Some transitive verbs typically have animate arguments, examples including the verbs in (4a); others typically have inanimate arguments, examples in (4b); and some occur as often with animates as with inanimates, as those in (4c).

---

<sup>3</sup> The brackets indicate that a prefix has two allomorphs: a syllabic form, which attaches to consonant-initial verbs/nouns, and a consonantal form which attaches to vowel-initial verbs/nouns.

- (4) a. *wei* ‘bathe him/her’  
*li’in* ‘invite him/her’  
*tewar* ‘walk with him/her/animal’  
*mis* ‘marry him/her’, ‘give birth to him/her’ (lit. ‘sit with/for him/her’)  
*fin* ‘catch him/her/animal’  
*tiar* ‘chase him/her/animal’  
*soi* ‘order him/her’  
*lal* ‘show to him/her’ (vs. *tub* ‘show, point it out’)
- b. *tutax* ‘cut it into very small pieces’  
*boqai* ‘cut it up in chunks’ (e.g. big fish, tree)  
*me’* ‘be located in/at; stay in/at’  
*parat* ‘tie it’  
*dau* ‘cook it’  
*ol* ‘buy it’  
*tub* ‘show it, point it out’ (vs. *lal* ‘show to someone’)
- c. *regan* ‘ask for him/her/it’  
*dee* ‘burn him/her/some animal/it’  
*mar* ‘follow/take him/her/some animal/it’  
*boxan* ‘guard him/her/some animal/it’  
*walas* ‘talk with him/her; talk about it, tell it’  
*er* ‘affect/make him/her/it’

Observe that the animate Ps of the verbs in (4a) have various semantic roles, including Patient, Benefactive, Recipient, Addressee, and Comitative. These roles overlap with those of the inanimate Ps of the verbs in (4b). Despite this overlap in semantic roles, a P is only prefixed when it has an animate referent (i.e. the verbs in (4a) take object prefixes; those in (4b) do not). Animacy is thus the relevant trigger for encoding an argument on the verb, and not its grammatical relation nor its thematic role.

The translations of some of the verbs in (4c) show that the animacy value of P can alter the interpretation of the verb. This is also illustrated in (5).

- (5) a. Na ga’an mar.  
 1SG 3 take  
 ‘I take/get it.’
- b. Na ga- mar.  
 1SG 3- follow  
 ‘I follow him/her.’

It should be remarked here that Teiwa has a tiny class of transitive verbs that always express P with a prefix. These verbs *also* distinguish Ps according to their animacy value by using different prefixes: a CVC prefix refers to animates, and a CV to inanimates. Examples include *wulul* ‘speak, talk, tell’ and *wultag* ‘talk’:

*The Role of Animacy in Teiwa and Abui (Papuan)*

- (6) a. *ga'*- *wulul* 'talk with/tell him/her'  
       *ga*- *wulul* 'talk about it, tell it'  
       b. *ga'*- *wultag* 'talk to/about him/her, tell him/her'  
       *ga*- *wultag* (or [gultag]) 'talk about it'

In sum, in Teiwa, verbal agreement does not mark objects as such, but is crucially determined by the animacy value of the object. The thematic role of the object is irrelevant, as long as it is not an agent.

Turning now from verbal to nominal agreement, let us consider Teiwa possessor marking. Teiwa body part nouns ('arm', 'leg', 'stomach') and kinship terms ('mother', 'son') have an inalienable possessor, and they can only occur with a possessor prefix. This is illustrated in (7a,b). Common nouns ('house', 'mountain', 'milk') have an alienable possessor, and as such, possessors are not obligatory; common nouns can occur in isolation without a possessor marker, illustrated in (7c,d).

- (7) a. *ga*- *xala'*                      b. \**xala'*  
       3- *mother*                        *mother*  
       'his/her/their mother'  
       c. *ga*- *yaf*                        d. *yaf*  
       3- *house*                        '(a) house/houses'  
       'his/her/their house'

As mentioned in the introduction, we suggest that there is a fundamental distinction between kinship terms and body part nouns on the one hand, and common nouns on the other: the former exist only in relation to a possessor that is animate, while the latter may have no possessor, or one that is inanimate. In Teiwa, this distinction is expressed morphosyntactically by using an obligatory vs. optional possessor prefix.

In sum, animacy plays a similar role in the verbal and nominal agreement of Teiwa: just as an animate P is expressed as a verbal core argument by an obligatory prefix on the verb, so is an animate possessor expressed as a nominal core argument by an obligatory prefix on the noun. The prefixes that are used are homophonous.

**2. Animate Objects and Possessors in Abui**

Abui has a group of transitive verbs that typically occur with an inanimate object, and do not require a P-marking prefix. Some examples are given in (8) (cf. the comparable Teiwa verbs in (4)).

- (8) *bang* 'carry it'  
       *telang* 'pull it'  
       *tadia* 'cut, slice it'

<i>tukong</i>	‘cut at it, cut in pieces’
<i>i</i>	‘put it’
<i>bai</i>	‘hit, grind it’
<i>takai</i>	‘steal it’
<i>meng</i>	‘wear it’
<i>mi</i>	‘take it’

Abui also has a class of verbs that can have an animate or inanimate object. These verbs do require a P prefix, and they mark the animate/inanimate character of the P by choosing a different prefix. Abui uses three sets of prefixes to encode P.<sup>4</sup> They are given in (9), which also presents the prefixes marking inalienable and alienable possession. Note that these are identical to two of the P-markers.

(9) Abui 3<sup>rd</sup> singular pronouns for P and Possessor

<i>P</i>			<i>Possessor</i>	
Patient	Other	Animate	Inalienable	Alienable
<i>ha-</i>	<i>he-</i>	<i>ho-</i>	<i>ha-</i>	<i>he-</i>

The first set of P-marking prefixes, glossed here as PAT(ient), marks prototypical Patients, entities that undergo a change of state or condition (cf. Dowty 1991:572-573, Van Valin and LaPolla 1997:85). This entity may be animate or inanimate. Sets two and three mark Ps that do not undergo a change of state. Set three is glossed ANIM(ate) since it typically marks animate Ps. Set two contrasts with the first PAT set because it marks Ps that do not undergo a change of state, and it contrasts with the third ANIM set because it typically (though not exclusively, see (12h)) encodes inanimates. To indicate that it marks neither typical patients nor typical animates, it is glossed here as OTHER.<sup>5</sup> In (10) the distinctions between the paradigms are represented.

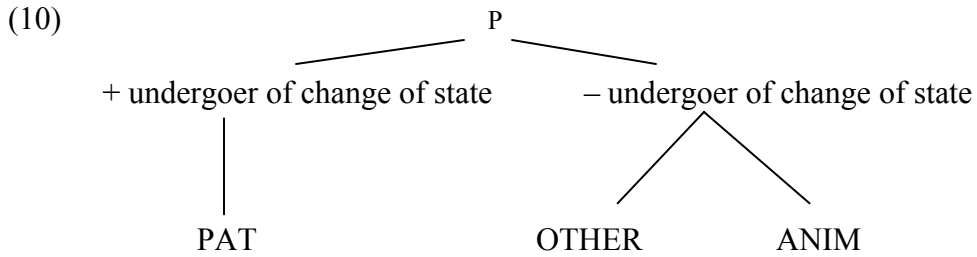
<sup>4</sup> The three sets probably derive from one original prefix, *h-*. Synchronically, Abui has a rich array of generic verbs that consist of a single consonant or vowel, including the verbs *a* ‘be at’, *e* ‘add’, and *o* ‘point’ (cf. Kratochvíl, in prep.). Originally, *h-* may have been used to mark Ps and Possessors, fusing over time with the generic verbs it often attaches to. As a result of reanalyzing the morpheme boundaries, as in (i) below, there are now three distinct prefixes, each with their own semantics.

- (i) a. *h-a fanga* > 3-BE.AT tell > *ha-fanga* ‘tell/order him’  
 b. *h-e fanga* > 3-ADD tell > *he-fanga* ‘say it’  
 c. *h-o fanga* > 3-POINT tell > *ho-fanga* ‘scold him’

One function of generic verbs in Abui is to introduce additional arguments into a clause. In the development of *ha-*, *he-*, and *ho-*, the semantics of the generic verb may have fused with the referential properties of the pronominal prefix. In other Trans New Guinea languages such as Dani (Lower Grand Valley, Bromley 1981) and Ekagi (Drabbe 1952), cognate forms of similar generic verbs have also fused with the pronominal prefix.

<sup>5</sup> This label is used for expository reasons here; Kratochvíl (in prep.) employs a different label.

*The Role of Animacy in Teiwa and Abui (Papuan)*



The verb *fanga* ‘say’ combines with all three prefixes. In (11a), the order is given and carried out by the patient Simon, who thus undergoes a change of state. In (11b), the prefix refers to the inanimate complement of ‘say’. In (11c), the prefix has an animate referent, the addressee or malefactive participant of the scolding.

- (11) a. He-      maama      Simon      **ha-**      fanga.  
          3.AL-      father      S.      3.PAT      say  
          ‘His father ordered Simon.’
- b. Ama      he-      kang      **he-**      fanga.  
          person 3.OTHER can      3.OTHER say  
          ‘People agree/approve.’ (lit. ‘Persons say it can.’)
- c. A      neng      loku      **ho-**      fanga.  
          you.SG man      PL      3.HUM      say  
          ‘You scold at the men.’

More examples illustrating the contrasts are given in (12).

- |         |   |   |   |
|---------|---|---|---|
| (12) a. | ha-fanga<br>3.PAT-tell/say<br>‘order him’ | he-fanga<br>3.OTHER-tell/say<br>‘say it’      | ho-fanga<br>3.ANIM-tell/say<br>‘scold (at) him’ |
| b.      | ha-lia<br>3.PAT-fly<br>‘shoot it’         | he-lia<br>3.OTHER-fly<br>‘fly at it’          | ho-lia<br>3.ANIM-fly<br>‘fly at him’            |
| c.      | *ha-faaling<br>3.PAT-listen               | he-faling<br>3.OTHER-listen<br>‘listen to it’ | ho-faling<br>3.ANIM-listen<br>‘listen to him’   |
| d.      | *ha-fahak<br>3.PAT-hug                    | he-fahak<br>3.OTHER-hug<br>‘embrace it’       | ho-fahak<br>3.ANIM-hug<br>‘hug him’             |

e.	ha-dik 3.PAT-prick 'pierce it through'	he-dik 3.OTHER-prick 'stab (at) it'	ho-dik 3.ANIM-prick 'prick/tickle him'
f.	ha-tang 3.PAT-release 'release him/it'	he-tang 3.OTHER-release 'transfer/pass it along'	*ho-tang 3.ANIM-release
g.	*ha-lel 3.PAT-impend	he-lel 3.OTHER-impend 'almost do it'	ho-lel 3.ANIM-impend 'threaten him'
h.	*ha-kafia 3.PAT-scratch	he-kafia 3.OTHER-scratch 'scratch for/instead of him'	ho-kafia 3.ANIM-scratch 'scratch him'

On the basis of these examples, the following observations can be made.

- (i) The distribution of the prefixes cannot be determined by looking at the thematic role of the argument because the roles of *he-* in the middle column overlap (to some extent) with the roles of *ha-* and *ho-*. The thematic role of *ha-* is patient; the thematic role of *he-* is theme in (a,c,f), goal in (b), location in (d), patient (without change of state) in (e), and benefactive in (h). The thematic roles referred to by *ho-* include malefactive, benefactive, or goal; all of the referents are animate.
- (ii) The *ha-* form refers to a real patient in the sense that it must undergo a change of state, while the *he-* form does not. This is clear from the ungrammatical forms in (c,d,h): the P of these verbs does not undergo a change of state, hence no *ha-* form is possible. In (e), the referent of *ha-* is the most affected P; it really underwent a change of state, while the referent of *he-* did not.
- (iii) Some forms are not allowed because they would denote semantically unusual concepts. For example, in (f) no *ho-* form is allowed because 'to release on someone' is semantically strange. In (g) no *ha-* form is allowed because an event that is 'about to happen' by definition does not have a Patient that undergoes a change of state because nothing has happened to P yet.
- (iv) *Ho-* only refers to animate Ps. However, not all animate Ps are marked by *ho-*: they are encoded by *ha-* when they undergo a change of state, as in (a,f), or when they are benefactive, as in (h). Thus, animacy is a crucial feature of the referents of *ho-*, but it does not play a role in the choice of

*The Role of Animacy in Teiwa and Abui (Papuan)*

*ha-*, while the choice of *he-* vs. *ho-* is determined by animacy to some extent.

Further evidence for the analysis of *ho-* as referring only to animates comes from the marking of the animate S of intransitive verbs that denote states or experiences. In (13a), the S is the animate experiencer of *lil* ‘hot’ and must be marked with *ho-*. In contrast, an inanimate S of the same verb is expressed as a lexical NP, as shown in (13b).

- (13) a. Ho-            lil-        a.  
           3.ANIM-    hot        be.at  
           ‘S/he feels hot.’
- b. Kopi         do        lil-        a.            \***ho-**/\***he-**/\***ha-**lila  
           coffee     DET     hot        be.at  
           ‘This coffee is hot.’

Note that the *he-* prefix may be used with *lil* when it refers to a possessor, as in (14a), and a *ha-* prefix refers to the patient of the derived construction ‘to give heat’ > ‘to heat up’, as in (14b).

- (14) a. He-            lil        -a  
           3            hot        be.at  
           ‘his warmth/blessing/knowledge of life’ (referent of *he-* is possessor)
- b. Ha-            lil        -r        -a.  
           3.PAT        hot        give     be.at  
           ‘Heat it up.’ (lit. ‘Give it heat.’) (referent of *ha-* is patient)

In the nominal domain, Abui distinguishes between alienably and inalienably possessed nouns. The inalienable nouns have an obligatory possessor. Body parts mark their possessor with *ha-*; kin terms mark it with *he-*. This is illustrated in (15a,b). The possessor prefix of alienable nouns is also *he-*, but unlike for kin terms, it is optional for common nouns. This is shown in (15c,d). Nominal attributive constructions as in (15e) may have a possessor interpretation, but do not have a possessor prefix.<sup>6</sup>

---

<sup>6</sup> In Teiwa and Abui, certain locations may be expressed with a nominal possessive construction involving a possessed body part noun (e.g. Abui: *fala ha-po* ‘house 3-forehead’ > ‘in front of the house’). In such constructions, the prefix obviously has an inanimate possessor referent. We see this as metaphorical extensions of the original possessor constructions with body part nouns, which always have animate possessors.

(15)	a.	<i>ha-to</i>	‘his penis’	<i>*he-to, *ho-to, *to</i>
		<i>ha-ne</i>	‘his/her name’	<i>*he-ne, *ho-ne, *ne</i>
		<i>ha-muk</i>	‘its horn’	<i>*he-muk, *ho-muk, *muk</i>
	b.	<i>he-wil</i>	‘his/her child’	<i>*ha-wil, *ho-wil, *wil</i>
		<i>he-hai</i>	‘his/her wife’	<i>*ha-hai, *ho-hai, *hai</i>
	c.	<i>he-fala</i>	‘his/her house’	<i>*ha-fala, *ho-fala</i>
		<i>he-konrek</i>	‘his/her shirt’	<i>*ha-konrek, *ho-konrek</i>
	d.	<i>fala</i>	‘house’	
		<i>konrek</i>	‘shirt’	
	e.	<i>aremang fala</i>		
		tribe house	‘tribal house’	

Abui possessor marking prefixes are thus homophonous with the prefixes that mark Ps. Note, however, that their function in the verbal and nominal domain is not parallel.

(16)		<i>Verbal Domain</i>	<i>Nominal Domain</i>
	<i>ha-</i>	patient	inalienable poss. of body parts: obligatory
	<i>he-</i>	various roles	inalienable poss. of kinship: obligatory
			alienable poss.: optional
	<i>ho-</i>	animates	<b>not used</b>

To conclude, inalienable nouns are inherently possessed by an animate entity, and such possessors are expressed as prefixes to the nouns. Similarly, animate Ps must be expressed as the core argument of Abui verbs.

### 3. Summary and Discussion

In Teiwa and Abui, the animacy of referents plays a crucial role in verbal and nominal agreement patterns. While verbal prefixes in Teiwa and Abui mark objects, not all objects are prefixed, so we cannot characterize the prefixing by referring to the grammatical role of the argument, nor does the semantic role of the argument play a determining role: non-agent argument of various kinds are prefixed. The traditional way to characterize agreement in terms of grammatical role or semantic properties (or a combination of these) is not adequate to describe the patterns found in Teiwa and Abui.

In both Teiwa and Abui, a structural relation exists between the object of verbs and the possessor of nouns: both arguments are expressed by identical prefixes, and when the referents are animate, the agreement is obligatory. Given the salient role that the animacy of entities plays in perception and cognition, it is not surprising that we find patterns like these where the animacy of a referent is the feature triggering verbal and nominal agreement.

In addition, similar connections have been noted in other languages. For example, polysemous possessive and benefactive morphemes are found in both Austronesian (Oceanic) as well as Papuan languages and are discussed in



Lichtenberk (2002). In Abui and Teiwa, the parallel goes even further, since in these languages not only benefactives, but *all* kinds of Ps are polysemous with possessors.<sup>7</sup>

Finally, we would like to point out that animacy also plays a role in the agreement patterns in other grammatical domains of Papuan languages—for example, in the marking of the (single) object of verbs translated as ‘give’. Many Papuan languages lack ditransitives altogether; many have at most one or two ditransitives (Foley 2000:377).<sup>8</sup> In languages which have verbs for concepts like ‘give’, the object marker usually marks the recipient or benefactive, not the patient (Foley 2000:378).<sup>9</sup> If we assume that animacy is the trigger for the marking of P, a pattern like this is expected: a benefactive or recipient (the person given to) is by definition a human and thus marked on the verb more systematically and frequently than the patient (the thing given).

Another feature that is often mentioned as typical for Papuan languages is the existence of so-called ‘experiential’ constructions (Reesink 2002:27). In such constructions the experiencer of an uncontrolled state verb is marked by a regular object affix. This experiencer may be part of an intransitive construction (‘me hungers’), or a transitive construction (e.g. ‘it hungers me’, ‘hunger does me’). In these cases the experiencer, an animate non-agent, is marked like P, like the other non-agent animate arguments.<sup>10</sup> Constructions like these exist in many Papuan languages and are additional indications that the animacy value of referents can play a pervasive role in shaping the agreement patterns of languages.<sup>11</sup>

## References

- Aissen, Judith. 2003. Differential Object Marking: Iconicity vs. Economy. *Natural Language and Linguistic Theory* 17:673-711.
- Bittner, Maria, and Ken Hale. 1996. The Structural Determination of Case and Agreement. *Linguistic Inquiry* 27(1):1-68.
- Bossong, Georg. 1991. Differential Object Marking in Romance and Beyond. In Douglas Kibbee and Dieter Wanner, eds., *New Analyses in Roman Linguistics*, 143-170. Amsterdam: John Benjamins.

---

<sup>7</sup> A Papuan language with similar patterns is Dani (spoken in the Balim Valley in the central part of West Papua), which also cross-references possessors and undergoers with identical forms. In Dani, forms with the vowel *o* typically mark the inalienably possessed kinship terms and body parts (Bromley 1981:190), while alienably possessed nouns are marked with forms containing an *a* root.

<sup>8</sup> For example, Teiwa has only one verb that optionally occurs with two objects in a ditransitive construction; all the other transitives are monotransitive.

<sup>9</sup> Often, the Patient occurs with its own predicate in a serial verb construction.

<sup>10</sup> In transitive constructions, the inanimate instigator of the experience is encoded like A.

<sup>11</sup> There may be other areas in Papuan languages where the animacy of a participant plays a crucial role. For example, it may determine the grammaticalization paths of certain verbs, as argued in Klamer (to appear).

- Bromley, H. Myron. 1981. *A Grammar of Lower Grand Valley Dani*. Canberra: Research School of Pacific Studies, Australian National University.
- Dowty, D. 1991. Thematic Proto-Roles and Argument Selection. *Language* 67(3):547-619.
- Etherington, Paul A. 2002. *Nggem Morphology and Syntax*. M.A. thesis, The Northern Territory University, Australia.
- Foley, William A. 2000. The Languages of New Guinea. *Annual Review of Anthropology* 29:357-404.
- Gentner, Dedre, and Lera Boroditsky. 2001. Individuation, Relativity, and Early Word Learning. In Melissa Bowerman and Stephen C. Levinson, eds., *Language Acquisition and Conceptual Development*, 215-256. Cambridge: Cambridge University Press.
- Klamer, Marian. In preparation. *A Grammar of Teiwa*.
- Klamer, Marian. To appear. Animacy and the Grammaticalisation of 'come' in Teiwa and Kaera. In Michael Ewing and Marian Klamer, eds., *Typological and Areal Analyses: Contributions from East Nusantara*. Leiden: KITLV Press.
- Kratochvíl, František. In preparation. *A Grammar of Abui*. Ph.D. diss., Leiden University.
- Lichtenberk, Frantisek. 2002. The Possessive-Benefactive Connection. *Oceanic Linguistics* 41(2):439-474.
- Ortmann, Albert. 1998. The Role of [+/-animate] in Inflection. In Ray Fabri, Albert Ortmann, and Teresa Parodi, eds., *Models of Inflection*, 60-84. Tübingen: M. Niemeyer.
- Reesink, Ger. 1987. *Structures and Their Functions in Usan, a Papuan Language of Papua New Guinea*. Amsterdam: John Benjamins.
- Reesink, Ger P. 2002. *Languages of the Eastern Bird's Head*. Canberra: Pacific Linguistics.
- Van Valin, Robert D., and Randy J. LaPolla. 1997. *Syntax: Structure, Meaning and Function*. Cambridge: Cambridge University Press.

Marian Klamer  
Leiden University  
Languages and Cultures of  
SE Asia and Oceania  
The Alor and Pantar Project  
(AAPP)  
Faculty of Arts  
P.O. Box 9515  
2300 RA Leiden  
The Netherlands

marian.klamer@let.leidenuniv.nl

František Kratochvíl  
Leiden University  
Languages and Cultures of  
SE Asia and Oceania  
The Alor and Pantar Project  
(AAPP)  
Faculty of Arts  
P.O. Box 9515  
2300 RA Leiden  
The Netherlands

f.kratochvil@let.leidenuniv.nl

# A Feature Geometry of the Tongan Possessive Paradigm

CATHERINE MACDONALD  
*University of Toronto*<sup>1</sup>

## 0. Introduction

Feature geometries adapted from Harley (1994) and others can be developed to account for the semantic distinctions in Tongan's massive possessive paradigm. These pronouns represent three distinct syntactic categories, and I propose a typology of these pronouns adapted from that proposed by Dechaine and Wilt-schko (2002), in which each feature geometry is subsumed under a different syntactic head.

## 1. The Tongan Pronominal Paradigms

Tongan personal and possessive pronouns distinguish four persons and three numbers. In the possessive paradigm, the person and number features are those of the possessor, not the possessum. In addition, the possessive pronominal forms encode the definiteness of the possessum, speaker's sympathy, and the relationship between possessor and possessum. Sympathy is encoded through the incorporation of a diminutive marker. The relationship between possessor and possessum is marked by an incorporated genitive case-marker: *'a* or *'o*.

### 1.1. Typology of Personal and Possessive Pronouns

There are two types of personal and three types of possessive pronouns in Tongan. Strong personal pronouns function like ordinary DPs. Clitic personal pronouns cross-reference or replace a strong pronoun which is a transitive or intransitive subject (A or O); they occur adjacent to the tense-aspect marker. "Ordinary" possessive pronouns precede the possessum, require an NP complement, and encode all distinctions described above. Postposed (adjectival) possessive pronouns follow the possessum and do not encode definiteness or sympathy. Emphatic possessive pronouns occur alone (instantiating the possessor and the possessum) or precede the possessum; they are always specific and definite, but

---

<sup>1</sup> I wish to thank Diane Massam, my supervisor, and Elizabeth Cowper, my instructor. Both have provided helpful criticism and encouragement. The paper has been changed from the version presented at BLS 2006, taking into account comments of other delegates.

never sympathetic. The paradigms of these pronouns are presented in the appendix.

## 1.2. Morphology of Possessive Pronouns

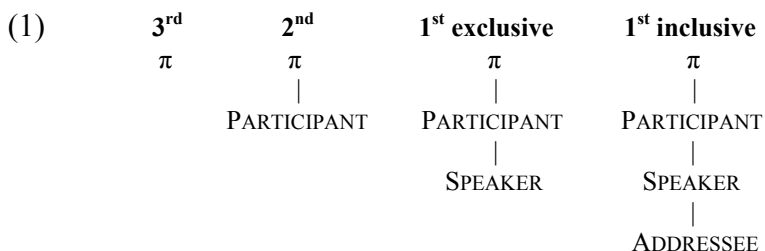
The morphology of Tongan possessive forms is strikingly transparent. Each form comes historically from the merger of a personal pronoun encoding the person and number of the possessor, a possessive case marker (*'a* or *'o*), a determiner (*ha* nonspecific or *he* specific), and optionally, a diminutive particle *si'i* ('small') denoting sympathy. The weak personal pronoun forms the basis of ordinary and postposed pronouns; the strong one forms the basis of emphatic pronouns.

## 2. Feature Geometries

### 2.1. Person

Personal and possessive pronouns in Tongan encode three persons: first inclusive, first exclusive, second, and third. Several proposals have been made for a geometric arrangement of person features in pronouns. Harley's (1994) proposal has been variously adapted by Harley and Ritter (2002) (henceforth, H&R), Cowper and Hall (2005), and McGinnis (2004). I adopt the strict hierarchical structure of Harley's (1994) model but H&R's features.

In this analysis, first and second persons (discourse participants) are distinguished from other pronominal referents by a feature PARTICIPANT. Among discourse participants, parties including the speaker (first person) are distinguished from those not including the speaker (second person) by the presence of SPEAKER. An inclusive/exclusive distinction is made by the presence of ADDRESSEE under SPEAKER. This feature is not present in the second person, even though its referent is an addressee, because the feature is available only when contrastive, i.e. when it distinguishes one possible referent from another. The geometry is given in (1).



### 2.2. Number

Personal and possessive pronouns in Tongan distinguish singular, dual, and plural. The feature hierarchy proposed by Harley (1994) has been adapted by H&R, Cowper (2003, 2005), and McGinnis (2004). Again, I adopt the geometry of Harley (1994) with H&R's features. INDIVIDUATION distinguishes between mass and countable entities. Its dependent, PLURAL, distinguishes non-singular from singular referents. In a three-way number system, PLURAL has a dependent,

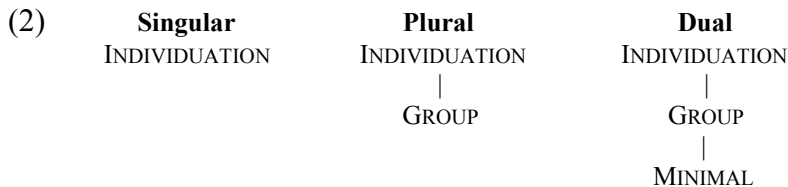
## *Feature Geometry of Tongan Personal Pronouns*

MINIMAL, which distinguishes “minimal groups” (i.e. groups of two) from others.

Underlying this proposal is the assumption that dual is more marked than plural, based on Greenberg’s (1963) observation that any language with a dual also has a plural. Dryer (2005) argues that appeals to this observation are vacuous; no language could logically make a singular-dual distinction without also making singular-plural and dual-plural distinctions. Rice (1999) notes that crosslinguistic markedness may not accurately diagnose featural markedness, noting that in phonology, central vowels are infrequent among the inventories of the world’s languages, but minimally marked where they occur. Cowper (2003, 2005) argues that instead of a feature marking groups of two as minimal, there is one marking larger groups as “> 2”, and thus dual is more “marked” than plural.

Cowper (2003) argues that syncretisms should provide evidence: singular-dual syncretisms will support her model, and singular-plural syncretisms will support that of Harley (1994). In Tongan, due to the paucity of agreement morphology, there is little such evidence, and it is unhelpful. Churchward (1953) lists ten nouns with marked dual and/or plural forms: five display dual-plural syncretisms; three singular-dual syncretisms; one a singular-dual syncretism; and one has singular and plural forms, either of which can be used for dual.

An alternative source of evidence is the pronominal dual and plural markers, which seem to suggest that dual is more marked than plural. Nevins (2006:2), citing Jakobson and Greenberg, notes that “if a certain category is marked, then one will find fewer oppositions for other categories within it”; Tongan has one dual and seven plural markers. I thus adopt the geometry in (2).



### **2.3. Specificity and Diminutivity**

The morphological transparency of the Tongan possessive pronouns reveals the presence of the incorporated determiners *he* (‘specific’) or *ha* (‘nonspecific’) and their respective “emotional” counterparts *si’i* and *si’a*. These four determiners arise from the four possible combinations of SPECIFIC and DIMINUTIVE.

Although *he* and *ha* are usually glossed as ‘a’ and ‘the’, respectively, Dukes (2006) argues that ‘specific’ and ‘nonspecific’ are more accurate. Likewise, Churchward (1953), who defines *he* as ‘semi-definite’ and *ha* as ‘completely indefinite’, notes that the latter can be used with things whose existence is questioned or denied (cf. his (7c,d)). Hendrick (2005) points out that *ha* and *he* are equivalent to the two senses of English *a* in the *de dicto* and *de re* readings of a sentence like *Kim is looking for a newt in the garden*, and that the equivalent of *the* is *he* plus the definitive accent, a stress shift targeting the final word in the definite DP (Churchward 1953, Dukes 1996, Hendrick 2005). It is only available

with *he* or a specific possessive pronoun, not with *ha*. Specific and nonspecific possessives are shown in (3).

- (3) a. he'eku                      helé  
       SPEC-POSS.A-1EX.SG knife  
       'the-my knife (the knife which is mine)'
- b. ha'aku                      hele  
       NONSPEC-POSS.A-1EX.SG knife  
       'a-my knife (a knife which is mine)'
- c. 'oku 'ikai ha'aku                      tohi.  
       PRES NEG NONSPEC-POSS.A-1EX book  
       'I haven't a book.' (lit. 'There is not a-my book.')
- (Churchward 1953:130)

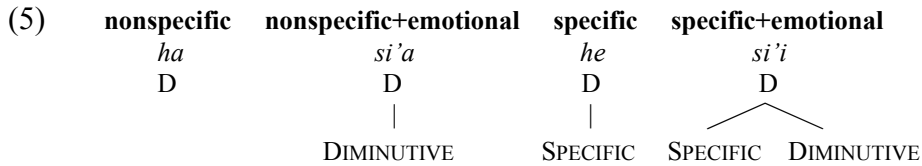
Clark (1974) argues that the definitive accent is the reflex of a demonstrative, \*-*a*. This clitic, having lost its phonological specification, is now a null mora affixed to the end of a DP, lengthening the final vowel and causing primary stress to be reassigned to it. In addition to its phonological specification, this clitic has lost meaning; it is now a “purely referential marker of definiteness” (Clark 1974:107).

Cowper and Hall (2003) propose the following geometry for determiner features: a D with no dependents is “nonspecific” (and indefinite). D may have a dependent SPECIFIC; it, in turn, may have a dependent DEFINITE. Tongan determiners do not mark definiteness, so DEFINITE is not a D-feature in this language; it is the interpretation of a neutral demonstrative in Dem<sup>0</sup>. I will return to this with the syntax of emphatic possessive pronouns in §3.

Emotional import, encoded in “ordinary” possessive pronouns and determiners, is realized by an incorporated diminutive marker *si'i* ('small'). Its presence indicates the speaker's “feelings of affection, friendship, pity, humility, or respect” (Churchward 1953:23). Illustrative examples are presented in (4). I propose that the “emotional” feature is DIMINUTIVE, a dependent of D. Since this is a sister to SPECIFIC, these features are independent, allowing all four possible combinations. The proposal is presented in (5).

- (4) a. kuo lavea **si'a**                      tamasi'i?  
       PERF hurt NONSPEC.EMOT child  
       'Has a (dim.) child been hurt?'                      (Churchward 1953:23)
- b. 'oku nofo masiva **si'eku**                      mātu'á.  
       PRES very poor SPEC.EMOT-POSS.A-1EX.SG parent  
       'My parents are very poor.'                      (Shumway 1971:513)

*Feature Geometry of Tongan Personal Pronouns*



**2.5. Subjective vs. Objective Possession**

Tongan encodes two types of possession with the case-markers ‘*a* and ‘*o*. It has been claimed that the choice of ‘*a* or ‘*o* has become rigidly lexicalized for most possessa, but Taumoefolau (1996) claims that while there may be a conventional preference for ‘*a* or ‘*o* with each noun, this can be overruled. Wilson (1982) argues that ‘*a* indicates possessor “control over the initiation of the possessive relationship”, and ‘*o* marks default possession. Others characterize the distinction as alienable/inalienable (Otsuka 2000) or subjective/objective (Churchward 1953). Moysse-Faurie (2000) argues that many factors contribute to the choice. It is difficult to know how to treat the feature(s) distinguishing ‘*a* and ‘*o* in Tongan.

It also needs to be determined which of ‘*a* or ‘*o* is marked. Although Wilson (1982) proposes that ‘*a* is marked, Clark (2000) notes that where languages have lost the ‘*a*/‘*o* distinction, ‘*a* has remained. This and the fact that ‘*o* only marks objects in nominalizations while ‘*a* marks both transitive and intransitive subjects suggest that ‘*o* is marked. I propose a feature OBJECT dependent on a possessive head between the determiner and the possessor.

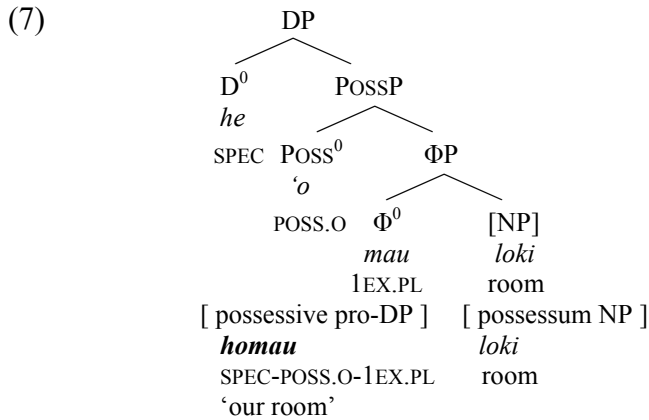
**3. Typology of Tongan Pronouns**

Dechaine and Wiltschko (2002) (henceforth D&W) propose three types of pronominals: pro-DPs, pro-φPs, and pro-NPs. Expanding their typology, I propose that the Tongan possessive pronouns are pro-DemPs (“emphatic”), pro-D<sup>0</sup>s (“ordinary”), and pro-PossPs (“postposed”). Each realizes a different syntactic structure; the terminal nodes host its morphosyntactic features.

**3.1. “Ordinary” Possessive Pronouns**

“Ordinary” possessive pronouns in Tongan consist of a determiner, a possessive case-marker, and a pronominal root encoding the φ-features of the possessor. They are pro-D<sup>0</sup>s, taking the possessum NP as complement. Evidence that the pronoun and its complement NP constitute a DP comes from their distribution: they function as arguments (6), not as predicates. I propose the structure in (7).

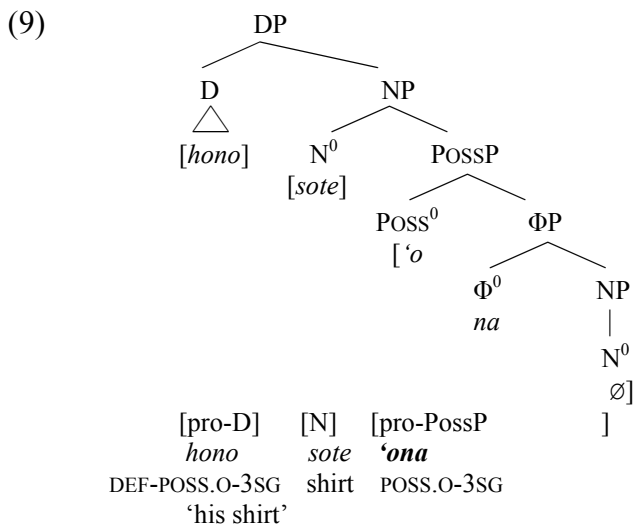
- (6) ‘oku ‘uli homau loki.  
 PRES dirty SPEC-POSS.O-1EX.PL room  
 ‘Our room is dirty.’ (Shumway 1971:275)



### 3.2. Postposed Possessive Pronouns

Postposed possessive pronouns are POSSP modifiers of NP. They take no NP complement, but have a nominal component— $\emptyset$  in the singular, *ua* ('two') in the dual, and *tolu* ('three') in the plural—which stands for the possessor. Morphologically, they resemble pro-D<sup>0</sup>s without a determiner; they do not encode definiteness or sympathy. Syntactically, they behave like adjectives (8). They are predicative (8a) and attributive (8b): when attributive, they are postnominal; when predicative, they may be nominalized. Their structure is given in (9).

- (8) a. 'oku 'ana eni.  
 PRES POSS.A-3SG this  
 'This is his.' (Churchward 1953:135)
- b. 'oku 'ikai ke u sai'ia he'ene founa 'ana.  
 PRES NEG C 1EX.SG like SPEC-POSS.A-3SG method POSS.A-3SG  
 'I don't like his particular method.' (Shumway 1971:434)

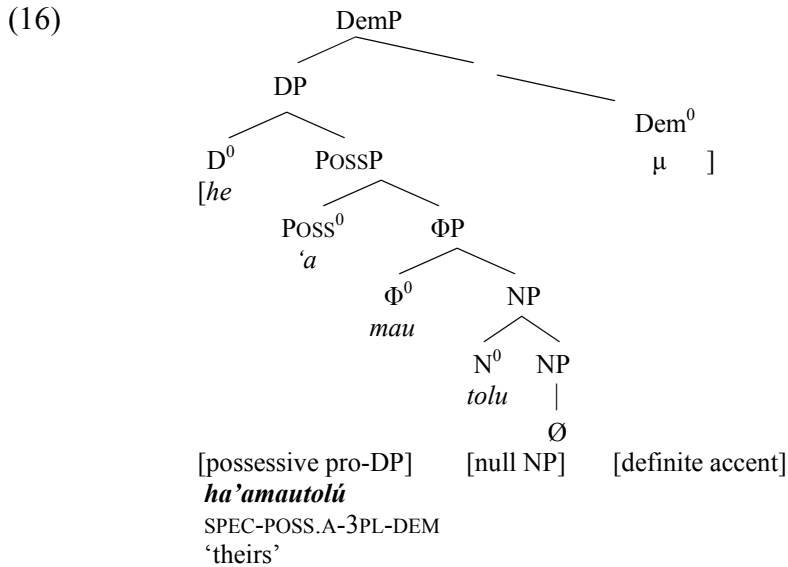




### 3.3. Emphatic Possessive Pronouns

Emphatic possessive pronouns are pro-DemPs, fully saturated with a null NP representing the possessum. Evidence for this saturation comes from their ability to occur as an argument (10) and the requirement that they bear the “definitive accent” which targets the right edge of DPs. The fact that they are always specific and never emotional, I propose, has to do with the selectional restrictions of Dem<sup>0</sup>. I propose the structure in (16).

- (15) ‘omi ha’akú.  
 bring POSS.SUBJ.1EX.SG  
 ‘Bring mine.’ (Churchward 1953:135)



### 4. Conclusions

Tongan’s rich possessive paradigm can be accounted for with a feature geometry adapted from those of Harley (1994) and Cowper and Hall (2003), to which I add DIMINUTIVE and OBJECT. These feature hierarchies are subsumed under syntactic terminal nodes. I adopt D&W’s proposal that “pronouns” represent a range of syntactic structures, adding pro-DemPs, pro-D<sup>0</sup>s, and pro-PossPs to their inventory. Evidence for these structures comes from the morphology of Tongan possessive pronouns, as well as from the features available for each pronominal series and their syntactic behaviors.

### References

- Churchward, C. Maxwell. 1953. *Tongan Grammar*. Nuku’alofa: Vava’u Press.  
 Clark, Ross. 1974. On the Origin of the Tongan Definitive Accent. *Journal of the*

- Polynesian Society* 83:103-108.
- Clark, Ross. 2000. Possessive Markers in Polynesian Languages. In Steven R. Fischer, ed., *Possessive Markers in Central Pacific Languages: Sprachtypologie und Universalienforschung* 53(3/4):258-268.
- Cowper, Elizabeth. 2003. Why Dual Is Less Marked than Plural. Ms., University of Toronto. <http://www.chass.utoronto.ca/~cowper/Whydual.pdf>.
- Cowper, Elizabeth. 2005. A Note on Number. *Linguistic Inquiry* 36(3):441-455.
- Cowper, Elizabeth, and Daniel Currie Hall. 2003. The Syntactic Manifestation of Nominal Feature Geometry. *Proceedings of Canadian Linguistics Association*. Montreal: Cahiers Linguistiques de l'UQAM.
- Cowper, Elizabeth, and Daniel Currie Hall. 2005. The Pieces of Pi. In Marie-Odile Junker, Martha McGinnis, and Yves Roberge, eds., *Proceedings of the 2004 Canadian Linguistics Association Annual Conference*.
- Dechaine, Rosemarie, and Martina Wiltschko. 2002. Decomposing Pronouns. *Linguistic Inquiry* 33(3):409-442.
- Dryer, Matthew. 2005. Vacuous Appeals to the Number Hierarchy. Paper presented at the 6<sup>th</sup> Biennial Meeting of ALT, Padang, Indonesia.
- Dukes, Michael. 1996. On the Nonexistence of Pronominals and Anaphors in Tongan. Ph.D. diss., University of California at Los Angeles.
- Fischer, Steven R. 2000a. Introduction. In Steven R. Fischer, ed., *Possessive Markers in Central Pacific Languages: Sprachtypologie und Universalienforschung* 53(3/4):229-250.
- Harley, Heidi. 1994. Hug a Tree: Deriving the Morphosyntactic Feature Hierarchy. *MIT Working Papers in Linguistics* 21:289-320.
- Harley, Heidi, and Elizabeth Ritter. 2002. Person and Number in Pronouns: A Feature-Geometric Analysis. *Language* 78(3):482-525.
- McGinnis, Martha. 2004. On Markedness Asymmetries in Person and Number. *Language* 81(3):699-718.
- Nevins, Andrew. 2006. Dual Is Still More Marked than Plural. Syracuse University Colloquium Series. <http://ling.auf.net/lingBuzz000235>.
- Otsuka, Yuko. 2000. Ergativity in Tongan. Ph.D. diss., Oxford University.
- Rice, Keren. 1999. Featural Markedness in Phonology: Variation – Part 1. *GLOT International* 4(7):3-6.
- Shumway, Eric B. 1971. *Intensive Course in Tongan*. Honolulu: University of Hawai'i Press.
- Taumoefolau, Melenaitē. 1996. Nominal Possessive Classification in Tongan. In John Lynch and Fa'afo Pat, eds., *Oceanic Studies: Proceedings of the First ICOL*. Pacific Linguistics Series C, No. 133: 293-304.
- Wilson, William H. 1982. *Proto-Polynesian Possessive Marking*. Pacific Linguistics Series B, No. 85. Canberra: Australian National University.

*Feature Geometry of Tongan Personal Pronouns*

Catherine Macdonald  
 Department of Linguistics  
 University of Toronto  
 130 St. George Street, Room 6076  
 Toronto, Ontario M5S 3H1  
 Canada

cath.macdonald@utoronto.ca

**Appendix: Paradigms of Tongan Personal and Possessive Pronouns**

*Personal Pronouns*

<b>Personal</b>		Strong	Weak
1 <sup>st</sup> exclusive	Singular	<i>au</i>	<i>ku</i>
	Dual	<i>kimaua</i>	<i>ma</i>
	Plural	<i>kimautolu</i>	<i>mau</i>
1 <sup>st</sup> inclusive	Singular	<i>kita</i>	<i>te</i>
	Dual	<i>kitaua</i>	<i>ta</i>
	Plural	<i>kitautolu</i>	<i>tau</i>
2 <sup>nd</sup>	Singular	<i>koe</i>	<i>ke</i>
	Dual	<i>kimoua</i>	<i>mo</i>
	Plural	<i>kimoutolu</i>	<i>mou</i>
3 <sup>rd</sup>	Singular	<i>ia</i>	<i>ne</i>
	Dual	<i>kinaua</i>	<i>na</i>
	Plural	<i>kinautolu</i>	<i>nau</i>

Catherine Macdonald

*Ordinary Possessives*

<b>Ordinary Possessive ('a)</b>		Neutral		Sympathetic	
		Definite	Indefinite	Definite	Indefinite
1 <sup>st</sup> exclusive	Singular	<i>he'eku</i>	<i>ha'aku</i>	<i>si'eku</i>	<i>si'aku</i>
	Dual	<i>he'ema</i>	<i>ha'ama</i>	<i>si'ema</i>	<i>si'ama</i>
	Plural	<i>he'emau</i>	<i>si'emau</i>	<i>ha'amau</i>	<i>si'amau</i>
1 <sup>st</sup> inclusive	Singular	<i>he'ete</i>	<i>ha'ate</i>	<i>si'ete</i>	<i>si'ate</i>
	Dual	<i>he'eta</i>	<i>ha'ata</i>	<i>si'eta</i>	<i>si'ata</i>
	Plural	<i>he'etau</i>	<i>ha'atau</i>	<i>si'etau</i>	<i>ha'atau</i>
2 <sup>nd</sup>	Singular	<i>ho'o</i>	<i>ha'o</i>	<i>si'o</i>	<i>si'ao</i>
	Dual	<i>ho'omo</i>	<i>ha'amo</i>	<i>si'omo</i>	<i>si'amo</i>
	Plural	<i>ho'omou</i>	<i>ha'amou</i>	<i>si'omou</i>	<i>si'amou</i>
3 <sup>rd</sup>	Singular	<i>he'ene</i>	<i>ha'ane</i>	<i>si'ene</i>	<i>si'ane</i>
	Dual	<i>he'ena</i>	<i>si'ena</i>	<i>ha'ana</i>	<i>si'ana</i>
	Plural	<i>he'enau</i>	<i>si'enau</i>	<i>ha'anau</i>	<i>si'anau</i>
<b>Ordinary Possessive ('o)</b>		Neutral		Sympathetic	
		Definite	Indefinite	Definite	Indefinite
1 <sup>st</sup> exclusive	Singular	<i>hoku</i>	<i>haku</i>	<i>si'oku</i>	<i>si'aku</i>
	Dual	<i>homa</i>	<i>hama</i>	<i>si'oma</i>	<i>si'ama</i>
	Plural	<i>homau</i>	<i>hamau</i>	<i>si'omau</i>	<i>si'amau</i>
1 <sup>st</sup> inclusive	Singular	<i>hoto</i>	<i>hato</i>	<i>si'oto</i>	<i>si'ato</i>
	Dual	<i>hota</i>	<i>hata</i>	<i>si'ota</i>	<i>si'ata</i>
	Plural	<i>hotau</i>	<i>hatau</i>	<i>si'otau</i>	<i>si'atau</i>
2 <sup>nd</sup>	Singular	<i>ho</i>	<i>hao</i>	<i>si'o</i>	<i>si'ao</i>
	Dual	<i>homo</i>	<i>hamo</i>	<i>si'omo</i>	<i>si'amo</i>
	Plural	<i>homou</i>	<i>hamou</i>	<i>si'omou</i>	<i>si'amou</i>
3 <sup>rd</sup>	Singular	<i>hono</i>	<i>hano</i>	<i>si'ono</i>	<i>si'ano</i>
	Dual	<i>hona</i>	<i>hana</i>	<i>si'ona</i>	<i>si'ana</i>
	Plural	<i>honau</i>	<i>hanau</i>	<i>si'onau</i>	<i>si'anau</i>

*Feature Geometry of Tongan Personal Pronouns*

*Postposed Adjectivals*

<b>Postposed (Adjectival) ('a)</b>	Singular	Dual	Plural
1 <sup>st</sup> Exclusive	'a'aku	'amaua	'amautolu
1 <sup>st</sup> Inclusive	'a'ata	'ataua	'atautolu
2 <sup>nd</sup>	'a'au	'amoua	'amoutolu
3 <sup>rd</sup>	'a'ana	'anaua	'anautolu
<b>Postposed (Adjectival) ('o)</b>	Singular	Dual	Plural
1 <sup>st</sup> Exclusive	'o'oku	'omaua	'omautolu
1 <sup>st</sup> Inclusive	'o'ota	'otaua	'otautolu
2 <sup>nd</sup>	'o'ou	'omoua	'omoutolu
3 <sup>rd</sup>	'o'ona	'onaua	'onautolu

*Emphatic Possessives*

<b>Emphatic Possessive ('a)</b>	Singular	Dual	Plural
1 <sup>st</sup> Exclusive	ha'aku	ha'amaua	ha'amautolu
1 <sup>st</sup> Inclusive	ha'ata	ha'ataua	ha'atautolu
2 <sup>nd</sup>	ha'au	ha'amoua	ha'amoutolu
3 <sup>rd</sup>	ha'ana	ha'anaua	ha'anautolu
<b>Emphatic Possessive ('o)</b>	Singular	Dual	Plural
1 <sup>st</sup> Exclusive	ho'oku	ho'omaua	ho'omautolu
1 <sup>st</sup> Inclusive	ho'ota	ho'otaua	ho'otautolu
2 <sup>nd</sup>	ho'ou	ho'omoua	ho'omoutolu
3 <sup>rd</sup>	ho'ona	ho'onaua	ho'onautolu



# On the Argument Structure of Complex Predicates in Kalam, a Language of the Trans New Guinea Family

ANDREW PAWLEY

*Australian National University*

## 0. Introduction

Complex predicates, characterized by having two or more heads, are problematic for models of language that assume a neat division of labor between syntax and lexicon (Alsina et al. 1997). On the one hand, such predicates usually have many of the properties of a typical lexical unit, for example in exhibiting a unified argument structure. On the other hand, syntactically the constituents of complex predicates function separately. Various theoretical frameworks have been applied in an effort to understand the workings of such constructions.

This paper examines argument structure in three construction types in Kalam, a language spoken around the junction of the Bismarck and Schrader Ranges in the southwest corner of Madang Province, Papua New Guinea.<sup>1</sup> Kalam belongs to the Trans New Guinea (TNG) family, which with upwards of 400 members is by far the largest of the many ‘Papuan’ (non-Austronesian) families of Melanesia. Except for a branch in the Timor area, the Trans New Guinea family is confined to New Guinea and a few offshore islands.

The three types are *serial verb constructions* (SVCs), *verb adjunct constructions* (VACs), and *involuntary (bodily or mental) experience constructions* (IECs), each of which has analogs in many other TNG languages. Each construction contains, or has claims to contain, a complex predicate, but differs sharply from the other two in grammatical and semantic structure.

Serial verb constructions contain complex predicates that can have indefinitely many verb roots as heads. The main challenge here is to define what

---

<sup>1</sup> I am indebted to John Bowden and Mark Donohue for comments on drafts of this paper. My work on Kalam has been supported by grants from the Wenner-Gren Foundation for Anthropological Research, the New Zealand University Research Grants Committee, the Papua New Guinea Biological Foundation, and Amrad Discovery Technologies. I began working on Kalam as a graduate student more than 40 years ago, soon after the first government patrol post was established at Simbai. Other linguists who have worked on Kalam include Lyle and Helen Scholz of the Summer Institute of Linguistics, who spent 28 years based at Simbai, Bruce Biggs, Talmy Givón, and Jonathan Lane.

coreference and event structure constraints must apply to the arguments of different verbs before they can be combined in a single SVC. SVCs require a level of event structure semantics that refers to discourse structure well-formedness as well as to clause-internal constraints on argument structure.

Verb adjunct constructions depict a single event using two heads which belong to different parts of speech: verb and verb adjunct. The main questions here are: How do the argument structures of the two heads interact? What is the role of the verb adjunct? Is it a co-primary predicate, secondary predicate, or an incorporated argument of the verb? Do ‘verb adjuncts’ form a single and distinct part of speech, separate from nouns and adverbs?

There is a class of involuntary experience constructions which is transitive, requiring at least two nominal participants. One participant is an Experiencer, which is encoded as Direct Object. The other denotes a bodily or mental condition or process that affects the Experiencer. The main question here is: What is the grammatical status of the condition nominal? Is it the Subject, is it a kind of verb adjunct which adds to the meaning and argument structure of the verb, or is it in some strange way both Subject and verb adjunct?

My account of these three constructions will necessarily be brief and informal. In addition to constituent structure and semantic structure I will assume a distinction between two other levels of linguistic representation: argument structure (thematic arguments or macro-roles) and functional structure (grammatical functions such as subject and direct object), though I will in the usual way sometimes use ‘argument structure’ to refer to the whole domain in which thematic arguments are linked to grammatical functions. It is useful to distinguish between core arguments (roughly those conceptually central to a verb or construction, and which are usually expressed as complements) and peripheral arguments (not conceptually central and usually encoded as adjuncts).

Some leading Papuanists have been reluctant to attribute the grammatical functions Subject, Direct Object, and Indirect Object to Trans New Guinea languages or indeed to other Papuan families. Thus, in his influential book *The Papuan Languages of New Guinea*, William Foley (1986) does not use these terms at all. When talking about grammatical relations for ‘core nominals’, he refers only to semantic macro-roles, such as Actor, Undergoer, and Beneficiary.

Mark Donohue (2005) has recently argued that Subject is not usually a highly grammaticalized or well-defined category in Papuan languages of New Guinea. The entities that some Papuan grammars call Subjects are actually a conflation of constructs that belong to two different systems in the grammar:

- (i) There are Logical Subjects, which belong to the semantics of the clause. ‘Subject’ is often applied to the NP that governs verbal agreement, the selection of actor suffixes on the verb. The verbal agreement system, Donohue says, can be described without reference to ‘Subject’. Instead, it is enough to refer to role prominence using a universal hierarchy of macro-roles. (By implication, this applies also to ‘Direct Object’.)



*Argument Structure of Complex Predicates in Kalam*

- (ii) There is the topical argument in a clause, a pragmatic or discourse-based construct. What is tracked by switch-reference morphology on the verb is typically the topic, not the most prominent macro-role. It happens that the two sometimes match, but they need not do so.

Donohue concludes (for most Papuan languages of New Guinea) that “clearly we do not find a constellation of syntactic phenomena each choosing the same grouping, and so lack any reason to postulate a notion of ‘subject’” (2005:215).

These arguments against the usefulness of recognising Subject and Object as opposed to ranking arguments on a universal hierarchy of thematic roles—and thus eliminating a level of functional grammatical relations—do not sit so well with Kalam. However, the Kalam data support Donohue on another point. There are certain constructions in TNG languages where it is difficult to determine subjecthood, in particular some involuntary experience constructions where the animate participant exerts no control over the event.

The merits of a lexicalist vs. a construction-based treatment of the argument structure of these kinds of predicate phrase will be touched on. In the lexicalist approach, argument structure is regarded as being inherent in the lexical verb. That is, a verb’s array of thematic roles is defined by the meaning of the verb. Sense differences in a verb, including those that entail differences in the number and configuration of arguments, are attributed solely to polysemy in the verb and not to the different constructions in which it occurs.

I will suggest that, in the case of Kalam complex predicates, there are certain advantages in a construction-based treatment in the sense of Goldberg (1995). First, this allows for the two (or more) elements in a complex predicate to jointly determine the argument structure of the predicate. Second, allowing the construction as a whole as well as its lexical constituents to contribute to meaning, the polysemy of verbs in complex predicates can be constrained in an intuitively satisfying way.

The following abbreviations are used in glossing Kalam examples.

D	dual	PF	perfect (denotes today’s past, present perfect or present-iterative)
DS	different Subject (from following verb)		
DUR	durative	PAST.HAB	past habitual
FUT	future	PRIOR	prior or preceding (the event denoted by following verb)
HORT	hortative	PRES	present
IMM	immediate past	PROG	present progressive
NDR	noun derivative suffix	S	singular
OBJ	Object (case)	SIM	simultaneous (with the event denoted by following verb)
OPT	optative		
P	plural	SS	same Subject (as following verb)
PAST	remote past (yesterday or earlier)		

**1. Background Notes on Kalam Grammar**

Kalam has two main dialects, structurally very similar but with many differences

in morphological forms and lexicon. Some of the data cited here come from the Etp dialect of Kaironk and some from the Ti dialect of Gobnem, both spoken in the Upper Kaironk Valley. Unless otherwise stated, examples are from the Etp dialect.

### 1.1. Word Classes

The principal parts of speech are verbs, nouns, verb adjuncts, adjectives, and adverbs. Brief remarks on the first three follow. The last two will be discussed in §3.

**Verb roots** are a small, closed class with about 130 members. There are no processes for deriving new verb stems. However, polysemy roughly triples the number of sense units among verb roots. Verbs are the only part of speech to carry inflectional suffixes marking tense, aspect or mood, Subject person-and-number, and switch reference (see §1.3). The small stock of verb roots is augmented by a large body of lexicalised complex predicates. Complex predicates make up between 25 and 30 percent of verbal predicates in text.

**Nouns** are a large class with many subclasses, including personal pronouns. There are two main sets of personal pronouns: the Subject (or nominative) set and Object (or accusative) set. The most basic forms in each set in the Etp dialect are given below.

	1S	2S	3S	1D	2D	3D	1P	2P	3P
Subject	<i>yad</i>	<i>nad</i>	<i>nuk</i>	<i>ct</i>	<i>nt</i>	<i>kikmay</i>	<i>cn</i>	<i>nb</i>	<i>kik</i>
Object	<i>yp</i>	<i>np</i>	<i>nup</i>	<i>ctp</i>	<i>ntp</i>	<i>kipmay</i>	<i>cnp</i>	<i>nbp</i>	<i>kip</i>

Object pronouns are used both for canonical Direct Objects and for Indirect and Dative Objects. Possessive pronouns are drawn from one or the other of these sets according to whether the possessive phrase is the grammatical Subject or Object.

**Verb adjuncts** are roots or derived forms that occur only or primarily as the partner of one or a very few verb roots to form a complex predicate. The characteristics of this category will be described in more detail in §3.

### 1.2. Verbal Clauses and Clause Sequencing

A verbal clause consists minimally of an inflected verb. Many verb roots can take two core arguments. A few occur with three, e.g. *ñ-* ‘give’, *yom-* ‘show’, and *taw-* ‘trade, buy, sell’. Locatives are distinguished syntactically by being able to occur freely either before or after the verb. All other NPs precede the verb.

If core arguments are present in transitive clauses, the canonical word order is S-IO-DO-V, as in (1) and (2).<sup>2</sup>

<sup>2</sup> The most detailed account of Kalam SVCs is given by Lane (1991), but other studies include Givón (1990), Pawley (1987, in press), and Pawley and Lane (1998).

- (1) Bin kaj-nup piow-ya-k.  
woman pig-it(OBJ) search-3P-PAST  
'The women searched for the pig.'
- (2) An np moni ñ-a-k?  
who you(OBJ) money give-3S-PAST  
'Who gave money to you?'

Departures from this word order are fairly common. For example, a focused Object (animate or inanimate) can precede an animate Subject. When the Object is animate and the Subject is inanimate, the Object usually comes first.

Except for personal pronouns, there is no case-marking on core arguments, peripheral arguments, or adjuncts.

Because zero anaphora is normal for established referents, transitive clauses often occur without an overt Subject or Object NP. Nouns representing Subject and Object often occur without a determiner, especially when the referent is indefinite but specific.

In a sequence of conjoined clauses, the inflected verb in each non-final clause carries a suffix marking two semantic relations to the next verb: relative tense (prior, simultaneous, or prospective) and same or different Subject (switch reference). The final verb is marked for absolute tense and Subject reference.

### **1.3. Criteria Defining Subjects, Objects, and Peripheral Arguments**

In clauses that have as their head a transitive verb (either a single root or a complex verb), Subject and Object are distinguished by a number of diagnostics.

#### **1.3.1. Subject**

Subjects link with the thematic roles of Agent, Effector, Experiencer, and Theme/Patient, probably in that order. I will refer to the highest-ranked role as the Logical Subject. The grammatical (surface) Subject of a verbal clause can be identified using the following formal criteria:

- (a) When the Subject is a free pronoun, it selects a pronoun from the Subject class. This class also marks extra-clausal topics as well as possessors that are not grammatical Objects.
- (b) Agreement in clauses with a simple verb: the Subject is the NP that is co-referential with the person-number suffix on the main verb. An animacy hierarchy operates in marking number. With human Subjects, the verbal suffix must distinguish singular, dual, or plural. When higher animals are Subjects, the suffix usually distinguishes number, but in some contexts need not do so, using the singular as the default. With lower animal Subjects, the suffix seldom distinguishes number, but may do so. With inanimate Subjects, the verbal suffix is always third person singular.

- (c) Subjects control verb morphology marking switch reference.
- (d) Only Subjects of transitive verbs can be antecedents of the marker *key* in its reflexive sense (with Agent-Patient verbs) and in its sole participant sense ('by oneself, acting alone') (with Effector or Experiencer Subject verbs and Agent-Theme verbs).
- (e) Subjects prefer first position in the clause.
- (f) The verbal negative pro-clitic *ma-* seldom, if ever, precedes Subjects, but occasionally (under some conditions) precedes Objects (when these are a full NP) and Instruments, and often precedes verb adjuncts.

### 1.3.2. Object

Objects link with Experiencer, Patient, Beneficiary, Theme, Place, and Goal. Formal tests for identifying Object NPs (Direct and Indirect Object) are as follows:

- (a) When the Object NP is a pronoun, the pronoun is drawn from the Object class.
- (b) When an animate Object is represented by a full NP, the NP is optionally followed by an Object pronoun.
- (c) The preferred position for Objects is after the Subject and before the verb. Direct Objects normally follow Indirect Objects.
- (d) Although it usually follows, the verbal negative pro-clitic *ma-* can precede a light Object NP other than a pronoun.

Subject and Object are at the very least convenient labels for grammatical relations that are marked by a cluster of morphological and syntactic features. It is true that these morphosyntactic features in most cases mark the same NPs that can be identified by a semantic criterion: the NPs with the highest and second highest ranked thematic role in the clause. However, there are exceptions to this generalization, to be discussed in §2 and §4.

## 2. Serial Verb Constructions

### 2.1. General Characteristics

Kalam serial verb constructions (SVCs) provide an elegant and streamlined mechanism for expressing certain kinds of event sequences or overlapping event clusters in a single clause.<sup>2</sup> The predicate (verb phrase) of an SVC has as its nucleus a *verb series*, in which one or more bare verb roots precedes an inflected verb root. There is no grammatical limit to the number of verb roots that can

occur in an SVC. In practice, if we exclude iteration of verb roots to show repetition or continuity, the limit seems to be about ten. The question is: What are the semantic constraints on possible SVCs?

Typically, each verb root in an SVC specifies a separate event in a sequence or an event that overlaps in time with another. The verbs must share the same tense/aspect/mood marking, and this appears only on the final verb. (In schematic representations of verb series, only the final verb is followed by a hyphen to indicate the placement of inflections.) In most types of SVC, the verb series falls under the scope of a single negator and is uttered within a single intonation contour. Some of the main structural features of SVCs are illustrated by (3) and the second clause of (4). Verb roots and their glosses appear in boldface.

- (3) **Am d aw-an!**  
**go get come-2S.IMP**  
'Fetch (it)!' (lit. 'Go get (it) and come!')

- (4) Wel **d-l,**  
oil **get-SS.PRIOR**  
yp **wik d ap tan d ap yap g-s<a>p.**  
me **rub get come ascend get come descend do-PRES.PROG<3S>**  
'He is massaging me with oil.' (or: 'Using oil, he is giving me a rub-down.')'

In terms of their semantic and syntactic structure, SVCs can be divided into two main types: *compact* and *narrative* SVCs.

## 2.2. Compact SVCs

Compact SVCs express a tight-knit event sequence, one where the constituent events are connected in a causal or close temporal relation. The verb series is also syntactically tight-knit. No arguments can be inserted within it. Locatives can either precede or follow the verb series; all other arguments must precede it. The negative clitic and any adverbial modifiers must precede the entire series. Compact SVCs can be regarded as lexicalized phrases and are often translatable by a single verb in English. Some examples are: *d am-* (get go) 'take', *d ap-* (get come) 'bring', *ptk am-* (fear go) 'flee', *ag ñ-* (say transfer) 'tell', *ag ask-* (say avoid) 'refuse', *ag ay-* (say stabilise) 'make an appointment, ask to stay', *ag ny-* (say perceive) 'ask', *d ny-* (touch perceive) 'feel', *ñb ny-* (consume perceive) 'taste', *pk cg-* (strike adhere) 'attach, stick on', *d ap tan d ap yap-* (hold come ascend hold come descend) 'move back and forth'. While most compact SVCs contain two to four verb roots, they are not limited to this range. Thus, the second clause of (4) contains an eight-verb sequence which is arguably a single compact SVC.

The constraints on the marking of grammatical relations in compact SVCs are like single-verb clauses: only one Subject and one Object NP can be expressed. The final verb in the series is always inflected to agree with the Logical Subject

(the most prominent thematic role) of the first verb in the series. In most kinds of compact SVC this requires no adjustment in the surface structure: all the verbs share the same Logical Subject. There is, however, one type where this perfect match between Surface Subject and Logical Subject does not hold: resultative or cause-effect SVCs, exemplified in (5).

(5)	<i>pk cg-</i>	(strike stick)	‘stick sth. on, cause sth. to adhere’
	<i>pk wk-</i>	(strike shattered)	‘knock sth. to bits, shatter sth.’
	<i>pk sug-</i>	(strike extinguished)	‘put out (a fire)’
	<i>pug sug-</i>	(blow extinguished)	‘blow out (a flame)’
	<i>puji ask-</i>	(pierce open)	‘prise sth. open’
	<i>puji lak-</i>	(pierce split)	‘split sth. by wedging or levering’
	<i>ag yok-</i>	(say move.away)	‘send s.o. away, dismiss s.o.’
	<i>d am yok-</i>	(get go move.away)	‘get rid of sth., take sth. away’
	<i>pk lug yok-</i>	(strike slide move.away)	‘brush or sweep sth. away’
	<i>puji ju-</i>	(pierce withdraw)	‘extract sth. with a probe’
	<i>taw lug-</i>	(step.on slide)	‘push sth. away with the feet’
	<i>taw pag yok-</i>	(step.on broken displace)	‘break sth. off by stepping on it’
	<i>tb kluk yok-</i>	(cut gouge displace)	‘gouge sth. out (the centre of sth.)’

In the simplest case, resultative SVCs contain just two verbs: V1 is transitive and specifies an activity, usually a verb of contact, and V2 is intransitive and specifies a state or movement. The meaning derived from the sequence is that the event denoted by V1 precedes and causes the event denoted by V2, the verb of result. The two verbs share an argument: the Logical Subject of V2, the entity which is in a state or which moves, is also the Patient of V1, the entity which suffers contact. When it comes to marking of Subject, the two verbs behave like a single compound transitive verb. V1 is bare, and V2 carries agreement suffixes referring only to the Logical Subject of V1. That is to say, the SVC adopts the functional structure of the initial, transitive verb, while that of the final, intransitive verb is ignored. This sort of reworking of mismatched lexical argument structures to create a unified structure in the complex resultative predicate is crosslinguistically familiar.

Resultative SVCs are not restricted to two verb roots. More than one verb root can fill the V1 slot, so long as the first verb is transitive and the other roots share with it the same Logical Subject. Examples in (5) above are *d am yok* (get go move.away) ‘get rid of s.th.’ and *pk lug yok* (strike slide move.away) ‘brush s.th. away’. More than one verb can fill the V2 slot, so long as all the verb roots are intransitive and share with it the same Logical Subject. An example in (5) is *taw pag yok-* (step.on broken move.away) ‘break s.th. off by stepping on it’.

### 2.3. Narrative SVCs

Narrative SVCs have a much more complex semantic and constituent structure than compact SVCs. They express in formulaic form a sequence of distinct events

*Argument Structure of Complex Predicates in Kalam*

or distinct event clusters, each taking place at different times and in different places, that together make up an episode, a short story.

Narrative SVCs can readily be paraphrased by a chain of clauses. However, a sequence of clauses can be reduced to an SVC only under certain conditions. The clauses must be coordinate; the initial verbs in each clause must share the same Logical Subject and the same Patient (if any); the clauses must report a sequence of events which taken together conforms to Kalam discourse structure conventions for a minimal well-formed report; and the report must be of a familiar class of episodes. For example, a well-formed report on someone's gathering of food or materials should mention four stages: (1) the actor's movement (if any) to the scene of the gathering, (2) the gathering activities, (3) transport of the goods to the scene of disposal (if it occurs), and (4) the disposal activities, such as cooking and eating, distribution, or storage.

Any single stage of a narrative may be represented by an SVC, but to qualify as a *narrative SVC*, two or more stages of the narrative must be included in the SVC. In example (6), all four stages of a gathering (hunting) episode are represented. In (7) and (8), stages 2-4 are present. Example (9) reports a different kind of episode: escape from confinement, where the component stages are somewhat different. Examples (6)-(9) are all in the Ti dialect.

- (6) Basd skop ...**am** kmn pak d ap ad ñb-elgp-al.  
 grandfather distant **go** animal **kill get come cook eat**-PAST.HAB.3PL  
 'Our distant ancestors...used to hunt [lit. go, kill, bring back, cook and eat] game mammals.'  
 (*kmn* 'the larger marsupials and arboreal rodents, which men hunt')
- (7) ...mj bep tk d ap nb okyañ **yok-l**,...  
 leaf plant **sever get come** place below **throw**-SS:PRIOR  
 '...having gathered *bep* leaves and brought and tipped (them) below'
- (8) ...kuñp ognap **tb d am** katp-at okok **l-l**,...  
 kuñp(leaves) some **cut get go** house-area about **put**-SS:PRIOR  
 '...having gathered *kuñp* leaves and taken to the house and put (them),...'
- (9) Ognap **am su ñb ñb** mgan okdañ nb pag **jak-l**,...  
 sometimes **go bite eat eat** inside there place **break rise.up**-SS:PRIOR  
 'Sometimes, having chewed their way through from inside (the burrow and protective foliage), they break out...'

Each stage in a narrative SVC is represented by a separate predicate phrase, which may be a compact SVC, a single verb, or certain other sequences of verb roots. The predicate phrases show some degree of syntactic independence from one another. Each predicate phrase can take separate adverbial modifiers, for instance, and one sometimes finds that two predicate phrases each has its own

Locative adjunct, as in (7). A negative proclitic having scope over the whole SVC may either precede the entire verb series or may precede the final verb. The negative clitic may also negate only the final verb, in which case it immediately precedes that verb.

In spite of their considerable semantic and syntactic complexity, narrative SVCs have a core argument structure that is no more complex than that of compact SVCs. No matter how many predicate phrases the narrative SVC contains, only one Subject NP and one Object NP can appear overtly. Subject agreement is only marked once on the final verb in the construction. The initial verbs of all predicate phrases in the narrative SVC must have the same Logical Subject. If more than one predicate phrase has an (understood or overt) Object, this must be shared.

### **3. Verb Adjunct Constructions**

#### **3.1. General Characteristics**

Verb adjunct constructions (VACs) have the following grammatical and semantic characteristics:

- (a) They consist minimally of a verb adjunct plus a verb. Verb adjuncts are an open class of roots and derived words, with several hundred recorded members which occur only as the partner of a verb in a complex predicate. (New members may be recruited to the class by derivational processes and/or by borrowing.)
- (b) The verb root is inflected, except when the VAC functions as a non-final serial verb.
- (c) They express a single event.
- (d) The verb root serves as a classifier. It marks the event as being of a certain general semantic type. The verb adjunct specifies the event as being of a particular subtype of this class or as adding an associated activity to that depicted by the verb root. Only about 20 verbs take part in verb adjunct constructions.
- (e) The verb adjunct contributes to the argument structure of the VAC and can be viewed as a co-predicate.

VACs are exemplified by (10)-(12). In these and later examples, the verb adjunct and verb roots appear in boldface, and the whole verb adjunct construction is bounded by square brackets. In an attempt to provide a gloss that is neutral between verb, noun, and adverb, a gerundive gloss (e.g. perching.together, begging, slithering) is used for some verb adjuncts.



*Argument Structure of Complex Predicates in Kalam*

- (10) Sawan [**guglum ag-ig**] k-j<a>p.  
Sawan **snoring say-SS.SIM** sleep-PRES.PROG<3SG>  
'Sawan is asleep, snoring.'
- (11) Yakt omay ok [**gub g-i**] md-p-it.  
bird two those **perching.together do-SS.PRIOR** stay-PF-3D  
'Those two birds are perched together.'

Usually the adjunct and verb root are contiguous, but certain kinds of material may intervene, e.g. locative phrases as in (12), question words as in (13), and modifiers and negators.

- (12) Tap tubtub-toktok wad g-i, [**dad ms amn-a-k**].  
things knick-knacks bag do-SS-PRIOR **carrying** outside **go-3S-PAST**  
'Having put his personal effects in a string bag, he carried them outside.'
- (13) Pa-skoy [**si etp-nen ag-a-k**]?  
girl-small **crying** what-for **say-3S-PAST**  
'Why did the girl cry?'

### 3.2. Classes of Verb Adjunct Construction

VACs can be classified in the first place according to the meaning of the verb adjuncts and the verb (or verbs) they select as co-predicate. Examples follow from each of the three largest classes.

A sizeable class of verb adjuncts selects a verb of locomotion and denotes the manner, direction, or concomitant of the actor's movement. The locomotion verbs that take the widest range of verb adjuncts are *am-* 'go' and *ap-* 'come, appear'. Some adjuncts also occur with *tag-* 'walk about, travel', *sayd-* 'depart', *yap-* 'go down, fall', *tan-* 'go up, climb', *ju-* 'withdraw', and *yok-* 'move away, be displaced'. Most verb adjuncts that occur with a verb of locomotion fall into one or another of the semantic categories shown in (14).

- (14) a. Manner of actor's movement along a path
- |                     |                          |                          |
|---------------------|--------------------------|--------------------------|
| <i>kleyd am-</i>    | (crawling go)            | 'crawl'                  |
| <i>mumlok d am-</i> | (rolling, tumbling go)   | 'fall over, tumble down' |
| <i>pug-tkd am-</i>  | (running go)             | 'run'                    |
| <i>tawd am-</i>     | (stepping go)            | 'walk'                   |
| <i>wald am-</i>     | (sliding, slithering go) | 'slide, slither'         |
| <i>wayd am-</i>     | (carried.by.water go)    | 'float along'            |
| <i>wy d am-</i>     | (flying go)              | 'fly'                    |
- b. Direction of actor's movement in relation to path or deictic centre
- |                        |                         |                         |
|------------------------|-------------------------|-------------------------|
| <i>gogeb-mageb am-</i> | (zigzagging go)         | 'zigzag along'          |
| <i>koday kodoy am-</i> | (hither and thither go) | 'go hither and thither' |

<i>payd am-</i>	(passing out of sight go)	‘pass from view’
<i>adkd am-</i>	(turning back, reversing go)	‘turn back’
<i>skd am-d</i>	(entering go)	‘go in, enter (an enclosure)’
<i>tkd am-</i>	(crossing over go)	‘cross over (a divide)’

- c. Manner in which a burden is transported or a person/animal is accompanied

<i>dad am-</i>	(carrying go)	‘carry sth.’
<i>pojɔd am-</i>	(guiding go)	‘guide s.o., take s.o. (to a place)’
<i>sb am-</i>	(brandishing go)	‘go brandishing sth. (usually a weapon)’
<i>kosiɔd am-</i>	(stacked, piled go)	‘go carrying sth. on the back’
<i>wtsek am-</i>	(chasing go)	‘chase, pursue s.o.’

The combined arguments of certain VACs exceed those of any of the constituent predicates. This applies to the VACs in (14c) where each of the verb adjuncts refers to a separate participant. In (15) a verb adjunct *pojɔd* ‘guiding, leading, accompanying (a person or animal)’ is paired with the verb *ap-* ‘come’.

- (15) *Yad b tud nup kotp-cn pojɔd ow-n-k.*  
 I man white him house-our **leading** **come-1S-PAST**  
 ‘I’ve brought the white man to our home.’

The adjunct here contributes two participants: a guider (thematic Agent, grammatical Subject) and a guidee (Theme, Object). The verb also contributes two participants: a comer (Agent, Subject) and a destination (Goal, Locative). The guider and the comer are coreferential. Together these two elements create a three-place predicate: guider/comer (Agent, Subject), guidee (Theme, Object), and destination (Goal, Locative).

In (16), the adjunct *dad* ‘carrying (inanimate object)’ is paired with the verb *n-* ‘join others on a trip, catch up with someone who has left’.

- (16) *Cn np mdak ayn-bogs dad n-ngp-un.*  
 we you(SG.OBJ) later iron-box **carrying** **join-FUT-1P**  
 ‘We will join you later bringing the metal trunk.’

Together they create a three-place predicate. The adjunct *dad* contributes a carrier (Agent, Subject) and a burden (Theme, Object). The verb *n-* contributes a joiner (Actor, Subject) and a joinee (Goal, Object). The carrier and the joiner are coreferential. In (16) the burden is the Direct Object, and the person joined is the Indirect Object.

*Argument Structure of Complex Predicates in Kalam*

There is a large class of verb adjuncts which select the verb *ag-* ‘talk, say, make a sound’. These verb adjuncts almost all refer to kinds of noises. As the sole head of a verb, *ag-* has both intransitive and transitive uses. Some VACs containing *ag-* are intransitive, as in (17a). Others, where the adjunct refers to a form of speech, are transitive, as in (17b).

(17) Some VACs containing *ag-* ‘say, make a sound, etc.’

a. intransitive constructions

<i>bu ag-</i>	(exploding say)	‘explode, burst’
<i>gigu ag-</i>	(rattling say)	‘rattle, jingle’
<i>guglak ag-</i>	(croaking say)	‘croak’
<i>guglum ag-</i>	(snoring say)	‘snore’
<i>mukbel ag</i>	(belching) say	‘belch’

b. transitive constructions

<i>asb ag-</i>	(begging say)	‘keep asking s.o. for things to be given’
<i>paj ag-</i>	(taunting say)	‘taunt s.o. over their misfortune’
<i>kub ag-</i>	(loud.calling say)	‘call loudly to s.o.’
<i>kuk ag-</i>	(alarm.calling say)	‘shout an alarm to s.o.’

Given that *ag-* out of context is of indeterminate transitivity, it seems that the transitivity of a given VAC string containing this verb is determined by the verb adjunct. In intransitive constructions, the verb adjunct is a single argument predicate referring to a kind of sound that is not speech and so is not (on a normal reading) addressed to anyone. In transitive constructions, the verb adjunct denotes a form of speech which implies an addressee.

The verb that takes the largest number of verb adjuncts is *g-* ‘act, function, do, make, perform’, the most general activity and process verb. This verb, too, has both intransitive and transitive uses. VACs that contain *g-* generally denote either (a) an intentional act, (b) a dynamic event attributed to an inanimate effector, or (c) a state. Examples are given in (18). Some verb adjunct phrases belong to more than one category, e.g. those in (18c) have a dynamic (and transitive) reading, but some also have a stative reading (18d). Although *g-* is glossed as ‘do’ throughout, in some contexts the gloss might equally well be ‘make’ and in others, ‘be’.

(18) Some verb adjuncts occurring with *g-* ‘do’

a. intransitive VACs denoting intentional acts

<i>gub g-</i>	(perching.together do)	(of birds) ‘perch together’
<i>jlan g-</i>	(nodding do)	‘nod the head rhythmically’
<i>gsey-bsey g-</i>	(hurry-scurrying do)	‘hurry, be in a hurry or rush’
<i>ñk g-</i>	(crouching do)	‘crouch, duck’

- b. transitive VACs denoting intentional acts
- |                       |                      |                            |
|-----------------------|----------------------|----------------------------|
| <i>bl g-</i>          | (abstaining do)      | ‘abstain from an activity’ |
| <i>gadal-badal g-</i> | (criss-crossing do)  | ‘lay things criss-cross’   |
| <i>gdey-bdey g-</i>   | (rumple-jumpling do) | ‘do something roughly’     |
| <i>saj g-</i>         | (compensating do)    | ‘pay compensation’         |
- c. transitive VACs denoting events with inanimate Effectors
- |                       |                         |   |
|-----------------------|-------------------------|---|
| <i>gtij-gtoj g-</i>   | (ding-donging do)       | ‘make a lot of noise, make a din or racket’ |
| <i>gley-wley g-</i>   | (clitter-clattering do) | ‘rattle, clatter, as bones, metal objects’  |
| <i>kawn g-</i>        | (swaying do)            | ‘swing, rock, sway, flap’                   |
| <i>kopay-mopay g-</i> | (blowing.fiercely do)   | (of storm winds) ‘blow fiercely’            |
| <i>lm g-</i>          | (shooting do)           | (of plant suckers) ‘shoot up’               |
| <i>wnwn g-</i>        | (peeling do)            | (of, e.g., bark) ‘peel off, fall to bits’   |
- d. intransitive VACs denoting states
- |                       |                      |                                       |
|-----------------------|----------------------|---------------------------------------|
| <i>gadal-badal g-</i> | (criss-crossing do)  | ‘be higgledy-piggledy, criss-crossed’ |
| <i>gdey-bdey g-</i>   | (rumple-dumpling do) | ‘be in disarray’                      |
| <i>gogeb-mageb g-</i> | (twisted do)         | ‘be twisted, crooked’                 |
| <i>gutgat g-</i>      | (drenched do)        | ‘be drenched and cold’                |
| <i>kolkol g-</i>      | (tangled do)         | ‘be tangled’                          |
| <i>jl g-</i>          | (loose-fitting do)   | ‘be loose, loose-fitting’             |

As with *ag-*, it seems that the verb adjunct determines the transitivity of VACs containing *g-*.

### 3.3. Are Verb Adjuncts a Distinct Part of Speech?

The verb adjunct category is problematic. Although some common features distinguish this class from other parts of speech, its members are far from homogeneous in semantic type or syntactic behavior. In meaning, some verb adjuncts resemble prototypical verbs, while others resemble nouns, adjectives, or adverbs. In their grammatical behavior, verb adjuncts are sharply distinct from verbs. Verb adjuncts cannot take verbal inflections and cannot occur alone as the head of a predicate phrase. They differ from adjectives, which can modify nouns.

Verb adjuncts resemble adverbs in being modifiers of verbs. However, they differ from typical adverbs in several respects: (a) an adverb can combine with many different verb roots whereas a verb adjunct can combine only with a very restricted set; (b) adding an adjunct in some cases creates a complex predicate with a different argument structure from the verb root alone, whereas adverbial modifiers do not affect the valency of the predicate phrase; (c) adjuncts can take

at least some adjectival modifiers but adverbs cannot; (d) adverbs modify the verb in a general sense and in a graded way, e.g. for speed, intensity, or purposefulness, whereas verb adjuncts typically specify a particular kind of action, process, or state that is a subtype of the category of event denoted by the verb; and (e) whereas adverbs may modify the whole VP, the scope of verb adjuncts is restricted to the verb.

Canonical verb adjuncts differ from nouns in that they are cannot be possessed or quantified and cannot be questioned, topicalized, or relativized. Some verb adjuncts resemble nouns in that they can be modified by intensifying adjectives (such as *yob* ‘big, loud’, *koŋay* ‘many’, *tmey* ‘bad, intensifier’, and *naban* ‘very’) and in that they represent a propositional participant of a performance verb such as *ag-* ‘say’ and *g-* ‘do’, i.e. the adjunct represents something said or done. One has to say something, and one has to do something. Thus, in (11), *gub* ‘perching together’ is the action performed, and in (13), *si* ‘crying’ is the sound that is uttered even though neither verb adjunct has strong claims to be the direct object. In these respects, they resemble incorporated objects in many languages, some coverbs in Jaminjung (Schultze-Berndt 2000:189-190), and to some extent nominals in some N+V complex predicates in Hindi which serve both as arguments and as predicates (Mohan 1997).

Foley (1986:117-119) would extend the class of verb adjunct constructions in Kalam to include sequences where the ‘adjunct’ is a typical noun which occurs before a verb to form a conventional expression, e.g. *wdn ny-* (eye perceive) ‘see’, *tmd ny-* (ear perceive) ‘hear’, *gos ny-* (thought/mind perceive) ‘think’, *wsn ny-* (dream perceive) ‘dream’, *mnm ag-* (speech say) ‘speak’, *kmap ag-* (song say) ‘sing’. In such expressions the nouns are non-referential, so are not arguments, and can be viewed as incorporated in the predicate phrase. (Nouns occurring in such sequences are referred to in Pawley et al. (2000) as ‘quasi verb adjuncts’, in contrast to ‘true verb adjuncts’.) A problem is that these nouns have different combinational privileges from true verb adjuncts. Unlike verb adjuncts, the nouns can take determiners, quantifiers, and possessors, and they can be arguments. Thus, *mnm* ‘speech’ and *kmap* ‘song’ can be the Direct Object of *ag-* ‘say, utter, make a sound’ and *wdn* ‘eye’ and *tmd* ‘ear’ can be treated as Locatives or Instruments (the place or means of perception) of *ny-* ‘perceive’.

Adopting a constructional approach might allow us to have the best of both worlds. Constructions where an incorporated non-referential noun modifies a verb show a strong family resemblance to, or formal overlap with, constructions where a non-referential verb adjunct modifies a verb. This family resemblance might be formally represented by positing a single major construction type which subsumes both verb adjuncts and incorporated nouns, while also recognizing that there are a number of subtypes in this family. The fact that these nouns can also occur as Direct Objects or Instruments would be captured by recognizing another set of constructions.

#### 4. Involuntary Experience Constructions

##### 4.1. General Characteristics

It is the third group of constructions, involuntary experience constructions (IECs), that is hardest to analyze.<sup>3</sup> IECs denote uncontrolled bodily or mental events, such as bleeding, sweating, shivering, feeling sick, hungry, or angry, needing to vomit, being overcome by laughter, or having boils, warts, or pimples. It is well-known that languages vary a good deal in how they describe such events, in particular in the grammatical status given to the Experiencer and to the constituent (be it verbal or nominal) denoting the bodily or mental processes.

Unlike English, Kalam expresses many uncontrolled bodily and mental experiences by transitive clauses, albeit non-typical ones. It is transitive IECs that we are concerned with here. These clauses require at least two nominal participants. One participant, which I will refer to by the cover term ‘Condition’, refers to a bodily or mental process or state that affects the Experiencer. The other represents either an Experiencer NP or a Location NP. The Location NP refers to the body part that is the locus of the Condition. It is usually possessed, or the possessor is recoverable from the context. When the Experiencer consists of a pronoun, it is always accusative. Likewise, when a body part nominal in an IEC is possessed, it is always possessed by an accusative pronoun.

The verb (which may be a single verb or a compact SVC) marks a relation between these participants, specifically the manner in which the Condition manifests itself in or affects the Experiencer or body part. For example, a Condition may ‘appear’, ‘form’, ‘fall’, ‘rise’, ‘grow’, or ‘act’ on, or it may ‘pierce’ or ‘eat’ the Experiencer. The usual order of constituents is LOCATION EXPERIENCER CONDITION VERB, as in (19)-(24). However, in some kinds of IECs, the three preverbal constituents can occur in any order. In (19-24) and later examples, the Condition nominal and the verb root(s) are shown in boldface.

(19) Np                    **wsb**     **jak**-s<a>p.  
       you.S(OBJ)     **sweat**   **rise**-PRES.PROG<3S>  
       ‘You are starting to sweat.’

(20) Nup    **suk**            **ow**-p.  
       him    **laughter**   **come**-PF-3S  
       ‘He is about to laugh.’/‘He feels like laughing.’

(21) Yp    **ydk**                    **g**-p.  
       me    **good-tasting**   **do**-PF.3S  
       ‘It tastes good to me.’

---

<sup>3</sup> A fuller account of Kalam IECs is given in Pawley et al. (2000). The only other detailed treatment of IECs in a TNG language that I am aware of is a very thorough analysis of Amele by Roberts (2001). Amele IECs work somewhat differently from those of Kalam.

- (22) Kud=yp            ak     **yuwt**     **g-s<a>p**.  
back=my(OBJ)    the     **pain**     **do-PRES.PROG<3S>**  
'It is my back that is hurting.'
- (23) Jun=np        **yuwt** tney **g-p**.  
head=her     **pain** bad **do-PF-3S**  
'She has a bad headache.'/'Her head is hurting badly.'
- (24) Tglm-wagn=yp     **puji**    **ñb-s<a>p**.  
ribs-base=my(OBJ) **pierce** **eat-PRES.PROG<3S>**.  
'There's a sharp pain in my ribcage'.

The grammatical status of the Experiencer and Location nominals in IECs is clear. Given that they select an accusative pronoun, these must be regarded as Objects of a sort. They cannot be the Subject, for three main reasons. First, the Experiencer is represented by a pronoun from the Object set. Second, the person-number suffix on the verb does not agree with the Experiencer NP; the verb is always marked for third person singular. Third, in this class of constructions, switch reference, a syntactic process sensitive to Subject identity or change across clauses, cannot be triggered by the Experiencer or Location NP.

Some IECs contain separate Experiencer and (non-possessed) Location NPs, as in (25). This co-occurrence requires a distinction to be made in such constructions between two kinds of Objects, call them 'Direct' and 'Locative' Objects.

- (25) Yp **mablep** kogm **ay-a-k**.  
me wart(s) knee **form-3S-PAST**  
'I had a wart/warts on the knee.' (lit. 'Wart(s) formed on me on the knee.')

#### **4.2. What Is the Status of the Condition Nominal (and Other Questions)?**

Various questions arise in the analysis of IECs. What is the grammatical role played by the Condition nominal? Is it the Subject? Is it a kind of verb adjunct, and so part of a complex predicate? Or does it contrive to be both verb adjunct and incorporated Subject? Do IECs have a Subject at all? Is a uniform analysis of IECs possible, or is this a disparate class? Another question concerns the source of the argument structure of IECs. Is this projected by the verb, or does it belong to the construction as a whole?

Foley (1986), writing about Kalam IECs, and Donohue (2005), writing about IECs in another TNG language, Lani, regard the Condition nominal as a kind of verb adjunct, forming a complex predicate together with the verb. On this analysis, Kalam IECs would have no Subject other than a dummy representation by a third-person singular suffix on the verb. It is true that all Condition nominals resemble verb adjuncts and differ from typical Subjects in that they typically occur immediately before the verb (rather than in canonical Subject position verb-initially). It is also true that some Condition nouns resemble verb adjuncts in that they co-

occur only with a single verb or a very few verbs, and in that they are non-referential, or at least cannot take a determiner or quantifier or occur as Direct Objects, topics, or heads of relative clauses.

However, Condition nominals are not a homogeneous class. Some Condition nominals score highly on standard tests for nouniness (§1.1) and show very few verb adjunct properties (§3.1). Others (call them ‘atypical’ nouns) score rather poorly on tests for nouniness and show more verb adjunct properties. Both sets of nouns meet the most important tests for subjecthood that are applicable to IECs in all constructions they occur in.

Unfortunately, two of the most reliable tests for identifying Subjects, namely (i) selection of a nominative pronoun, and (ii) Subject-verb agreement, are indeterminate for IECs. The pronominal test does not work because Condition NPs (unlike Experiencers) cannot be pronominalized in this context. The Subject agreement test is indeterminate because when the Subject NP is inanimate, the verb agreement suffix is invariably third person singular, regardless of whether the NP is singular, dual, or plural.

We are left with the four other morphosyntactic tests mentioned in §1.3 and with a semantic criterion: position in the hierarchy of thematic roles.

#### 4.2.1. Switch Reference as a Test of Subjecthood

It was noted earlier that verbs that belong to coordinate clauses carry suffixes marking change of Subject (DS) or same Subject (SS). Control of switch reference is the most powerful of the available grammatical tests for subjecthood in IECs.

All Condition nominals appear to be able to trigger switch reference. The complex sentences (26) and (27) each consist of a pair of IECs. In (26), the first clause describes a bodily process which precedes (and, by implication, causes) a sensation described in the second clause. A change of Subject is signalled in the first clause. The only candidates for Subject are the Condition nominals, denoting a boil and pain, respectively.

- (26) **Sɲl** nñ=yp **ay-e-k**, (yp) **yuwt g-p**.  
**boil** arm=my **form-DS.PRIOR-3S** (me) **pain do-PF.3S**  
 ‘A boil has formed on my arm and it’s painful (to me).’

In (27), the two clauses describe simultaneous bodily processes: a growling stomach and being hungry. The first condition is, by implication, a symptom or consequence of the second. Again, the only candidates for Subject are the Condition nominals.

- (27) **Sb-wt=cnp gullg ag-a-knɲ**, **cnp yuan g-s-a-p**.  
 innards=our **rumbling say-3S-DS.SIM** us **hunger act-PRES.PROG<3S>**  
 ‘When our stomachs are growling, we are hungry.’  
 (‘When rumblings sound in our innards, hunger has acted on us.’)



*Argument Structure of Complex Predicates in Kalam*

In the next example, the two clauses share the same Experiencer, but not the same Subject. The first clause is an IEC with an Experiencer Object and describes an illness (*tap*) that affected the Experiencer. The second clause is not an IEC. Here the Experiencer is the Subject, and there is no Condition nominal; instead, the verb sequence describes the condition: remaining indisposed (unable to work) for a period. The only nominal available to be Subject in the IEC clause is *tap*.

- (28) Nd nup **tap** g-e-k, (nuk) kum  
 first him **sick(ness)** **do**-DS.PRIOR-3S.PAST (he) indisposed  
 md-igp.  
 stay-PAST.HAB.3S  
 ‘First he got sick, then (he) wasn’t able to do anything for some time.’

In the next example, the first clause has an Agent Subject, the eater. In the second clause, an IEC depicts a subsequent event in which the eater is the (understood) Object, experiencing pleasure. The only available candidate for Subject in the second clause is *tep* ‘good, pleasure, pleased’.

- (29) Ñb-e-n (yp) **tep** g-p.  
 eat-DS.PRIOR.1S me **good** **act**-PF.3S  
 ‘It tastes good.’ (lit. ‘I having eaten, good/pleasure acts (on me).’)

There is a qualification to make about the switch reference test. In some circumstances a verb can be marked for same-Subject-as-next-clause even when the following clause has a different referent as Subject. This option is available when the first clause depicts an event with an Effector (inanimate cause) as Subject and when the following clause depicts a similar class of event which is regarded as a natural consequence of the first, e.g. lightning flashes, then thunder sounds. Examples (30)-(31) illustrate this point. Each contains a pair of IECs containing different Condition nominals, and the second IEC describes an event consequent upon the preceding event.

- (30) **Sñl** (nñ-yp) **ay-i** (yp) **yuwt** g-p.  
**boil** (arm-my) **form**-SS.PRIOR (me) pain act-PF.3S  
 ‘A boil has formed (on my arm) and it’s painful (to me).’

- (31) Yp wog **yuwt-bt** g-i, yp **ytuk** g-p.  
 me work **exhaustion** **do**-SS.PRIOR me lethargy act-PF.3S  
 ‘I’m worn out from working.’ (lit. ‘Work weariness having acted on me, lethargy has affected me.’)

However, marking for different subject is always an option in these circumstances (with no change in meaning), so the switch reference test remains valid.

#### 4.2.2. Coreference with Reflexive/Sole Participant Marker *key*

It was noted in §1.3 that only Subjects can be the antecedent of *key* when this word marks reflexive (with Agent-Patient verbs) or refers to the sole participant in an event (Effector or Experiencer Subject and Agent-Theme verbs). The Experiencer in an IEC cannot be coreferential with *key* in either of these senses. However, the Condition nominal behaves like other verbs with Effector Subjects:

- (32) **Soy** mebi **key** **ay-p.**  
**ulcer** here by.itself **form.PF.3S**  
 ‘This sore here formed by itself.’ (e.g. there was no prior wound)

#### 4.2.3. Other Tests

The other two morphosyntactic tests for subjecthood (preferred order of major constituents in the clause and placement of the negative pro-clitic *ma-*) are not worth a great deal. In IECs, the Condition nominal can occur clause-initially in the preferred position for Subjects, as in (33a). However, its preferred position is immediately before the verb, following the Object (33b), and in this respect Condition nominals resemble verb adjuncts. But while this is an unusual position for animate Subjects, it is fairly common for an inanimate Subject (e.g. an arrow) to follow an animate Object (e.g. a person wounded by the arrow).

- (33) a. **Yuan** yp **g-p.**  
 hunger me do-PF.3S  
 ‘I am (feeling) hungry.’
- b. Yp **yuan** **g-p.**  
 me hunger do-PF.3S  
 ‘I am (feeling) hungry.’

The negative proclitic normally attaches to the verb and follows the Condition NP. However, it can precede it.

- (34) a. Yp **yuan** **ma=g-p.**  
 me **hungry** not=**do**-PF.3S  
 ‘I am not (feeling) hungry.’
- b. Yp **ma=yuan** **g-p.**  
 me not=**hunger** **do**-PF.3S  
 ‘I am not (feeling) hungry.’

The order in (34b), which parallels the fairly frequent occurrence of *ma-* before verb adjuncts, might be taken as evidence that the Condition nominal is not a Subject, but is part of a verb phrase constituent consisting of the verb plus verb adjunct (if any) plus non-Subject arguments. However, the negative clitic is rather

flexible in its placement. It is occasionally found before canonical Direct Objects (if they are not pronouns) and before Instruments, and some informants even accept its placement before a Subject when this is an inanimate nominal (e.g. *yakam* ‘arrow’) that is placed immediately before a verb.

Appealing to the ranking of thematic roles is somewhat unsatisfactory insofar as it invites circularity. The ranking of roles is determined in the first place by which nominals one regards as being a Subject or not a Subject, and so on, in particular cases. Nevertheless, there are clear cases, and these can be used as a reference point when dealing with unclear cases.

#### **4.2.4. Cases Where the Condition Nominal Is a Canonical Noun**

Test for nouniness allow us to gauge where particular Condition nouns stand on the scale of normal to peculiar Subjects. In some types of IEC, the Condition nominal is a normal noun and behaves as a normal Subject nominal apart from a preference for occurring next to the verb. These are chiefly Condition nominals that refer to a tangible, substantial bodily condition. A selection of cases follows.

- (a) Stable visible conditions marked by *ay-* ‘put, form, stabilise, become’

Condition nouns denoting a stable, visible condition typically occur with the verb *ay-* ‘put, form, stabilize, become, turn into’, often with a body part nominal specified. Examples of such Condition nouns are *kñowj* ‘birthmark’, *magi-wt* ‘scar’, *sbek* ‘pimple’, *slañ* ‘scab’, *syl* ‘boil, abscess’, *soy* ‘sore, ulcer’, *tmd sb* ‘ear-wax’, and *wdn-sgalb* ‘sleep (dry secretion in eyes)’. These are canonical nouns which can take determiners, quantifiers, and possessors and can occur as Objects, topics, or heads of relative clauses. In (35), *syl* ‘boil’ takes a determiner and is modified by two adjectives, the second a quantifier.

- (35) Nup **snl** yob omɲal ak alkjon **ay**-a-k.  
him **boil big** two those armpit **form**-3S-PAST  
‘He had those two large boils in the armpit.’  
(lit. ‘Those two large boils formed on him in the armpit.’)

It is noteworthy that in this type of construction the verb *ay-* does not carry its most common sense ‘put’, but has a sense roughly translatable as ‘be in a stable condition or form’. That is to say, sense selection in the verb in this case is attributable to occurrence in a particular kind of IEC.

- (b) Onset of a condition or process associated with a bodily product, marked by *ap-* ‘come, appear’

The verb *ap-* ‘come, appear’ occurs both with Condition nouns that refer to stationary, visible bodily conditions (such as those that occur with *ay-*), but also with visible bodily products that are watery and mobile such as *lkañ* ‘blood’, *si-ñg*

‘tears’, *kuñk* ‘saliva’, and *wsb* ‘sweat’. In such constructions, *ap-* indicates the onset of the condition, as in the two following examples.

(36) Yp **si-ñg ow-p.**  
 me **tears come-PF.3S**  
 ‘I feel like crying/I’m ready to cry.’

(37) Yp **sbek ow-p.**  
 me **pimple come-PF.3S**  
 ‘I am getting pimples.’

Except in IECs, *ap-* ‘come’ is an intransitive verb, or at least takes only a Destination as Object, not an Experiencer. In IECs, however, it is transitive and marks the condition as affecting or beginning to affect the Experiencer. This element of meaning is more naturally attributed to the construction as a whole than to the verb alone.

(c) Processes marked by *yap-* ‘fall’, *jak-* ‘rise’, and *tan-* ‘grow, climb’

The verb *yap-* ‘fall, descend’ combines with certain terms for bodily products to indicate a process in which the product falls, actually or metaphorically.

(38) Yp **lkañ yow-p.**  
 me **blood fall-PF.3S**  
 ‘I am bleeding (freely).’ (lit. ‘Blood has fallen on me.’)

(39) Yp **sb yow-p.**  
 me **excrement fall-PF.3S**  
 ‘I need to defecate/I feel like defecating.’

When the body part or product has to do with body wastes (*ss* ‘urine’, *sb* ‘excrement’, *ss-kogi* ‘bladder’), the conventional implicature associated with the construction is that the speaker has become aware of the need for elimination, as in (39). Other visible bodily conditions occur with *jak-* ‘rise’ and *tan-* ‘grow, climb’. As further evidence that the bodily product is the Subject in IEC expressions such as (38) and (39), it is worth noting that if I am walking under a tree and blood or bird droppings fall on me, it would be normal to use (38) to say ‘Some blood has dropped on me’ and (39) to say ‘Some excrement has dropped on me’.

#### 4.2.5. Cases Where the Condition Nominal Is an Atypical Noun

Let us now turn to some cases where the Condition nominal is not a typical noun. These are chiefly expressions where the nominal denotes an intangible bodily or mental condition, such as a sensation or internal state or process.

A number of such nominals occur with *ap-* ‘come, appear’, e.g. *slg* ‘cramp’,

*Argument Structure of Complex Predicates in Kalam*

*suk* ‘laughter’, and *wsn* ‘sleep’. In such expressions, the use of *ap-* indicates that the Experiencer is starting to feel or be affected by the condition (the onset of cramp, wanting to sleep or to laugh, as in (20)).

The largest class of atypical nouns, however, occurs with *g-* ‘happen, do, make, act or work on, affect something’. This verb combines with verb adjuncts, quasi-adjuncts, and adjectives representing several types of Condition. Prominent among these are terms for sensations (or processes that involve sensations) or feelings, e.g. *nabŋ* ‘shame’, *yuan* ‘hunger’, *yuwt* ‘pain’, *yuwt-bt* ‘exhaustion’, *ytuk* ‘lethargy’, *pboŋ* ‘heat, hot’, *ygen* ‘cold (from wind)’, *takl* ‘cold’ (general), *km* ‘bitter’, *ydk* ‘good taste, tasty’, *slk* ‘(i) hot-tasting, pungent, (ii) itching’, *ñekñek* ‘hiccup’, *jlken* ‘cough, head cold’, *tap* ‘sick, sickness’, *kajknm* ‘wince’. Names of sensations are not as nouny as most other Condition nouns. However, they can be modified for manner or intensity, as in (40)-(42).

(40) Yp **takl tmeŋ g-p.**  
me **cold awfully act-PF.3S**  
‘I am terribly cold.’

(41) Yp **nabŋ yob g-a-k.**  
me **shame big do-3S-PAST**  
‘I was very ashamed/shy/embarrassed.’

(42) Yp **tep yb g-a-k.**  
me **good true do-3S-PAST**  
‘I felt truly happy/pleased.’

It is doubtful whether most, if any, Condition nouns denoting sensations and other invisible processes can be possessed. When an accusative pronoun immediately follows such a noun, e.g. (43), it is likely that the pronoun is not the possessor, but belongs to a separate phrase representing the Experiencer and is synonymous with (44):

(43) **Yuwt yp sayn g-p.**  
pain me **weak do-PF.3S**  
‘My pain has eased.’/‘The pain has eased in me.’

(44) Yp **yuwt sayn g-p.**  
me me **weak do-PF.3S**  
‘The pain has eased in me.’

#### **4.2.6. Clauses Where There Are Two Condition Nominals**

Some IEC clauses contain two Condition nominals, neither of which is possessed. In such cases, the second nominal denotes a sensation or other invisible process or condition.

- (45) Yp **sq̣l** **yuwt** **g-s-a-p.**  
 me **boil** **pain** **act-PROG-3S**  
 ‘A boil is causing me pain.’
- (46) Yp **jlken** **mñak** **g-p.**  
 me **cough** **sickness** **do-PF.3S**  
 ‘I am a bit sick with a cough.’

What is the structure of such sequences? Three different interpretations suggest themselves. One is that one nominal modifies the other, forming a complex NP which is the Subject. In (45), this would require *sq̣l yuwt* to be interpreted as ‘(a) painful boil’ and in (46), *jlken mñak* as ‘(a) cough sickness’ (cf. *mñak* ‘non-serious, short-lived sickness’). There is, however, clear evidence against this interpretation, namely that the Experiencer can intervene between the two Condition nominals in IECs and that in neither case do these two nominals form an acceptable complex noun phrase in other contexts.

A second and more plausible interpretation—one that accords with my understanding of what (45) and (46) mean—would treat the first Condition nominal as specifying the cause of the Experiencer’s pain or sickness, answering the question ‘Why are you in pain/sick?’ or ‘What is causing you to be in pain/sick?’. Such questions are normally asked by preposing the interrogative *etp-nen* ‘why?, what for?’ to the single Condition nominal, as in (47a). The cause of a bodily condition may be named as another, more specific ailment, or it may be named, for instance, as a craving for something, as in (47b).

- (47) a. Np **etp-nen** **yuan** **g-p?**  
 you(OBJ) **what-for** **hunger** **act-PF.3S**  
 ‘What are you hungry for?’
- b. Yp **pis-nen** **yuan** **g-p.**  
 me **tinned.fish-for** **hunger** **act-PF.3S**  
 ‘I’m hungry for tinned fish.’

In this analysis, the first Condition nominal is an Adjunct, a peripheral argument. The second Conditional nominal retains its claims to be (an atypical) Subject. However, unlike nouns that denote tangible conditions and bodily products, nouns denoting intangible processes and states must remain close to the verb in constructions with two Condition nouns. The order of the Condition nominals cannot be reversed, i.e. the intangible process/state nominal cannot precede the tangible condition nominal. And whereas a body part nominal can intervene between visible Condition noun and verb, as in (25), (30), and (35), none can occur between invisible process/state noun and verb.

This distributional constraint is consistent with the third possible interpretation of IECs with two Condition nominals, namely that the nominal closest to the

verb functions both as a Subject and as a verb adjunct. In terms of argument structure, it behaves like the Subject because it controls switch reference and fills the Subject slot. But in terms of predicate structure, Condition nouns denoting an intangible process/state behave rather like verb adjuncts, lacking many of the combinatorial privileges of typical nouns, co-occurring with a very small range of verbs, and preferring a position close to the verb.

#### **4.3. Residual Problems**

My brief discussion of transitive IECs has left many details unexplored. In particular, the status of certain Condition nominals remains problematic. The Condition nominal is the strongest candidate for Subject of an IEC because it controls switch reference (and therefore we can also assume it controls Subject agreement, which is formally indeterminate). However, Conditional nominals are a very mixed bag. Some (chiefly those denoting tangible conditions) behave like canonical nouns, and others behave more like verb adjuncts.

Mohanan (1997) has proposed that one class of N+V sequences in Hindi require multidimensional representations to reflect a mismatch between argument structure and grammatical category structure. Similarly, in Kalam IECs there are indications that some Condition nouns manage at the same time to be Subjects, in the domain of argument structure, and verb adjuncts, in the domain of constituent structure. However, these nouns behave differently from true verb adjuncts (§3.3) in a number of respects and we have not established that they form a complex predicate with the verb. Further exploration of these matters will have to left for another occasion.

#### **5. Touching on the Virtues of a Construction-Based Treatment of Kalam Complex Predicates**

The merits of a constructionist treatment of the argument structure of these three types of kinds of construction have been touched on in passing at various points in the discussion. Evidence was noted that the different heads in a complex predicate jointly determine the argument structure of the clause. Another point in favor of the constructionist position is that Kalam verb roots out of context cannot be fully specified for the range of semantic and grammatical properties they show in different contexts. Most common verbs have a number of different senses, each with their own grammatical subcategorizations, and these particularities of sense and subcategorization are associated with particular constructions. For example, although all verbs in IECs occur with two core arguments, some of these verbs are intransitive in other constructions, or at least cannot take an Object which is an Experiencer. This is the case, for instance, with *ap-* ‘come, appear’, *yap-* ‘fall’, and *jak-* ‘stand, rise’. These verbs of motion ‘take on transitivity’ by virtue of occurring in a particular form-meaning constructional pairing in which the Subject slot is filled by a Condition nominal and the Object slot by an Experiencer, and the role of the verb is to indicate roughly that the ‘Condition affects/manifests itself in the Experiencer’.

## References

- Alsina, Alex, Joan Bresnan, and Peter Sells, eds. 1997. *Complex Predicates*. Stanford: Center for the Study of Language and Information.
- Donohue, Mark. 2005. Configurationality in the Languages of New Guinea. *Australian Journal of Linguistics* 25(2):181-218.
- Foley, William. 1986. *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press.
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
- Lane, Jonathan. 1991. Serial Verb Constructions in Kalam. M.A. thesis, University of Auckland. [in press in *Pacific Linguistics*]
- Mohanan, Tara. 1997. Multidimensionality of Representation: NV Complex Predicates in Hindi. In Alsina et al., eds., 431-471.
- Pawley, Andrew. 1987. Encoding Events in Kalam and English: Different Logics for Reporting Experience. In Russell S. Tomlin, ed., *Coherence and Grounding in Discourse*, 329-360. Amsterdam: John Benjamins.
- Pawley, Andrew. In press. Compact versus Serial Verb Constructions in Kalam. In Gunter Senft, ed., *Serial Verb Constructions in Austronesian and Papuan Languages*. Canberra: Pacific Linguistics.
- Pawley, Andrew, and Jonathan Lane. 1998. From Event Sequence to Grammar: Serial Verb Constructions in Kalam. In Anna Siewierska and Jae Jung Song, eds., *Case Typology and Grammar*, 201-227. Amsterdam: John Benjamins.
- Pawley, Andrew, Simon Peter Gi, Ian Saem Majnep, and John Kias. 2000. Hunger Acts on Me: The Grammar and Semantics of Bodily and Mental Process Expressions in Kalam. In Videia P. de Guzman and Byron W. Bender, eds., *Grammatical Analysis: Studies in Honor of Stanley Starosta*, 153-185. Honolulu: University of Hawai'i Press.
- Roberts, John. 2001. Impersonal Constructions in Amele. In Alexandra Y. Aikhenvald, R.M.W. Dixon, and Masayuki Onishi, eds., *Non-Canonical Marking of Subjects and Objects*, 201-250. Amsterdam: John Benjamins.
- Schultze-Berndt, Eva. 2000. Simple and Complex Verbs in Jaminjung. Ph.D. diss., Nijmegen: MPI Series in Psycholinguistics.

Andrew Pawley  
Department of Linguistics, RSPAS  
Australian National University  
Canberra, ACT 200  
Australia

andrew.pawley@anu.edu.au



# Apical Obstruents in Pre-Proto-Tangkic and the Origins of the Non-Zero Absolutive

ERICH R. ROUND

*Yale University*

## 0. Introduction

This paper presents an extract from a larger research project (Round, in prep.) aimed at reconstructing the history of nominal morphology in the non-Pama-Nyungan, Tangkic languages of Queensland, Australia. Below I follow one of the subplots that runs through the long-term evolution of the system, namely the disappearance and then reappearance of non-initial apical oral stops, [t] and [t̪]. The story of word-medial [t̪] is particularly interesting because its reappearance occurs during a sequence of events which culminates in the emergence of the young, non-zero Tangkic absolutive marker, a marker whose precise historical origins have proven elusive for a number of decades now.

After this introduction, the paper is arranged into three main sections. Section 1 outlines some of the key characteristics of the Tangkic languages; section 2 covers the early change in which original apical oral stops are lost from certain positions; and section 3 presents a new account of the emergence of the Tangkic absolutive marker, clarifying the circumstances which led up to it and explaining why the allomorphs which result are those given in (1).

### (1) Stem-conditioned allomorphy of the Tangkic non-zero absolutive

-ja	after [i]	/-ka/	after underlying velar C
-wa	after [u]	/-ta/	after underlying coronal C
-ɟa	after [a]		(but [-a] after /r/)

In existing speculations on the origins of this allomorphy, Evans (1995:138) and Blevins (1997) presume that *-ta* and *-ka* appeared first as part of an emergent vowel-final constraint on words, with the glide-initial allomorphs following later. The actual source of *-ta* and *-ka* has remained a mystery though.<sup>1</sup> Below I provide

---

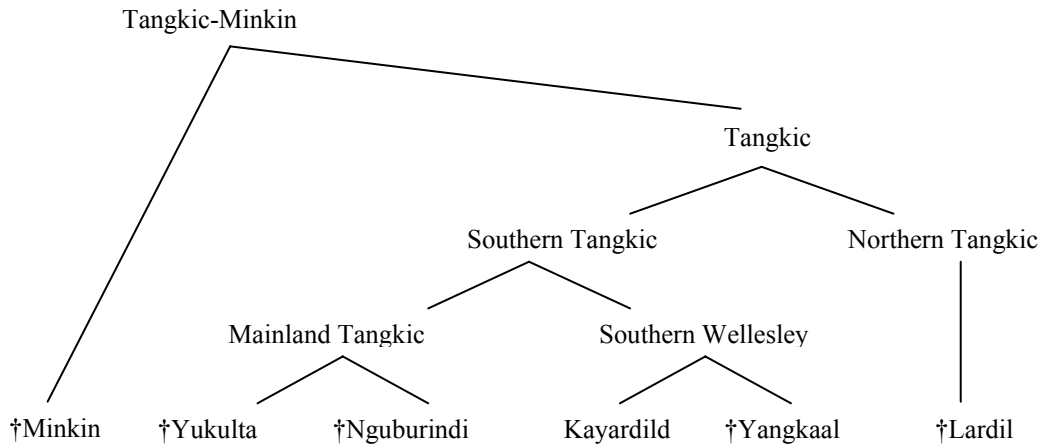
<sup>1</sup> One question is whether Tangkic *-ta* and *-ka* are comparable to the *-pa* augments of languages such as Pitjantjatjara, Inggarda, and Warlpiri (see, e.g., Hale 1973). Here, Evans (1996:137) posits

a source for all of the allomorphs in (1) and offer a revised hypothesis concerning their relative order of appearance.

**1. An Overview of the Tangkic Languages**

(2) presents the postulated genetic tree for the Tangkic-Minkin family (see Evans 1980, 1995:9-13).<sup>2</sup> Most of the changes discussed in this paper are restricted to Tangkic proper.

(2) Postulated genetic tree<sup>3,4</sup>



Given in (3), the phonemic inventory varies only marginally during the reconstructed history of Tangkic-Minkin. The main features are all unexceptional in the Australian context: there are six places of articulation (four coronal; velar; and labial); coronal places divide into natural classes defined by the active articulator

---

an erstwhile discourse particle =*pa* which became an obligatory augment to originally consonant-final words (and thence part of the stem in Warlpiri), and suggests that Tangkic *-tal/-kal/-jal/-wal/-ta* may have had similar roots (as *\*=ka > \*wa / V\_*; then place assimilation). The main problem facing this analysis is its implication for laminal finals, that *\*c=ka* and *\*t=ka > ta*, which is not derivable by assimilation. Now, one could suppose that there were in fact two original clitics, *\*=ka* and *\*=ta*, but although a focus clitic =*ka* does exist in modern Tangkic, there is no attested =*ta*, and worse, one does find =*ta*, in which case, given the general homorganic conditioning of the absolutive allomorphs, it is odd that it should have been *\*=ta* and not *\*=ta* which attached to laminal finals. The present account avoids these problems.

<sup>2</sup> Principal references: Minkin (Evans 1980), Yukulta (Keen 1983), Nguburindi – a word list (Roth 1897), Kayardild and Yangkaal (Evans 1992, 1996), Lardil (Hale 1967, 1975; Klokeid 1976; Ngakulmungan Kangka Leman 1997).

<sup>3</sup> While Minkin has been viewed as related to Tangkic since Evans (1980), the label Tangkic-Minkin is proposed here for the first time.

<sup>4</sup> The Southern Tangkic subgroup has not been rigorously established, but is posited on the basis of lexicostatistics and the fact that only Lardil developed extensive phonological truncation processes. Since the latter observation bases the subgroup on shared retentions rather than shared innovations, a definitive subgrouping of Tangkic remains a task for future research.

*Pre-Proto-Tangkic Apicals and the Absolute*

(apical versus laminal); there is no contrastive voicing, and nasal stops appear at all six places of articulation; and there are three glides, two laterals (both phonemically apical), and a trill. Minor variations are listed at the bottom of (3).

(3) Tangkic phonemic inventory

	apico-		lamino-		velar	labial
	alveolar	post-alveolar	dental	palatal		
obstruent	t	ʈ	ʈ	c	k	p
nasal	n	ɳ	ɳ	ɲ	ŋ	m
trill	r					
lateral	l	ɭ				
glide		ɹ		j		w

vowels

i, i:	u, u:
a, a:	

Minor variations:

- possible original \*/ɭ/, \*/ɻ/ collapse with /l/, /ɹ/
- late merger of /ɭ/ and /ɹ/ in Lardil and Kayardild
- long vowels may at one point have been absent
- late development in Lardil of /e/, /e:/
- on distributional grounds, it could be said that /t/ and /ɾ/ merged for a short period

Turning from the underlying inventory to the phonological constraints operating in the languages, the most important are as follows.

Formulated by Hamilton (1989), HAMILTON'S GENERALIZATION states that for any consonant cluster, the sequence of segments is strictly ordered: apicals are first, followed by laminals, then velars, and then labials, cf. (4a). This constraint is found in many Australian languages; it is unviolated in all the of modern Tangkic languages, and there is no reason to reconstruct any period at which this was not also the case.

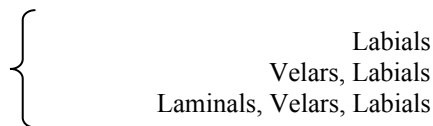
There are two corollaries of Hamilton's generalization which are relevant in this paper. The first is that deletions, whether seen as diachronic correspondences or synchronic processes, fall into the natural classes given in (4b). Consequent to the first is the second, long-term diachronic corollary—that inventories of phonotactically attested or permissible final consonants (whether word-final or stem-final) also fall into natural classes, given in (4c).

(4) Hamilton's generalization (Hamilton 1989) and its corollaries

a. *Order must be:* Apical > Laminal > Velar > Labial

b. *Immediate corollary:*

Natural classes of deletions:



Erich R. Round

c. *Long-term diachronic corollary / Typological generalization:*

Most common inventories of stem- & word-final consonants:	}	Apicals Apicals, Laminals Apicals, Laminals, Velars Apicals, Laminals, Velars, Labials
---	---	---

The only other constraints which will concern us here are given in (5a,b): there are no geminates, no glides in codas, and no reason for reconstructing them. The constraints are satisfied as in (5c).

(5) Other key constraints

- a. *\*Geminate:*  $/C_{\alpha} + C_{\alpha}'/ > [C_{\alpha}]$
- b. *Ordering by manner/sonority*  
 Order must be: 

Liquid / Nasal > Glide	}	∴ no glides in codas
Liquid > Nasal > Obs		
- c. *Satisfaction of constraints*  
 Achieved by:
  - i. regressive assimilation of Nasal + Obs
  - ii. deletion of C from right-edge of morpheme  $M_1$  in  $/M_1+M_2/$

Turning to morphology, what we are concerned with here are nominals, a part of speech in Tangkic that covers roughly the same functional range as English nouns and adjectives. In (6) I list the morphological environments in which one finds nominal stems in Tangkic. I will refer below particularly to nominal and verbal compounds. Nominal compounds (6c) are right-headed except if they end in the morpheme *tajka* (the root from which Tangkic draws its name). In verbal compounds, a nominal stem precedes a verbal stem (6d).

(6) Morphological distribution of nominal stems

- a. *Simple inflected stems:* STEM<sub>N</sub>-INFL
- b. *Case stacking:* [STEM<sub>N</sub>-INFL]-INFL...  
 e.g. Kayardild  $[[tajka-kara]-\eta uni]-jca$   
 $[[man-GEN]-INSTR]-OBL$
- c. *Nominal compounds*
  - i. N<sub>HEAD</sub>+N
 

e.g. Kayardild	<i>ma<sub>t</sub>-cu<sub>ɲ</sub>ara</i>	‘big handed’
	hand-big	
Yukulta	<i>mankar-pa<sub>t</sub>a<sub>ɲ</sub>u</i>	‘well built’
	body-big	
Lardil	<i>/mil-ki<sub>t</sub>i<sub>k</sub>i<sub>t</sub>i/</i>	‘halo around moon’
	edge-moon	
  - ii. N + /*tajka*/<sub>HEAD</sub>

e.g. Kayardild	<i>kunpa-tajka</i>	‘person from Kunba’
	Kunba-person	
Lardil	<i>/julma-tajka/</i>	‘avoided person’
	restricted-person	

*Pre-Proto-Tangkiic Apicals and the Absolutive*

d. <i>Verbal compounds</i>	e.g.	Kayardild	<i>tami-ɬa:c-</i> blunt-spear	‘spear blunt’
		Yukulta	<i>ɟaʃil-karmac-</i> hip-grab	‘have hands on hips’
		Lardil	<i>/maɬ-kupa/</i> hand-become.good	‘take care of’

To complete the overview of Tangkiic, the final phenomenon of direct relevance to the reconstruction below is Tangkiic ‘truncation’ (Hale 1973). Although the synchronic process of truncation is most famously known from Lardil, it can also be understood as a general Tangkiic characteristic involving: (i) deletion of a final vowel, and (ii) the consequent deletion of phonotactically banned final consonants. The parameters for the modern languages are given in (7a), and what I reconstruct for Proto-Tangkiic in (7b); note that all of the modern languages have simplified the ancestral system in some way, whether by restricting it or generalizing. I will be proposing below that the young Tangkiic absolutive marker emerged during a series of reanalyses of this system, including rule inversion.

(7) Tangkiic truncation and its parameters

	Domain of truncation	Deleting vowels	Illegal final Cs <sup>†</sup>			
a.	Lardil	word-final a, i, u	<table style="border-left: 1px solid black; border-right: 1px solid black; border-collapse: collapse;"> <tr> <td style="padding: 0 5px;">clusters</td> </tr> <tr> <td style="padding: 0 5px;">w, j; p, k, c, ɬ; m, ŋ, ɲ, ɳ</td> </tr> <tr> <td style="padding: 0 5px;">– i.e. all non-apicals</td> </tr> </table>	clusters	w, j; p, k, c, ɬ; m, ŋ, ɲ, ɳ	– i.e. all non-apicals
clusters						
w, j; p, k, c, ɬ; m, ŋ, ɲ, ɳ						
– i.e. all non-apicals						
	Kayardild	phrase-final word-final a, only after:	none* w / u__ (i.e. /uwa/ > [u] / __#)			
	Yukulta	DOES NOT OCCUR	—			

\* There is good evidence from frozen forms that p was illegal at an earlier stage.

† Old, word-final \*ŋ, \*ɲ, \*ɳ had been irregularly lost by Proto-Tangkiic, \*m completely so.

Proposal:

b.	Proto-Tangkiic	phrase-final (obligatory) word-final (optional)	a	w, j, ɟ; p, k; m, ŋ – i.e. all glides & non-coronals
----	----------------	--	---	---

**2. The Early Loss of Final Apical Stops from Tangkiic Nominal Stems**

Apical oral stops are synchronically absent from the final position in Tangkiic nominal stems, and I propose this absence is due to historical loss at an early stage. This presupposes that modern Tangkiic languages do in fact lack stem final /t/ and /tʰ/; hence, I address first some early analyses of Lardil and Yukulta in which this was not taken to be the case.

All modern Tangkiic languages have stems whose absolutive (or nominative)<sup>5</sup> form ends in [ta] and whose other case inflections end in some laminal plus other

<sup>5</sup> The nominative case of Lardil and Kayardild is cognate with the Yukulta absolutive, and all derive from the earlier Proto-Tangkiic absolutive (Evans 1996:423ff).

material, as illustrated in (8). The original analyses for both Lardil and Yukulta were that these stems underlyingly end in apical stops which undergo a process of ‘laminalization’ outside of the absolutive/nominative.

(8) The underlying-apical analysis of Lardil, Yukulta

Lardil	{	a.	/ɲut/~ɲit/	ɲuta	/ɲut + a/	<table border="0" style="width: 100%; text-align: center;"> <tr> <td style="width: 50%;">LOC</td> <td style="width: 50%;">FUTURE.OBJ</td> </tr> <tr> <td>ɲic-e</td> <td>ɲit̪-uɽ</td> </tr> <tr> <td>/ɲit + e/</td> <td>/ɲit + uɽ/</td> </tr> <tr> <td>pic-e</td> <td>pit̪-uɽ</td> </tr> <tr> <td>/pit + e/</td> <td>/pit + uɽ/</td> </tr> </table> <p>laminalization: /t/ &gt; [c] / __i, e; &gt; [t̪] / __a, u</p>	LOC	FUTURE.OBJ	ɲic-e	ɲit̪-uɽ	/ɲit + e/	/ɲit + uɽ/	pic-e	pit̪-uɽ	/pit + e/	/pit + uɽ/
		LOC	FUTURE.OBJ													
ɲic-e	ɲit̪-uɽ															
/ɲit + e/	/ɲit + uɽ/															
pic-e	pit̪-uɽ															
/pit + e/	/pit + uɽ/															
b.	/pit/	pita	/pit + a/													
(cognate examples)																
Yukulta	{	c.	/ɲit/	ɲita	/ɲit + a/	<table border="0" style="width: 100%; text-align: center;"> <tr> <td style="width: 50%;">ERG/LOC</td> <td style="width: 50%;">COMITATIVE</td> </tr> <tr> <td>ɲic-iya</td> <td>ɲic-u u</td> </tr> <tr> <td>/ɲit + iya/</td> <td>/ɲit + u u/</td> </tr> <tr> <td>pit̪-iya</td> <td>pit̪-u u</td> </tr> <tr> <td>/pit + iya/</td> <td>/pit + u u/</td> </tr> </table> <p>/t/ &gt; [t̪], [c] idiosyncratically, dependent on stem</p>	ERG/LOC	COMITATIVE	ɲic-iya	ɲic-u u	/ɲit + iya/	/ɲit + u u/	pit̪-iya	pit̪-u u	/pit + iya/	/pit + u u/
		ERG/LOC	COMITATIVE													
ɲic-iya	ɲic-u u															
/ɲit + iya/	/ɲit + u u/															
pit̪-iya	pit̪-u u															
/pit + iya/	/pit + u u/															
d.	/pit/	pita	/pit + a/													

As indicated by the arrows in (8), the analysis holds water for Lardil because the laminal of any stem’s non-nominative forms is predictable from the following vowel. In Yukulta, though, this is not the case: laminals are lexically idiosyncratic. A superior analysis by Evans (1995) is that for Yukulta, and also for Kayardild and proto-Tangkic,<sup>6</sup> it is the laminal which is underlying (as in (9)). Evans replaces the laminalization process with de-laminalization; however, even this is not needed. By Hamilton’s generalization, an underlying laminal + apical cluster undergoes deletion of the laminal: the alternation is predicted by the regular phonology of the languages.

<sup>6</sup> The underlying-apical analysis succeeds in Lardil because the two original stem-final laminals have collapsed, removing the lexical contrast. Even in Lardil, though, it seems possible to posit underlying laminals and derive ‘delaminalization’ from nominal /-ta/ and Hamilton’s generalization (see also Hale 1973:437ff for the inconclusiveness of one argument from analytical elegance for underlying /t/; on Hale’s appeal to facts in Damin, see further Round, in prep.).

*Pre-Proto-Tangkic Apicals and the Absolutive*

(9) Evans's (1995) analysis [modified to incorporate Hamilton's generalization]

Lardil	{	a. /ɲuc/~ɲic/ 'firewood'	ɲuta /ɲuc + ta/	ɲic-e /ɲic + e/	ɲiɰ-uɰ /ɲic + uɰ/
		b. /pic/ 'smell'	pita /pic + ta/	pic-e /pic + e/	piɰ-uɰ /pic + uɰ/
Yukulta	{	c. /ɲic/ 'firewood'	ɲita /ɲic + ta/	ɲic-ia /ɲic + ia/	ɲic-uɰ /ɲic + uɰ/
		d. /piɰ/ 'smell'	pita /piɰ + ta/	piɰ-ia /piɰ + ia/	piɰ-uɰ /piɰ + uɰ/

Admitting this analysis, the inventory of final consonants for nominal stems reconstructed for Proto-Tangkic is given in (10).

(10) Final consonants/clusters of nominal stems in Proto-Tangkic

	apical	laminal	velar
obstruent		t̚      c	k
nasal	n      ɲ		ŋ
liquid	r, l		
clusters			rk, lk, [k; rŋ, lŋ]

Note that in relation to the typologically normal inventories of final consonants given earlier in (4c), the Tangkic inventory exhibits gaps: it lacks apical oral stops, it lacks laminal nasals, and arguably non-velar final clusters. In this paper, I focus on the oral stops.<sup>7</sup>

Since we aim, all things equal, to reconstruct typologically regular proto-systems over irregular ones, an obvious question is: can one find evidence that the stem-final apical oral stops were lost at some stage? The answer is yes.

In (11), I give the developments which removed original stem-final /t/ and /t/ from the system. None of these individual changes is particularly remarkable: apical stops flap after vowels; flaps fall together with trills if no vowel follows; and the remaining apico-alveolar (non-retroflex) flaps fall together with the trill.

<sup>7</sup> However, see (16) for some indication of how the disappearance of the laminal nasals is precipitated by the accretion of the *-ta* absolutive suffix. The reconstructed history of cluster-final stems is complex and cannot be treated here, but see Round (in prep.).

(11) Changes resulting in the loss of nominal stem-final /t/ and /t̥/

	t / V__V	t / V__ {#,C}	t / n__	t̥ / V__V	t̥ / V__ {#,C}	t̥ / n__
1. post-vocalic flapping	t > r	t > r		t̥ > ɾ	t̥ > ɾ	
2. flap > trill / __ {#,C}		r > r			ɾ > r	
3. [r] > trill [r]	r > r					
4. [ɾ] > [t̥]				ɾ > t̥		
outcome	$\underbrace{t > r \quad t > r \quad \text{no change}}_{t > r / V\_}$			$\underbrace{\text{no net change} \quad t̥ > r \quad \text{no change}}_{t̥ > r / V\_ \{#,C\}}$		

[r] apico-alveolar flap; [ɾ] apico-postalveolar (retroflex) flap; [r] apico-alveolar trill

These changes leave behind certain traces, as presented in (12a-d). The best evidence is found in compounds given in (12c,d), taken from modern Kayardild.

(12) Evidence in modern Tangkic for loss of final /t/ and /t̥/

a. Distribution within the morpheme:

	<i>initial</i>	<i>medial</i>	<i>final</i>
t	yes	almost never	no
t̥	yes	yes	no

b. Relative frequencies of root-finals: [r]-finals preponderate.

c. Individual lexical evidence for VrV < \*Vr # tV or \*Vt # tV

- i. *muniric-* ‘put to one’s breast’ cf. *munir* ‘breast’; *-tic* ‘sit’
- ii. *ɟaric* ‘lay (egg)’ cf. *ɟar-* ‘low; on the ground’; *-tic-* ‘sit’

d. Individual lexical evidence for VrV < \*V # tV

- i. *malarayka* ‘sea-object’ cf. *mala* ‘sea’; *tayka* ‘thing’<sup>8</sup>
- ii. *walmuric-* ‘look up’ cf. *walmu* ‘up; high’; *-tic-* ‘sit’

The changes to be considered next are reconstructed as having occurred after the Minkin branch splits from Tangkic, and are restricted to the latter. I also skip over a series of changes reconstructed for pre-Proto-Tangkic (see Round, in prep.) in which final segments were lost in a number of domains, summarized in (13).

<sup>8</sup> The element *tayka*, which in modern Tangkic languages generally means ‘person, man’, is reconstructed here as ‘thing’. In support of *tayka* not originally meaning ‘man’: (i) the Minkin lexemes glossed as ‘man’ fail to include *tayka*, even though *maku* ‘woman’ is identical to Tangkic; (ii) Tangkic ‘female’ is *maku*, but ‘male’ is not *tayka* but *\*wu[ka]*, a potential cognate of which is found in Minkin (*wurar* ‘boy’, with correspondences *k* :  $\emptyset$  / LIQUID\_\_ and *l* : *r/l*). In support of *tayka* originally meaning ‘thing’: (i) cf. *malarayka* above; (ii) *tayka* still carries this meaning in certain constructions in Kayardild:

Kayardild: *Dathina riinda [jungarra dangkaa] kalaj*. ‘There’s a [big one] flying in the east.’  
*[Ngaaka dangkaa] dathin?* ‘[Which one] is that?’ (pointing to a plant)



(13) Right-edge deletions during early Proto-Tangkic

<i>Level of regularity</i>	<i>Domain</i>	<i>Segments lost</i>	
complete loss	word-final	labials	p, m
productive, postlexical process	phrase-final	velars	k, ŋ
irregular loss, lexicalized	word-final	non-apical nasals	ŋ, ɲ, ŋ

Note that at this stage, it will only take the deletion of phrase-final [a] to set up precisely the pattern of truncation which was proposed for Proto-Tangkic in (7b).

**3. Truncation, the Absolutive, and the Return of Word-Medial /t/**

In this final section, I detail a novel proposal for the source of the Tangkic absolutive. Entering the earliest period of truncation, the inherited, Proto-Tangkic-Minkin absolutive case marker was zero, i.e. absolutives were bare stems.<sup>9</sup> As suggested above, truncation proper is set in motion by the loss of phrase-final short [a] in words of three or more syllables, with the concomitant loss, due to phonotactic constraints, of glides (cf. (5)) and of labials and velars (labials by word-level phonotactics, velars by phrase-level; see (13) above). Given this situation, it is not uncommon for entire syllables to be truncated phrase-finally. In (14) are listed what would have been the most frequently truncated syllables, together with their most frequent, leftward phonological environments.

(14) Most frequent occurrences of each type of deleted syllable

-ja	after	i	in the ergative/locative case marker /-kija/
-wa	after	u	in the absolutive of long stems ending in /-uwa/
-ɟa	after	a	in the absolutive of 'north' /cirkaɟa/ <sup>10</sup>
-pa	after	r	in the case marker and derivational suffix /-(ŋ)arpa/
-(ŋ)ka	after	V	in the absolutive of long stems ending in /-(ŋ)ka/

Soon hereafter, major syntactic changes began, restructuring biclausal structures into single clauses (Evans 1995). Prosodic restructuring would have ensued, with erstwhile multiphrasal units collapsed into single phrases.<sup>11</sup> As a residue, some truncations now occur at particular, syntactically defined positions which are not

<sup>9</sup> Neither modern Tangkic nor Minkin preserve this reconstructed Tangkic-Minkin absolutive form (Minkin appears to append a phonological augment [-a] to consonant-final stems, cf. Evans 1980:187-8); however, considerable evidence can be mobilized in favor of reconstructing the absolutive as a bare stem (Round, in prep.).

<sup>10</sup> The lexeme *cirkaɟa* 'north' appears in the list here because in Tangkic, as in many other Australian languages, cardinal directions are employed at all scales of spatial reckoning; hence, words for 'north', 'south', 'east', and 'west' have a very high discourse frequency.

<sup>11</sup> See Fletcher et al. (2002) for the typologically marked status of Kayardild intonation in Australia. By my own observations, its distinctness is shared by Lardil; these unusual features would find a natural explanation in their having arisen diachronically through the restructuring of multi-phrasal, edge-marked units into single phrases with mobile prominence marking.

final in the phrase, but medial, and when further syntactic change renders their original conditioning opaque, these phrase-medial truncations are reinterpreted simply as optional.

The next development is a reanalysis of optional truncation from a deletion of syllables to a simple alternation between long and short forms; that is, truncation originally creates shorter forms of long words, but speakers initiate a reverse process which creates longer forms for short words by augmenting them with unetymological syllables. The source of these new ‘augments’ is the set of most frequent, original truncated syllables listed above in (14), but with one further restriction: only those syllables are adopted which form augments homorganic with their phonological stem (that is, [-pa] is not generalized as an augment after [r]). Examples from this new system are given in (15).

(15) Old effects of truncation, new effects of augmentation

Truncation	/tulkija/	:	[tulki]	<	[tulki-ja]
	/kuwakuwa/	:	[kuwaku]	<	[kuwaku-wa]
	/cirkaɟa/	:	[cirka]	<	[cirka-ɟa]
	/malaraŋka/	:	[malara]	<	[malara-ŋka]
Augmentation	/piɽi/	:	[piɽi]	>	[piɽi-ja]
	/maku/	:	[maku]	>	[maku-wa]
	/wara/	:	[wara]	>	[wara-ɟa]
	/yaraŋ/	:	[yara]	>	[yara-ŋka]

Note that the only stems which now lack augments are coronal-finals (there being no labial-final stems at this point); also, other than this gap in the paradigm, the form of the augment is already identical to what will soon become the Proto-Tangkic absolutive.

What I propose happened next was that the gap in the augment paradigm was filled, specifically by co-opting a truncated form of *taŋka* ‘thing’ which was found in productive, left-headed nominal compounds ((6c); fn. 8 above). Note that since such compounds would necessarily comprise at least three syllables, final /-taŋka/ would always be truncated to [-ta]. In compounds where the first element was attributive, the semantics of this [-ta] would be almost empty, as in (16). Generalizing from such constructions,<sup>12</sup> this ‘meaningless’ [-ta] was spread to fill the gap in the augment paradigm.

<sup>12</sup> The key context would have been predicative nominal clauses such as \*[maku-wa] [kalkanta]: [woman-AUGMENT] [sick(one)] ‘The woman is (a) sick (one)’.

(16) Truncation of N<sub>ATTRIBUTIVE</sub>+*taŋka*:

/kalkan-taŋka/	>	[kalkan-ta]	‘sick (one)’	with /r + t/ > [r] as a relic of the changes in (11) exhibiting effects of Hamilton’s generali- zation
/campaŋ-taŋka/	>	[campaŋ-ta]	‘hollow (one)’	
/tiral-taŋka/	>	[tiral-ta]	‘slippery (one)’	
/tawu-taŋka/	>	[tawu-ta]	‘new (one)’	
/julmpur-taŋka/	>	[julmpur-a]	‘long (one)’	
/ba-taŋka/	>	[ba-ta]	‘(one) in the west’	
/waŋji:c-taŋka/	>	[waŋji:-ta]	‘strange (one)’	exhibiting effects of Hamilton’s generali- zation
/niwaŋ-taŋka/	>	[niwan-ta]	‘his (one)’	

With the adoption of [-ta] as the augment to all coronal-final stems, the full allomorphy of the non-zero absolutive was in now place. The only changes left to occur were those in which optional augment syllables became obligatory (i) after all mono- and di-syllables, a natural strategy for stress clash avoidance in a language where primary stress was predominantly word-initial, and (ii) after all consonant-final stems, making every word vowel-final.

As a final observation, we can note that this hypothesis on the origins of the Tangkiç absolutive explains not only the reasons for the distinctive allomorphy of the marker, but also two further quirks of its synchronic behavior. In the first of these, the Tangkiç absolutive always follows, without any semantic content, any consonant-final case suffix which otherwise would stand at the right edge of the word, i.e. it still acts as a phonological augment (Evans 1995:136-7).

The second quirk surfaces in certain Kayardild nominative NPs<sup>13</sup> (the Kayardild nominative being the reflex of the Proto-Tangkiç absolutive), namely those NPs with the form [STEM<sub>N,ATTRIBUTIVE</sub> *dangka-a*]. If this unit were a single word, then it would be over two syllables long and the final element should be the zero-marked *-dangka* (< \**taŋka*), not *dangka-a* (< \**taŋka-ŋa*). Nevertheless, the first element is entirely uninflected, as if the whole unit really were a compound. I have only encountered such NPs where the first STEM is coronal-final. Synchronically this is hard to explain, and diachronically it only receives a coherent explanation if Tangkiç absolutive [-ta] does derive from /-taŋka/. In the latter case, these odd NPs in Kayardild with their zero-marked, coronal-final absolutive STEM can be seen as preserving the old situation, before [-ta] was incorporated into the augment system,<sup>14</sup> and they do so in the one construction into which [-ta] was not extended, but in which it was original, as /taŋka/.

#### 4. Summary

Apical oral stops were lost from final position in nominal stems during Proto-Tangkiç-Minkin in a merger with trilled /r/. Apico-alveolar /t/ was also lost word-medially, but reappears late in Proto-Tangkiç in the phonological augment

<sup>13</sup> Ambiguities in the morphological status of N+/*taŋka*/ also arise in Lardil (Klokeid 1976:68-72).

<sup>14</sup> Although the original construction, as one long word, would not have borne the marker \**-ŋa* at the end.

syllable [-ta]. This [-ta] is in fact the last piece of an otherwise complete paradigm of augments which had recently arisen out of the partial, diachronic inversion of the Tangkic truncation rule, a paradigm which later becomes the Tangkic absolutive. The syllable [-ta] was itself derived by regular truncation from the morpheme /taŋka/ ‘thing’, appearing in left-headed nominal compounds  $N_{\text{ATTRIBUTIVE}^+}/\text{taŋka}/$ , where its semantic contribution was close to zero.

### References

- Blevins, Juliette. 1997. Rules in Optimality Theory: Two Case Studies. In Iggy Roca, ed., *Derivations and Constraints in Phonology*, 227-60. Oxford: Oxford University Press.
- Evans, Nicholas. 1990. The Minkin Language of the Bourketown Region. In Geoffrey N. O’Grady and D. T. Tryon, eds., *Studies in Comparative Pama-Nyungan*, 173-207. Canberra: Pacific Linguistics.
- Evans, Nicholas. 1992. *Kayardild Dictionary and Thesaurus*. Melbourne: University of Melbourne Linguistics Department.
- Evans, Nicholas. 1995. *A Grammar of Kayardild: With Historical-Comparative Notes on Tangkic*. Berlin: Mouton de Gruyter.
- Fletcher, Janet, Nicholas Evans, and Erich Round. 2002. Left-Edge Tonal Events in Kayardild (Australian): A Typological Perspective. In B. Bel and I. Marlien, eds., *Proceedings of the First International Conference on Speech Prosody*, 295-8. Provence: Laboratoire Parole et Langage.
- Hale, Kenneth. 1967. Some Productive Rules in Lardil Mornington Island Syntax. *Papers in Australian Linguistics*, No. 2, 63-73. Canberra: Pacific Linguistics, A-11.
- Hale, Kenneth. 1973. Deep-Surface Canonical Disparities in Relation to Analysis and Change: An Australian Example. In T. A. Sebeok, ed., *Current Trends in Linguistics, Vol. XI: Diachronic, Areal and Typological Linguistics*, 401-58. The Hague: Mouton de Gruyter.
- Hamilton, Philip. 1989. Australian Phonotactics and the Internal Structure of the Place Node. M.A. thesis, University of Toronto.
- Hamilton, Philip. 1996. Phonetic Constraints and Markedness in the Phonotactics of Australian Aboriginal Languages. Ph.D. diss., University of Toronto.
- Keen, Sandra. 1983. Yukulta. In Robert M. W. Dixon and Barry J. Blake, eds., *Handbook of Australian Languages, Vol. 3*, 190-304. Amsterdam: Benjamins.
- Klokeid, Terry J. 1976. Topics in Lardil Grammar. Ph.D. diss., Massachusetts Institute of Technology.
- Ngakulumungan Kangka Leman (compilers). 1997. *Lardil Dictionary*. Gununa, Queensland: Mornington Shire Council.
- Roth, W. E. 1897. Obarindi (Word List). In uncatalogued MS216, Mitchell Library, Sydney.
- Round, Erich R. In preparation. A Reconstruction of Pre-Proto-Tangkic Nominal Morphology.

*Pre-Proto-Tangic Apicals and the Absolute*

Erich R. Round  
370 Temple Street, Room 204  
PO Box 208366  
New Haven, CT 06520-8366

[erich.round@yale.edu](mailto:erich.round@yale.edu)