Grammatical Analysis

Morphology, Syntax, and Semantics

Studies in Honor of Stanley Starosta

Edited by

Videa P. De Guzman and Byron Bender

Grammatical Analysis

Grammatical Analysis

Morphology, Syntax, and Semantics

Studies in Honor of Stanley Starosta

Edited by

Videa P. De Guzman and Byron W. Bender

University of Hawai'i Press

Honolulu

Camera-ready copy for this book was prepared under the
supervision of the series editor.

CONTENTS

EDITOR'S NOTE

Volumes such as this serve to mark, for all to see, a relationship that has existed between a scholar and those who have encountered that scholar's ideas and person. Some may have sat at the scholar's feet, others down the hall or across the conference room, but for each of us represented in this volume there has occurred a significant exchange of ideas. Stanley Starosta came to the University of Hawai'i in 1967, but has remained ever since a most peripatetic scholar. Germany, India, Australia, Scotland, Thailand, Taiwan, and Japan are some of the more exotic places where he has spent an appreciable amount of time (while always maintaining his Hawai'i connection). A listing of appointments (domestic and foreign), conferences attended, and papers read would cover altogether too many pages to list here.

For those who have been exposed to his work, what has been seen of note? A call to make Linguistics a true science, whose results are as equally replicable as those of, say, physics (Stan's undergraduate major). Definitions that work and that do not go round in circles. Entities and relations that can be stated formally, and are, to such an extent that the formalisms become a language of their own. The continual testing of those formalisms against a wide variety of languages (an endeavor in which each student becomes a partner). A minimum of abstractness in the structures posited between form and meaning, leading to skepticism of unnecessary or empty categories. One former student describes her early experiences as follows:

> I'd come from a masters degree at McGill, eager to sink my analytical teeth into Pacific languages again after working with indigenous Canadian languages for three years. One of my courses that first semester was [one in which] Stan was teaching us lexicase theory and we were using it to solve syntax problems. . . . Each week the students were given a new set of data to analyze. Oh, how I struggled with Swahili noun classes. But the process that led to the answers was so rewarding. We discovered the truth of the data, poring over sentences and arguing heatedly with each other and with Stan, too. To me it was the best part of being in graduate school. . . . That semester Stan presented the fundamentals of theories in general and Lexicase in particular:
>
> • The more powerful the theory, the less explanatory value it has.
> • All theories should be falsifiable.
> • There is only one level of syntactic structure—the surface level.
> • Syntactic relationships between constituents in a sentence can be explained in terms of dependency relations.
> • Relationships between syntactically-related sentences can be explained by derivational relationships between verbs.
> • Derivations create new lexical items in categories that always include root forms.
> • All lexical items can be identified as members of a limited, universal class of parts of speech.
>
> Lexicase was and is much more than these few principles, and since then, the theory has developed in many ways as is attested by the analyses presented elsewhere in this volume. Nevertheless, these few principles are sufficient to establish a very limited set of parameters for analysis. What these principles do, in fact, is hold the lin-

guist accountable to the data. There are no opportunities to perform sleight-of-hand by positing some abstract elements or underlying representations.

For the [lexicase grammarian], the theory determines a narrow range of possible explanations. The limited possibilities were, for me, a gift. . . . Given that I was wearing "lexicase lenses," the only place to look for answers was at the data itself. The constraints of the theory and its formalisms limited the possible analyses and helped me to reach the answers. At each turn in the data, I was surprised at what the theory predicted. It helped me to know where to look for the answers. If I found a verb with certain features, then given the formalism of the theory, there should also be verbs without that feature: for every "plus" feature, there had to be a minus. This bifurcation of classificatory features forced me to look carefully at the data, and each time I was rewarded. There were indeed verbs to fill in every node on the branches. . . .[1]

The classwork here reported occurred in the mid-eighties. Lexicase has continued to develop since then. But its development has been consistent and in keeping with these earlier principles. And efforts along these lines have not precluded new endeavors. I have personally been impressed with how much of Stan's work in the past decade has been in morphology rather than in syntax, where he has been using his bent for formalisms to sharpen and answer a number of questions that concern whole-word morphology.[2]

It is with great pleasure that the editorial board of OLSP publishes this collection as a tribute to the work of Stanley Starosta.

Byron W. Bender
General Editor
OLSP

---

1. Louise Pagotto, "On being a Lexicasualty [1998]" (Kapi'olani Community College, photocopy), 1–2.

2. Much of this is still unpublished but in press. I have seen advance copies of parts, and also had the good fortune to sit in on a class on word formation he taught in Spring 1995. My own paper in this volume owes much to the stimulation of his 1992 talk, included in the references.

PREFACE

This volume is dedicated to Professor Dr. Stanley Starosta of the University of Hawai'i. We, the contributors, wish to commemorate his remarkable influence on us in our respective academic careers and/or continuing research work.

The contents of this book are to a large extent representative of Stan's enduring interest in grammatical theory and its implications for the various aspects of syntax and the lexicon. Part 1 deals with some theoretical, historical, or pragmatic aspect of linguistic analysis. Part 2 includes essays on semantics, morphology, and/or syntax. The languages treated in the various grammatical analyses range across the different groups of the Austronesian family and extend to a few non-Austronesian languages, reflecting the individual interests and expertise of the contributors. In chapter 1, O'Grady proposes a syntactic representation in his version of categorial grammar (compared with Starosta's Lexicase dependency grammar) that accommodates principles that account for such linguistic phenomena as agreement, anaphora, and negative polarity. Bender (chap. 2) addresses the question of the formal role of paradigms in a whole-word approach to morphology. He shows how this can provide an explanation for why morphological solutions prevail over phonological solutions when the two are in conflict. Two essays on grammaticalization from a historical (chap. 3) and a pragmatic perspective (chap. 5) are presented. Reid's chapter 3 applies Starosta's lexicase dependency grammar framework, showing that the development of at least some of the prenasalized obstruents in word-initial position in Proto-Oceanic resulted from a process of grammaticalization in the structure of the noun phrase. Yang's chapter 5 offers both cognitive and syntactic arguments to support the case of the emerging particle *poko* in Korean. In chapter 4, Clark explores the primary deictic and anaphoric concepts—proximal and distal—largely illustrated in Hmong and Vietnamese, and contends that they parallel "prelinguistic" (i.e., cognitive) processes. In chapter 6, Kullavanijaya shows how a personal pronoun in Thai may have a contradictory usage arising from semantic features that have developed through time.

In Part II, Stan's indelible stamp is clearly visible as his lexicase framework is applied analytically to account for specific problems in lexical categorization in Thai (Prasithrathsint's chapter 9; Savetamalya's chapter 10). Likewise, it provides the framework for analyzing syntactic constituents in Thai such as nonverbal predicates (Indrambarya's chapter 12) and double objects (Wilawan's chapter 13), and also verb agreement in Shoshoni, a Central Numic language of the Uto-Aztecan family (Lindsey's chapter 17).

The remaining chapters involve aspects of grammar in particular languages that are also close to Stan's research interests. One such language group is the Formosan languages. Chapter 7 by Li expounds on the focus, case, pronoun, and aspect systems of a nearly extinct language known as Pazeh. He points out some similarities and differences between Pazeh and the other Formosan languages. Tsuchida in chapter 8 identifies lexical prefixes in Siraya, an extinct language of Formosa, in relation to various classes of verbs. The two other studies on Formosan languages

are Zeitoun's chapter 15 and Yeh's chapter 16. The former investigates one type of possessive construction in ten Formosan languages and shows that where two *have*-structures exist, they differ not only semantically but also syntactically. The latter study examines the eight different forms of negators in Saisiyat in terms of their distribution and function and the semantic features that influence the selection of introducers of cooccurring predicates.

Finally, two essays that fall within the sphere of Stan's interest in grammatical and case/semantic relations are Pawley, Gi, Majnep, and Kias's chapter 11 and De Guzman's chapter 14. Pawley et al. examine the way bodily and mental processes are represented in Kalam, a language of Papua New Guinea. De Guzman presents morphological and syntactic evidence to support the grammatical distinctions that must be made between the two identically marked nominal phrases—the nonabsolutive agent phrase and the nonabsolutive patient phrase—in Tagalog, a major language of the Philippines.

I want to acknowledge the help of some of the contributors and other colleagues—Professors E.-D. Cook, Ray Freeze, and Robert Murray—in reviewing the essays. I also want to thank the following: Dr. Josie P. Clausen and Wendy Onishi for helping me at the outset of this project track down all of Stan's advisees; Aleli, Stan's wife, for providing other relevant materials; and Ritsuko Kikusawa for helping us find a suitable photograph. Lastly, my deep appreciation and gratitude go to Professor Byron W. Bender for agreeing to be coeditor of this special publication. Without his expertise in computer work as well as in editing, coupled with unselfish dedication, this volume would not have become a reality.

Videa P. De Guzman
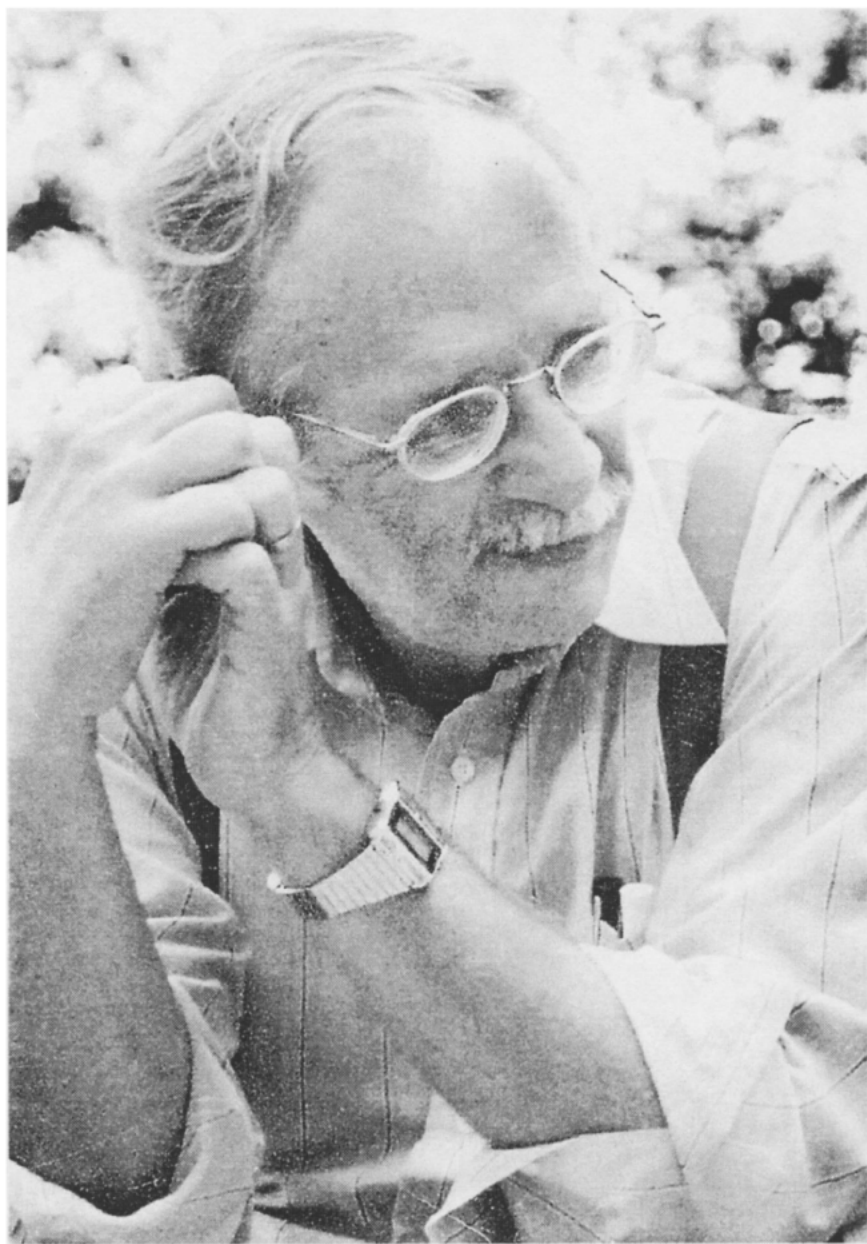
Stanley Starosta (shown here in 1998) has had an enduring influence on his students and colleagues throughout his academic career.

SELECTED PUBLICATIONS OF STANLEY STAROSTA

1969

Review of Tung T'ung-ho, *A descriptive study of the Tsou language, Formosa. Language* 45.2:439–444.

Recent schisms in transformational theory. *Philippine Journal of Language Teaching* 7.3–4:64–75.

1971

Some lexical redundancy rules for English nouns. *Glossa* 5.2: 167–201.

Review of John Lyons, *Introduction to theoretical linguistics. Language* 47.2:429–447.

Derivation and case in Sora verbs. *Indian Linguistics* 32.3: 194–206.

Linguistics and anti-science. *Philippine Journal of Linguistics* 2.1:13–27.

1973

The faces of case. *Language Sciences* 25.1–14. French translation, "Les visages du cas," appeared in *Langage* 38:104–128, 1975.

1974

Causative verbs in Formosan languages. *Oceanic Linguistics* 13.1–2:279–369.

1975

Case in the lexicon. In *Proceedings of the Eleventh International Congress of Linguists,* vol. 2, ed. by Luigi Heilmann, 805–813. Bologna: Societa Editrice il Mulino.

1976

A place for case. *Language Learning* 26:1–36. Appeared in *University of Edinburgh Work in Progress* 7:18–34, 1974.

With Philip N. Jenner and Laurence C. Thompson, eds. *Austroasiatic Studies.* Oceanic Linguistics Special Publication No. 13. Honolulu: University of Hawai'i Press.

Case forms and case relations in Sora. In *Austroasiatic studies,* pt. 2, ed. by Philip N. Jenner, Laurence C. Thompson, and Stanley Starosta, 1069–1107. Oceanic Linguistics Special Publication No. 13. Honolulu: University of Hawai'i Press.

1978

The one per Sent solution. In *Valence, semantic case, and grammatical relations,* ed. by Werner Abraham, 459–576. Studies in Language Companion Series No. 1. Amsterdam: John Benjamins. German translation, "Die 1-Pro-Sent Lösung," appeared in *Beiträge zum Stand der Kasustheorie,* ed. by Jochen Pleines, 45–148. Tübingen: Gunter Narr, 1981.

1980

Review of Neil Smith and Deirdre Wilson, *Modern linguistics: The results of Chomsky's revolution. Language Sciences* 2.2:339–342.

1981

Review of John M. Anderson, *On case grammar. Language* 57.3:137–140.

Case relations, perspective, and patient centrality. *Osmania Papers in Linguistics* 7–8:1–21.

1982

Lexical decomposition: Features or atomic predicates? *Linguistic Analysis* 9.4:379–393.

With Andrew K. Pawley and Lawrence A. Reid. The evolution of focus in Austronesian. In *Papers from the Third International Conference on Austronesian Linguistics*, vol. 2: *Tracking the travelers*, ed. by Stephen A. Wurm and Lois Carrington, 145–170. Pacific Linguistics C-75. Canberra: Australian National University.

Affix hobbling. LAUT (Linguistic Agency, University of Trier) Series A, Paper No. 95. Trier: University Trier. Appeared in University of Hawai'i *Working Papers in Linguistics* 9, 1:62–158, 1977.

Lexicase: An attempt at a generative reformation. *Kiswahili* 49.2:62–79.

1983

Patient centrality and English verbal derivation. In *Proceedings of the 13th International Congress of Linguists*, ed. by Shiro Hattori, Kazuko Inoue, Tadao Shimomiya, and Yoshio Nagashima, 489–492. Tokyo: International Congress of Linguists' Office.

1984

With Hirosato Nomura. *Lexicase and Japanese language processing*. Technical report no. 21887, Musashino Electrical Communications Laboratory. Tokyo: Nippon Telegraph and Telephone Public Corporation.

1985

The locus of case in South Asian languages. In *Proceedings of the Conference on participant roles: South Asia and adjacent areas*, ed. by Arlene R. K. Zide, David Magier, and Eric Schiller, 211–246. Bloomington: Indiana University Linguistics Club.

Relator nouns as a source of case inflections. In *For Gordon Faibanks*, ed. by Richard L. Leed and Veneeta Z. Acson, 111–133. Oceanic Linguistics Special Publication No. 20. Honolulu: University of Hawai'i Press.

Mandarin case marking: A localistic lexicase analysis. *Journal of Chinese Linguistics* 13.2:215–266.

Verbal inflection versus deverbal nominalization in PAN [Proto-Austronesian]: The evidence from Tsou. In *Austronesian linguistics at the 15th Pacific Science Congress*, ed. by Andrew Pawley and Lois Carrington, 281–312. Pacific Linguistics C-88. Canberra: Australian National University.

The end of phrase structure as we know it. LAUDT (Linguistic Agency, University of Duisburg, previously Trier) Series A, Paper No. 147). Duisburg: University of Duisburg. Earlier version appeared in University of Hawai'i *Working Papers in Linguistics* 11, 1:59–76, 1979.

With Louise Pagotto. Lexicase grammatical theory and its application to English and Japanese grammar. Interim report; first annual report. Tokyo: Nippon Telegraph and Telephone Public Corporation Electrical Communication Laboratory.

1986

Focus as recentralization. In FOCAL I: *Papers from the Fourth International Conference on Austronesian Linguistics*, ed.by Paul Geraghty, Lois Carrington, and S. A. Wurm, 73–95. Pacific Linguistics C-93. Canberra: Australian National University.

Lexicase parsing algorithm IV. Technical report no. 1, MENTAT Project, Pacific International Center for High Technology Research, Honolulu.

Rules and lexical entries for the major classes of English determiners and nouns. Technical report no. 2, MENTAT Project, Pacific International Center for High Technology Research, Honolulu.

With Hiroshi Nomura. Lexicase parsing: A lexicon-driven approach to syntactic analysis. In *Proceedings of the Eleventh International Conference on Computational Linguistics*, ed. by M. Nagao, 127–132. Bonn: Association for Computational Linguistics.

With Hisami Konishi Springer. Outline of lexicase grammatical theory (in Japanese). *Gengo Kenkyu* 90:1–26.

A place for (lexi)case. In *Concepts of case*, ed. by René Dirven and Günther Radden, Tübinger Beiträge zur Linguistik (SeG 4). Tübingen: Gunter Narr.

With Louise Pagotto. Lexicase grammatical theory and its application to English and Japanese grammar. Second annual report. Tokyo: Nippon Telegraph and Telephone Public Corporation Electrical Communication Laboratory.

With Hirosato Nomura, Louise Pagotto, and Hisami Springer. Lexicase grammatical theory and natural language processing (in Japanese). Technical Report No. 23427. Tokyo: Musashino Electrical Communication Laboratory, Nippon Teleraph and Telephone Corporation.

1987

Generative syntax: Hypothetico-deductive science of language. Taipei: The Crane Publishing Company.

The case for lexicase: An outline of lexicase grammatical theory. Open Linguistics Series, ed. by Robin Fawcett. London: Pinter Publishers.

A grammatical typology of Formosan languages. Fang-kuei Li memorial volume. *Bulletin of the Institute of History and Philology,* vol. 59, part 2, 541–576. Taipei: Academia Sinica.

1989

With Louise Pagotto. The grammatical genealogy of Chamorro. In VICAL 2: *Papers from the Fifth International Conference on Austronesian Linguistics,* ed. by Ray Harlow and Robin Hooper, 319–348. Te Reo Monographs. Auckland: Linguistics Society of New Zealand.

Sora combining forms and pseudo-compounding. In *Proceedings of the Symposium on Austro-Asiatic Languages,* ed. by David Thomas. *Mon-Khmer Studies* 18–19:77–105.

Generative methodology. LAUD (Linguistic Agency, University of Duisburg) Series A, Paper No. 239. Duisburg: University of Duisburg.

1990

Review of H. L. Somers, *Valency and case in computational linguistics. Machine Translation* 5.1:79–96.

Subgrouping by lexical similarity. *Bulletin of the Institute of History and Philology* 61.4:836–840. Taipei: Academia Sinica.

With Jeanne D. Gibson. Ergativity east and west. In *Linguistic change and reconstruction methodology,* ed. by Philip H. Baldi. *Trends in lingusitics: Studies and Monographs* 45:195–210. Berlin: Mouton de Gruyter.

1991

Natural language parsing and linguistic theories: Can the marriage be saved? In *Natural language parsing and linguistic theories,* ed. by U. Reyle and Christian Rohrer, 175–197. Studies in linguistics and philosophy, vol. 35. Studies in Language 15. Dordrecht: D. Reidel.

Grammar, perception, and reality. In *Currents in Pacific linguistics: Papers on Austronesian languages and ethnolinguistics in honor of George W. Grace,* ed. by Robert Blust, 477–500. Pacific Linguistics C-117. Canberra: Australian National University.

The great AUX cataclysm: Diachronic justification for a synchronic analysis. In *Studies in Dravidian and general lingusitics: A festschrift for Bh. Krishnamurti,* ed. by B. Lakshmi Bai and B. Ramakrishna Reddy, 505–515. Hyderabad, India: Osmania University.

1993

Word order and focus in constrained dependency grammar. In *Proceedings of the conference: Functional description of language,* ed. by Eva Hajičová, 237–252. Prague: Faculty of Mathematics and Physics, Charles University.

The case marking system of Proto-Formosan. In *Pan-Asiatic Linguistics, Proceedings of the Third International Symposium on Language and Linguistics,* vol. 3, ed. by Sudaporn Luksaneeyanawin, 1207–1221. Bangkok: Chulalongkorn University Printing House.

1994

Lexicase. In *The encyclopedia of language and linguistics,* vol. 4, ed. by R. E. Asher and J. M. Y. Simpson, 2167–2174. Oxford: Pergamon Press.

A grammatical subgrouping of Formosan languages. In *Austronesian Studies Relating to Taiwan,* ed. by Paul J.-K. Li, Cheng-hwa Tsang, Ying-kuei Huang, Dah-an Ho, and Chiu-yu Tseng, 683–726. Symposium series of the Institute of History and Philology no. 3. Taipei: Academia Sinica.

1995

The Chinese-Austronesian connection: A view from the Austronesian morphology side. In *The ancestry of the Chinese language,* ed. by William S-Y. Wang, 373–392. Monograph series no. 8. *Journal of Chinese Linguistics.*

1996

The position of Saaroa in the grammatical subgrouping of Formosan languages. In *Pan-Asiatic linguistics: Proceedings of the Fourth International Symposium on Languages and Linguistics,* ed. by Sulilai Premsirat, 944–966. Salaya, Thailand: Institute of Language and Culture for Rural Development, Mahidol University at Salaya.

1997

Formosan clause structure: Transitivity, ergativity, and case marking. In *Chinese languages and linguistics 4: Typological studies of languages in China,* ed. by Chiu-yu Tseng, 125–154. Symposium Series of the Institute of History and Philology no. 2. Taipei: Academia Sinica.

Sora noun inflection. In *Languages of tribal and indigenous peoples of India: The ethnic space,* ed. by Anvita Abbi, 263–306. MLBD Series in Linguistics 10. Delhi: Motilal Banarsidass Publishers.

Control in constrained dependency grammar. In *Reconnecting language: Morphology and syntax in functional perspectives,* ed. by Anne-Marie Simon-Vandenbergen, Kristin Davidse, and Dirk Noël, 99–138. Amsterdam: John Benjamins.

With Chhany Sak-Humphry and Kitima Indrambarya. Flying in and out in Khmer and Thai. In *Southeast Asian linguistic studies in honour of Vichin Panupong,* ed. by Arthur Abramson, 209–220. Bangkok: Chulalongkorn University Press.

1998

With Koenraad Kuiper, Siew Ai Ng, and Zhi-qian Wu. On defining the Chinese compound word: Headedness in Chinese compounding and Chinese VR compounds. In *New Approaches to Chinese word formation: Morphology, phonology, and the lexicon in Modern and Ancient Chinese*, ed. by Jerome L. Packard, 347–370. Trends in Linguistics, Studies and Monographs. Berlin: Mouton de Gruyter.

Lexicase revisited. In *Studies in linguistics in honour of Bela Korponay: A special issue*, ed. by Jozsef Andor, Bela Hollosy, Tibor Laczko, and Peter Pelyvas, 279–334. Debrecen: Kiadta a Kossuth Lajos Tudomanyegyetem, Angol-Amerikai Intezete.

With Hisami Springer. A dependency analysis of coordination in English and Japanese. In *Proceedings of the 16th International Congress of Linguists*, ed. by B. Caron. Paris: Pergamon/Elsevier.

Ergativity, transitivity, and clitic coreference in four Western Austronesian languages. In *Case, typology and grammar: In honour of Barry J. Blake*, ed. by Anna Siewierska and Jae Jung Song, 277–307. Typological Studies in Language. Amsterdam: John Benjamins.

To appear:

Control in Gujarati V-i V constructions: A constrained dependency grammar analysis. In [Professor H. S. Biligiri commemorative volume], ed. by D. N. S. Bhat, K. S. Gurubasave Gowda, K. S. Nagaraja, and K. Narayana.

Dependency grammar and monostratal transfer. In [C. Ramarao festschrift], ed. by B. Lakshmi Bai. Hyderabad: Osmania University.

Gujarati morphological causatives: A word-and-paradigm analysis. In *Proceedings of the Eighteenth South Asian Languages Analysis Roundtable* (SALA 18). Delhi: Jawaharlal Nehru University.

With Cathy Sin Ping Wong. How can syntactic change diffuse through the lexicon?: The case of Cantonese comparative constructions. In *Proceedings of the Fifth Annual International Conference on Chinese Linguistics* (ICCL-5), ed. by H. Samuel Wang and Chin-fa Lien. Hsinchu, Taiwan: Tsing Hua University.

The identification of word classes in Thai. In *Proceedings of the Thai languages in the twentieth century symposium*, ed. by M. R. Kalaya Tingsabadh and Arthur Abramson. Bangkok: Chulalongkorn University Press.

Nonreferential arguments and control in constrained dependency grammar. In [Festschrift for E. Annamalai], ed. by Kedutso Kapfo. Manasagangotri, Mysore, India: Central Institute of Indian Languages.

Lexicase revisited. *Encyclopedia of languages and linguistics*. Pergamon Press.

Part 1
Theory, History, Pragmatics

1:   THE ARCHITECTURE OF SYNTACTIC REPRESENTATIONS: BINARITY AND DECONSTRUCTION

William O'Grady
University of Hawai'i

1. Introduction[1]

Unlike many of the contributors to this volume, I approach syntax from the perspective of categorial grammar rather than dependency grammar. In various ways, however, the two theories are alike. Indeed, it has sometimes been proposed that categorial grammar and dependency grammar are "sister" theories of language, and even that categorial grammar is a type of dependency grammar (Wood 1993:1, 3).

Like dependency grammar, categorial grammar is monostratal in that there is only one level of representation (and no movement or deletion rules), it is word-based in that a sentence's form is largely determined by properties of its component words, and it is conservative with regards to what counts as a syntactic category in that most work does not treat elements such as "tense" or "agreement" as heads. Of course, there are also differences between the two theories, including the fact that sentence structure has a binary architecture in categorial grammar. (However, nothing in principle rules out the possibility of more elaborate types of n-ary branching. See Bach 1988:6 and Wood 1993:23.)

The purpose of this essay is to explore a theory of sentence structure that draws on certain assumptions associated with both categorial grammar and dependency grammar but also incorporates ideas that are accepted by neither. In particular, I will follow the consensus in both frameworks in assuming that syntactic representations are a "projection" of the lexical properties of individual heads and that movement rules play no role in their formation. However, I will depart from traditional work in this and most other frameworks in arguing that the grammar operates from left-to-right. (Among other things, this entails a departure from the commonly held view in categorial grammar that structure is built from the bottom up, more or less from right to left.) I will focus in this study on the application of these ideas only to SVO languages.[2] (See O'Grady 1997 for a detailed discussion of SOV and VSO languages.)

---

1. In developing the ideas presented here, I benefited from comments by Byron Bender, Videa De Guzman, Kazue Kanno, Gregory Lee, and an anonymous referee. Special thanks are due to Stanley Starosta, not only for many discussions concerning this essay but also for his friendship and collegiality over the past fifteen years.

2. In fact, I focus here on "pure" SVO languages, in which nothing intervenes between the subject and the verb. French is in fact a better example of such a language than is English, because it does not permit even an adverb to occur between the subject and the verb (cf. Fr. *Jean souvent travaille* vs. Eng. *John often works*). Nonetheless, for the sake of simplicity and convenience, I will draw on English data for the purposes of this study.

2. The Binarity Issue

I take as my starting point an idea that has long been a cornerstone of Stanley Starosta's work (for example, Starosta 1988) and that is now more or less universally accepted in the field: words carry information that ultimately determines a sentence's form. For now, we can focus on two such pieces of information—the word's syntactic category and the number and type of elements with which it must combine. The word *leave*, for example, carries the information that it is a verb and that it takes an agentive nominal argument. This information can be represented as follows, with "V" standing for verb, "N" for nominal, and "ag" for "agent."[3]

    (1)   *leave*: V, <N>
                      ag

A simple transitive verb such as *meet* will have the lexical properties depicted in (2) ("th" stands for "theme").

    (2)   *meet*: V, <N, N>
                  ag th

The most direct way to satisfy a word's argument requirements is through combination—as when an intransitive verb combines with its subject. In the structure depicted below, for example, the verb *left* combines with the nominal *Harvey*, satisfying or "canceling" its argument requirement and giving a sentence in which the property of having left is attributed to Harvey. (In the interests of expository simplicity, I do not include information about thematic roles in syntactic representations.)

    (3)

$$
\begin{array}{ccc}
 & V & \\
\diagup & & \diagdown \\
N_i & & V \\
| & & <N_i> \\
Harvey & & left
\end{array}
$$

Following a convention also employed in recent work by Starosta and various of his students (for example, Lee 1989, Springer 1993), I indicate that an argument requirement has been satisfied by copying the referential index of the relevant nominal onto the corresponding symbol in the verb's argument grid.

    Two other points need to be made about this representation. First, I assume that representations indicate the category membership of words and phrases, but not their "bar level"—a point long maintained by Starosta as well as by many categorial grammarians and, more recently, by Chomsky as well (see, for example, 1995:246). Thus, N stands for both "noun" and "noun phrase"; V for both "verb" and "verb phrase" and so on. Second, sentences containing a verb are verbal projections—another idea long advocated by Starosta as well as by many others working outside transforma-

---

3. As they are used here, thematic role labels are not intended to be primitives of linguistic analysis. Rather, I follow Jackendoff (1990:55) in assuming that thematic roles are really just names for particular argument positions in conceptual structure.

tional grammar—for example, in categorial grammar (Bar-Hillel 1953), generalized phrase structure grammar (Gazdar et al. 1985:61), head-driven phrase structure grammar (Pollard 1988:398), and construction grammar (Fillmore 1988:43).

In other respects, however, the view of structure building that I wish to put forward departs from the assumptions on which dependency grammars—including Lexicase—are built. One of these differences shows up when we consider the structure of transitive clauses. Here, consistent with the assumption that structure-building operations apply to pairs of elements and that a verb's agent argument is more prominent than its theme in English active sentences, I adopt the representation in (4) for the sentence *Harvey met Mary.*

(4)

$$
\begin{array}{c}
V \\
\diagup \diagdown \\
N_i \quad V \quad N_j \\
| \quad | \\
\text{Harvey} \quad <N_i N_j> \\
| \\
\text{met} \quad \text{Mary}
\end{array}
$$
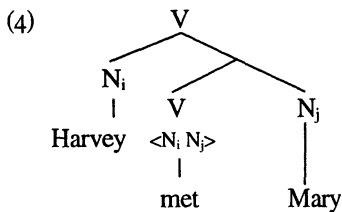
In contrast, the same sentence would have a representation resembling (5) in Lexicase dependency grammar (omitting information not directly relevant to the point at hand).
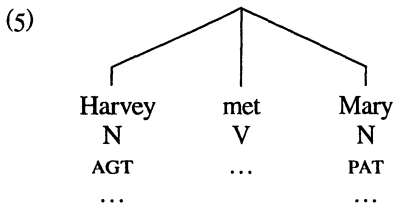
(5)

Harvey	met	Mary
N	V	N
AGT	...	PAT
...		...

An apparent advantage of the structure I propose is that it has just the architecture needed to accommodate traditional accounts of so-called "c-command asymmetries"— that is, the fact that a subject can serve as antecedent for a direct object reflexive and can license a negative polarity item in the direct object position, but not vice versa.

(6) a. The subject serves as antecedent for a direct object reflexive:
Harvey$_i$ admires himself$_i$.

    b. The direct object serves as antecedent for a subject reflexive:
*Himself$_i$ admires Harvey$_i$.

(7) a. The subject licenses a negative polarity item in the direct object position:
No one criticized anyone.

    b. The direct object licenses a negative polarity item in the subject position:
*Anyone criticized no one.

Given representations with the architecture in (4), contrasts such as these can be explained in terms of simple principles such as the following.

(8)   Prominence Constraint:
        A reflexive pronoun requires a structurally higher antecedent.

(9)   Constraint on Negative Polarity Items:
        A negative polarity item can be licensed by a structurally higher negative.[4]

Starosta (1997 class lectures) proposes that the reflexivization facts follow from a principle that requires a reflexive pronoun to be lower than its antecedent in a case hierarchy of the form AGT > PAT > LOC, COR, MNS. This principle certainly works in the case of the contrast illustrated in (6), where a PAT reflexive can take an AGT antecedent but not vice versa, and its effects compare favorably with those achieved both by Pollard and Sag's (1992) relational hierarchy and by the sort of thematic hierarchy proposed by Jackendoff (1972), Wilkins (1988), and Grimshaw (1990), among others. Moreover, both the Prominence Constraint and Starosta's case hierarchy correctly predict that linear order is not responsible for the contrasts exemplified in (6). Thus they both account for the fact that in VOS languages such as Malagasy, the second nominal can serve as antecedent for the first, but not vice versa. (The following data are from Keenan 1976:263.)

(10) a. DIRECT OBJECT REFLEXIVE—SUBJECT ANTECEDENT
        namono tena Rabe
        killed   body Rabe
        'Rabe killed himself.'

     b. DIRECT OBJECT ANTECEDENT—SUBJECT REFLEXIVE
        *namono   an-dRabe   (ny)   tena(-ny)
         killed   Ac-Rabe   the   body his
        'Himself killed Rabe.'

The second nominal in these sentences is structurally more prominent than the first in representations that employ binary branching, such as the one shown in (11). And it is higher in Starosta's hierarchy as well, since it is an AGT. Thus, both approaches predict that the first sentence should be acceptable and the second sentence unacceptable.

(11)

        V   N   N
       <N N> | |
         | | |
        namono   tena   Rabe
        killed   body   Rabe
   'Rabe killed himself'

---
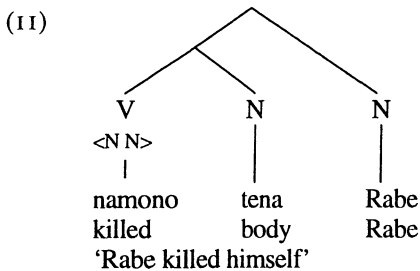
4. *AnyX* can be licensed by elements other than a negative (for example, a modal such as *can*—as in *Anyone can do that* vs. **Anyone did that*). The key point for the purposes of our discussion is simply that when *anyX* is licensed by a negative word, that negative must be structurally higher.

Crucially, however, the case hierarchy analysis fares less well with the negative polarity facts. It might appear that a negative polarity item can be licensed only by an element that is higher in the hierarchy (hence the AGT negative in (7a) licenses the PAT *anyone*, but the PAT in (7b) cannot license the AGT). However, sentences such as the following undermine this approach.

(12)   John never helps anyone.

Here the negative adverbial *never* (a noncomplement in Starosta's system) licenses the PAT *anyone*, which is surely higher than it in the hierarchy. No such problem arises in the configurational approach adopted here, since (12) has the structure depicted below, in which *never* is structurally higher than the negative polarity item in accordance with the principle proposed in (9).

(13)
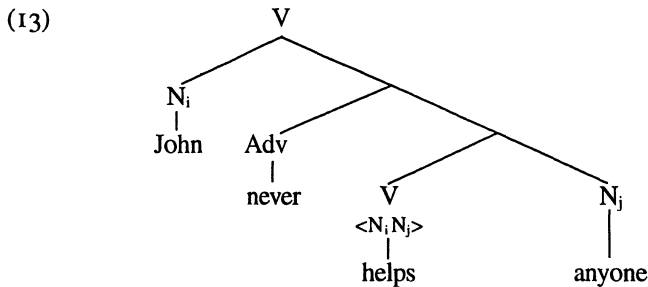
These considerations notwithstanding, dependency grammar representations or "stemma" have a number of advantages. In addition to being "minimalist" in an interesting way, they allow a simple and unified characterization of the set of potential agreement triggers (e.g., subjects and direct objects)—which end up as "sisters" of the head on which the agreement affixes appear (for example, Starosta 1988:86). At first glance, this generalization seems impossible in the sort of syntactic representation I have proposed, since only the direct object is a sister of the verb. However, as I will show in the next section, this is not in fact the case: contrary to appearances, the verb's arguments are all sisters of the verb at different points in the sentence formation process.

3. Deconstruction

Any adequate account of how human language works must ultimately confront what might be called the "linearity problem"—the fact that sentences are produced and perceived from left-to-right, essentially one word at a time. The traditional literature on grammar sidesteps this problem by distinguishing between a theory of sentence structure, which is concerned with the architecture of syntactic representations, and a theory of processing, which deals with the manner in which structures with an appropriate architecture are built in the course of producing or perceiving a sentence. Almost without exception, the question of how syntactic structure is built is relegated to the latter enterprise. In contrast to this consensus, I take the position that the mech-

anisms responsible for determining the internal architecture of a sentence are designed to operate in a linear manner.

If this idea is on the right track, then we will stand at the threshold of being able to unify the study of sentence structure with the study of processing. In order to demonstrate the viability of this proposal, it will ultimately be necessary to consider not only evidence involving the admissibility of particular forms and interpretations relevant to the theory of sentence structure, but also psycholinguistic data involving garden path effects, reaction time, short-term memory phenomena, and other factors that shed light on the production and perception of sentences. If I am right, the two types of evidence should converge to support the particular type of structure-building system that I propose.

Space does not permit me to even begin to undertake such an enterprise here (but see O'Grady 1997). Instead, I will focus on one small part of this sort of approach—a structure-building mechanism that offers a simple solution to the linearity problem while at the same time offering insights into the nature of phenomena such as agreement, binding, and scope, which have long been the central concern of research into sentence structure.

In the theory I propose (and in contrast with traditional categorial grammar), sentences are "generated" from left to right, with argument requirements being satisfied as quickly as possible in accordance with the following principle.

(14)   The Immediacy Requirement:
          Argument requirements must be satisfied immediately, if possible.[5]

I assume that it is possible to satisfy an argument requirement immediately if the relevant argument has already been encountered in the course of the left-to-right structure-building process. Thus, in an SVO language, a verb must combine immediately with its subject since this element precedes it in the linear string and therefore has already been encountered. This in turn entails that there is a point in the formation of the sentence *Harvey met Mary* where the grammar generates the structure depicted in (15).

(15)   STEP 1: Combination of the verb with its subject argument
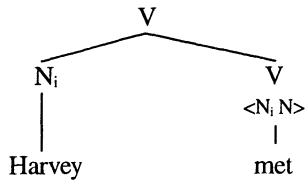
In the next step, the direct object argument is added to the syntactic representation. As depicted below, an operation that I will call DECONSTRUCTION allows it to combine

---

5. Where English departs from the SVO template, as in *wh* questions, Adv S V patterns, topicalizations, and so forth (see also note 2), argument requirements often cannot be resolved immediately—necessitating the use of a push-down storage mechanism not discussed here.

directly with the verb even though this element is already part of a larger constituent that includes the subject.

(16) STEP 2: Combination of the direct object argument and the verb

```
                    V
          ┌─────────┴────
        Nᵢ            ╲
         │         ┌───╲────
      Harvey       V      Nⱼ
                <Nᵢ Nⱼ>     │
                   │        │
                  met      Mary
```
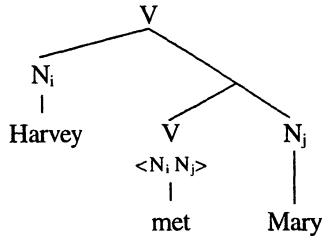
In other words, in the course of adding each new dependent as a sister of the head, the sentence-building system pushes each previously attached dependent one level higher. Thus, the initial combinatorial operation in the example above makes the subject the verb's sister, but subsequent attachment of the direct object NP "pushes" the subject to a higher position and replaces it as sister of the verb.[6] The end result is a representation in which the subject is higher than the direct object, as required to account for asymmetries involving reflexive pronoun interpretation and negative polarity. At the same time, however, there is a point in the sentence formation process at which the verb's subject argument is its sister (see [15] above) and a later point at which the direct object is its sister (see [16]). Assuming that agreement is triggered (or at least checked) at those respective points in the sentence formation process, we can maintain the view that agreement triggers are sisters of the verb without sacrificing the equally important subject-object asymmetry.

4. Double Object Patterns

An operation similar in its effects to deconstruction has long been assumed in the literature on sentence processing and parsing (see, for example, Marcus 1980:79–80 and Levelt 1989:242), and is a necessary part of any theory of syntax that seeks to address the linearity problem.[7] The key innovation proposed here is to incorporate deconstruction into the theory of sentence structure rather than using it only as part of a processing or parsing module. The justification for this is simple: without deconstruction, it would not be possible to build the representations needed to account for phenomena that are a central part of traditional syntactic theory. This can be seen by considering so-called "double object" patterns.

In a traditional dependency analysis, a sentence such as *John gave Mary advice* would be assigned the following representation.

---

6. Starosta (1988:104) remarks that syntactic representations in dependency grammar are "viewed as a network of dependencies obtaining between . . . *pairs* [emphasis added] of lexical items" (e.g., a verb and its subject, a verb and its direct object, and so forth). The representations I propose directly capture the pair-wise nature of these relationships.

(17)

```
        John    gave    Mary   advice
         N       V       N       N
        ...     ...     ...     ...
```
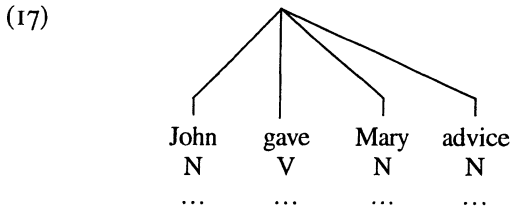
In a traditional phrase structure analysis, in contrast, the sentence would have the structure depicted below.

(18)

```
                    S
              ┌─────┴──┐
                      VP
            ┌───┬──────┴────┐
           NP   V     NP    NP
            |   |      |     |
          John gave  Mary  advice
```
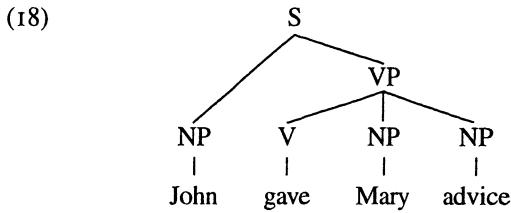
The system of structure building I propose postulates yet another structure.

(19)   STEP 1:   Combination of the verb and its subject argument.

```
                V
           ┌────┴────┐
          N_i         V
           |        <N_i N N>
           |           |
         John        gave
```

---
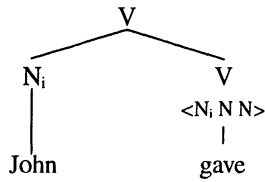
7. This may be true of Lexicase as well, as can be seen by considering the formation of biclausal sentences such as *John thinks [Harvey met Mary]*. Assuming the sort of left-to-right structure-building needed to resolve the linearity problem, the syntactic representation has the form depicted in (i) at the point just before the direct object in the embedded clause is encountered.

(i)

```
        John    thinks
         N        V
                        ┌────┴──┐
                      Harvey   met
                        N       V
```
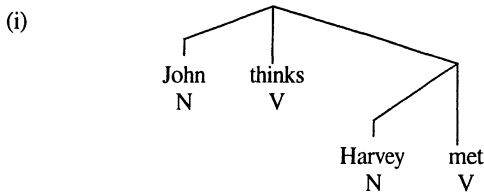
Upon encountering the nominal *Mary*, the structure-building mechanisms must add this element to the clause headed by *meet*, which is itself embedded inside a larger clause. The operation required to do this is in essence deconstruction since it allows the final constituent to be attached as sister of an element that is embedded within a larger phrase. However, as Stanley Starosta has pointed out to me, no sisterhood relations are altered in the Lexicase stemma—in contrast with what happens in the syntactic representations that I propose.

10

STEP 2: Use of deconstruction to allow the verb to combine directly with its first object.

$$
\begin{array}{c}
V \\
N_i \qquad \\
| \qquad \\
\text{John} \qquad V \qquad\qquad N_j \\
\langle N_i\, N_j\, N\rangle \qquad | \\
| \qquad\qquad \text{Mary} \\
\text{gave} \\
\end{array}
$$
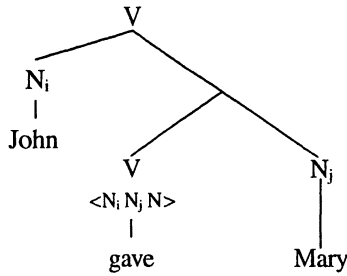
STEP 3: Deconstruction applies once again, allowing the verb to combine directly with its second object.

$$
\begin{array}{c}
V \\
N_i \\
| \\
\text{John} \qquad\qquad N_j \\
| \\
\text{Mary} \\
V \qquad\qquad\qquad N_k \\
\langle N_i\, N_j\, N_k\rangle \qquad | \\
| \qquad\qquad \text{advice} \\
\text{gave} \\
\end{array}
$$
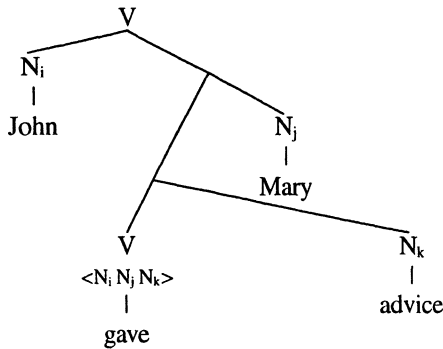
As shown here, each of the verb's arguments is temporarily its sister, with each subsequent combinatorial operation pushing it higher in the tree structure in order to allow the next argument to combine directly with the verb (thanks to deconstruction). The end result is that an argument's relative prominence is determined by its linear position: each argument is structurally lower than any preceding arguments in its clause—exactly the conclusion argued for on independent grounds and implemented in a very different way by Kayne (1994).

Which of these three representations is correct? One way to approach this question is to once again employ the phenomena of anaphora and negative polarity as tests for structural asymmetries. Recall that these phenomena are subject to the principles in (8) and (9), repeated from above.

(8)   Prominence Constraint:
       A reflexive pronoun requires a structurally higher antecedent.

(9)   Constraint on Negative Polarity Items:
       A negative polarity item can be licensed by a structurally higher negative.

We can thus use asymmetries such as the following, first noted by Barss and Lasnik (1986), to determine the relative prominence of the verb's complements in double object patterns.

(20)   Interpretation of a reflexive pronoun

    a. The first object serves as antecedent for the second object:
I showed Mary herself (in the photo).

    b. The second object serves as antecedent for the first object:
*I showed herself Mary (in the photo).

(21)   Licensing of a negative polarity item

    a. The first object licenses a negative polarity item in the second object:
I gave no one anything.

    b. The second object licenses a negative polarity item in the first object:
*I gave anyone nothing.

The contrasts exemplified here point to the conclusion that the first object is asymmetrically higher than the second object, just as the subject is asymmetrically higher than the direct object. Crucially, such asymmetries are found only in representations such as (19), whose formation reflects a commitment to binary branching and to the uniform use of deconstruction. (Starosta's case hierarchy also works here, but remember that it cannot accommodate examples like [12] above.)

Phenomena such as binding and scope are at the heart of much contemporary work in syntactic theory. The fact that the representations needed to permit a simple characterization of these and other phenomena presuppose the use of deconstruction provides strong evidence that this operation should be part of the theory of sentence structure and not just part of a theory of parsing or processing. Put more generally, there is reason to believe that the linearity problem should be treated as a central concern of syntactic theory and that its resolution has direct relevance for the analysis of a number of classic syntactic problems.

5. Concluding Remarks

A major challenge for all work on syntax is to provide the syntactic representations needed to accommodate principles that account for linguistic phenomena such as agreement, anaphora, and negative polarity. The proposal that I have put forward seeks to maintain something like the sisterhood constraint on agreement while retaining the binary architecture needed to represent structural asymmetries. It does this by building sentences from left-to-right with the help of a deconstruction operation that allows a head that is already part of a larger phrasal unit to combine directly with subsequent elements. In this system, then, each of the verb's arguments is its sister at some point in the sentence formation process, but the end result is a binary branching representation with the asymmetries appropriate for phenomena such as binding and negative polarity.

Of course, as with any proposal of this sort, there are numerous problems to work out. Nonetheless, there is perhaps enough promise in the ideas outlined here to justify further exploration of the territory that lies beyond the possibilities envisioned by

classical monostratal theories of syntactic representation such as categorial grammar and dependency grammar.

REFERENCES

Bach, Emmon. 1988. Categorial grammars as theories of language. In *Categorial grammars and natural language structures*, ed. by R. Oehrle, E. Bach, and D. Wheeler, 17–34. Dordrecht: Reidel.

→ Bar-Hillel, Yehoshua. 1953. A quasi-arithmetical notation for syntactic description. *Language* 29:47–58.

Barss, Andrew, and Howard Lasnik. 1986. A note on anaphora and double objects. *Linguistic Inquiry* 17:347–353.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, Mass.: MIT Press.

Fillmore, Charles. 1988. The mechanisms of "Construction Grammar." *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, 35–55.

Gazdar, Gerald, Ewan Klein, Geoffrey Pullum, and Ivan Sag. 1985. *Generalized phrase structure grammar*. Cambridge, Mass.: MIT Press.

Grimshaw, Jane. 1990. *Argument structure*. Cambrdige, Mass.: MIT Press.

Jackendoff, Ray. 1972. *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press.

————. 1990. *Semantic structures*. Cambridge, Mass.: MIT Press.

Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, Mass.: MIT Press.

Keenan, Edward. 1976. Remarkable subjects in Malagasy. In *Subject and topic*, ed. by C. Li, 247–301. New York: Academic Press.

Lee, Nagiko I. 1989. Complementation in Japanese: A lexicase analysis. Ph.D. dissertation, University of Hawai'i.

Levelt, Willem. 1989. *Speaking: From intention to articulation*. Cambridge, Mass.: MIT Press.

Marcus, Mitchell. 1980. *A theory of syntactic recognition for natural language*. Cambridge, Mass.: MIT Press.

O'Grady, William. 1997. Modes of sentence formation. Unpublished ms. Department of Linguistics, University of Hawai'i.

Pollard, Carl. 1988. Categorial grammar and phrase structure grammar: An excursion on the syntax-semantics frontier. In *Categorial grammars and natural language structures*, ed. by R. Oehrle, E. Bach, and D. Wheeler, 391–415. Boston: Reidel.

Pollard, Geoffrey, and Ivan Sag. 1992. Anaphors in Engish and the scope of binding theory. *Linguistic Inquiry* 23:261–303.

Springer, Hisami. 1993. Perspective-shifting constructions in Japanese: A lexicase dependency analysis. Ph.D. dissertation, University of Hawai'i.

Starosta, Stanley. 1988. *The case for lexicase: An outline of lexicase grammatical theory*. London: Pinter Publishers.

Wilkins, Wendy. 1988. Thematic structure and reflexivization. In *Syntax and semantics*, vol. 21: *Thematic relations*, ed. by W. Wilkins, 191–213. New York: Academic Press.

Wood, Mary. 1993. *Categorial grammars*. New York: Routledge.

2: PARADIGMS AS RULES

Byron W. Bender
University of Hawai'i

1. Introduction[1]

Word and Paradigm (WP) as a model of morphological description has been around at least since the fifties, when Hockett (1954) mentioned it in passing, and Robins (1959) presented its major outlines. The two terms, "word" and "paradigm," properly understood, convey its two most distinctive characteristics. The lexicon is primarily a collection of words—the role of morphemes is secondary. Inflection is accomplished within paradigms in those languages that have inflection. But what, in more detail, is the role of paradigms within inflection? This is the question I attempt to address in these pages.

1.1 Dealing with Whole Words

Much of linguistic analysis has to do with taking words (and larger structures) apart and developing instructions about how to put them back together so that they can be restored exactly, or converted into related structures. This seems to be based on the assumption that speakers build words up from smaller pieces. The WP model makes the opposite assumption—that speakers deal with whole words exclusively, and that they store them ready-made and refashion them as necessary, based on other words as models, but do not take them apart, or build them up from scratch.

Matthews (1991) says, "We have seen how word-forms can be built up from their roots. . . . But there is an alternative method, whose sources lie in the work of the ancient grammarians of Greek and Latin. This is simply to relate words as wholes. . . . Which method is best? . . . The modern method has already been explored, and its attractions do not need to be laboured further. But there are at least three reasons why the opposite approach should not be neglected. Firstly, it conforms very closely to the method by which languages of this kind are traditionally taught. Pupils begin by memorising *paradigms*. These are sets of WORDS AS WHOLES, arranged according to grammatical categories. . . . Secondly, it is not clear that, when native speakers learn a flectional

---

language, they do not themselves learn words as wholes. . . . [F]inally, both Latin and Ancient Greek had native grammarians; and it is significant that, as native speakers writing for and teaching other native speakers, they too dealt with words as wholes."

The idea of dealing with words as wholes seems to fly so much in the face of what we do in linguistics that a word or two further on the subject may be in order before we proceed to look at the paradigm side of the WP model. If we do not build words up from their roots in daily use, why is it that they come apart so easily? Why is it that we are able to analyze them to the extent we are? First of all, they do not come apart all that easily in many languages, and they do not come apart that regularly in most. Second, there is at least one seam implicit in many words that reflects what we do in modeling them on other words. So it is not as though they cannot be taken apart at all. It is just that we do the modeling—the analogizing—so smoothly that we are not conscious of having taken them apart. But the main answer to the question as to why we as linguists are able to take words apart to the extent we are is that that is how they were built up originally, when they were first formed, and what we are doing in our analysis of them is actually internal reconstruction, a type of linguistic archaeology. We are uncovering that part of their history that has not been scrambled or otherwise eroded. The pieces of this history that can be recovered—the morphemes—make it easier for us to store and remember the words as wholes, and sometimes give us clues as to their meaning and grammatical classification. For discussion and some examples of this role in lexical morphology, see Bender 1998. Here, in section 2, we turn to the matter of inflection, hoping to get a better idea of how it may be possible to relate whole inflectional wordforms to each other and model them on one another, especially in a language that is both highly inflected and fusional—in this instance, Latin.

2. Paradigms Tell the Story of Inflection

Inflections are the ways in which a word changes its shape as its morphosyntactic features are changed. In fusional or flectional languages, the categories of these features often intersect, and when they do, the various inflections can be displayed as word-shapes in rows and columns headed and labeled by the intersecting features. Such displays require as many dimensions as there are categories intersecting. When the categories are simply person and number, for example, two suffice. When there are more, as there are for the Latin verb, which is inflected for a total of six categories (tense, aspect, voice, mood, person, and number), six dimensions are required. In a two or three-dimensional world, other devices must be found for their display.

2.1 The Case for Paradigms

Both structural and generative linguistics, intent on analysis, have been slow to recognize the paradigmatic nature of inflection. Thus Spencer, as recently as 1991, asks, in part rhetorically, "Where does this leave the paradigm? If the lexicalist stance is taken, then the paradigm will remain an epiphenomenon of the morphosyntactic feature system, and therefore of no intrinsic interest" (Spencer 1991:224). Spencer goes

on to review the work of Natural Morphologists such as Wurzel, and of Carstairs, in support of the paradigm as much more than an epiphenomenon. Strongest support comes from Carstairs's Paradigm Economy Principle, which summarizes the constraints placed on inflectional systems by paradigms. According to this principle, "the number of (macro)paradigms found in the language won't exceed the number of different affixes for the feature combination with the greatest variety of affixes. In our Hungarian example there are two combinations with maximal variation, the 1SG. and 3SG. forms, and these have two distinct affixes each. Hence, there can be only two different paradigms [rather than the four possible if unconstrained]" (Spencer 1991:229). (In the example referred to, of the four possible combinations of person-number endings [A, B, C, D] only two [A, D] are found.)

	A	B	C	D
1SG	-ok	-ok	-om	-om
3SG	-∅	-ik	-∅	-ik

Matthews (1991:197) identifies the same effect of paradigms by saying simply that "one inflection tends to predict another" and goes on to exemplify it with data from Latin noun inflection. "Traditionally, it is the basis for the method of exemplary paradigms. If the alternations were INDEPENDENT, these would have to be numerous. One class of Nouns would have a Genitive Singular like *dominus,* but all its other endings like *flos;* another would have the endings of *flos* in every form except the Dative/Ablative Plural, and so on for every possible combination. But since they are INTERDEPENDENT, the number can be very small. In tradition there are five" [emphasis mine—BWB].

2.2 What an Inflectional Rule Looks Like

Starosta (1991) proposes, in the spirit of WP, that both derivation and inflection are accomplished on a whole-word basis by analogy, thereby obviating the agglutinative morphemes, boundaries of various types, and other mechanisms generally posited by morphologists for building up morphologically complex words in a compositional approach. The forms of the rules he uses with this approach bear some similarity to those used by Matthews (1991:193) to show how members of a paradigm may be related analogically. The example in (1), from Matthews, is a "morphological transformation" he gives for forming the genitive singular from the nominative singular of Latin masculine nouns in *-os,* based on Priscian.

$$(1) \quad \begin{bmatrix} \text{Masculine in } \textit{-os} \\ \text{Nominative} \\ \text{Singular} \\ X + s \end{bmatrix}_N \quad \rightarrow \quad \begin{bmatrix} \text{Genitive} \\ \\ X + \text{ris} \end{bmatrix}$$

However, whereas this rule is unidirectional, Starosta's (1995) are bidirectional, as are those also of Ford, Singh, and Martohardjono (1997). An example from the latter for French, termed a "morphological strategy," is given in (2):[2]

$$(2) \quad /Xal/_{N\ SG.} \quad \leftrightarrow \quad /Xo/_{N\ PL.}$$

cheval, chevaux; mal, maux (Ford, Singh, and Martohardjono 1997:11)

16

Could Matthews's rule be made bidirectional by giving the arrow a double head and more fully specifying the right side, so that it would serve for forming the nominative if one knew the genitive and were starting there, as in (3), or as in (4) (which is [3] restated in the format of Ford, Singh, and Martohardjono)?

(3)
$$\left[\begin{array}{l} \text{Masculine in } \textit{-os} \\ \text{Nominative} \\ \text{Singular} \\ \text{X + s} \end{array} \right]_{N} \quad \leftrightarrow \quad \left[\begin{array}{l} \text{Genitive} \\ \text{Singular} \\ \text{X + ris} \end{array} \right]$$

(4) $/Xs/_{N \text{ NOM. SG.}} \quad \leftrightarrow \quad /Xris/_{N \text{ GEN. SG.}}$
flōs, flōris; mōs, mōris

As it turns out in this particular instance, the rule is not reversible unless it is somehow delimited to include nouns like *flōs, flōris* 'flower', but exclude those like *victor, victōris* 'victor', also a masculine noun of the third declension.[3] This is easily done if the rule is part of a set that operates for a class of verbs of specified membership, those that follow a given paradigm. Let us consider whether, and if so how, the bidirectional rules of Starosta or the bidirectional strategies of Ford, Singh, and Martohardjono might operate to specify all the members of an inflectional paradigm.[4]

2.3 Strategies within a Paradigm

In Latin verb inflection, the usual starting point is the Present Active Indicative. These inflections of a First Conjugation verb such as *amō* 'love' are given in a paradigmatic array according to person and number in (5).[5]

---

2. Say Ford, Singh, and Martohardjono (1997) of their strategies, "A morphological strategy . . . captures the morphological relatedness amongst the words in the 'lexicon' and allows a speaker to morphologically analyze a word she may not have analyzed before or to create a new word that she may not have yet encountered or may have forgotten temporarily. The listing of the morphological strategies of a language constitutes a part of the description of that language. It is, therefore, an aspect of linguistic competence, a component of grammar" (Ford, Singh, and Martohardjono 1997:2).

3. Matthews's stipulation on the nominative "Masculine in *-os*" serves this purpose in the original direction, but a similar stipulation of gender and form on the genitive—"Masculine in *-oris,*" perhaps—cannot serve in the opposite direction because of the coincidence of the genitive forms.

4. Neither Starosta (1991) nor Ford, Singh, and Martohardjono (1997) addresses this matter, because the paradigm has no status as an entity within either of their theories. Ford, Singh, and Martohardjono state this most explicitly: "No sub-category of strategies is exclusively identified morphologically. Hence, no 'conjugation/declension', or 'intra/extra-paradigmatic' kind of typology. . . . We claim that there is no need for the outlawed restrictions. The burden of proof is, clearly, on those who want to dispense with these constraints and introduce additional devices to account for the facts" (Ford, Singh, and Martohardjono 1997: 3–4).

5. My reasons for shifting from nominal to verbal morphology at this point in the discussion are primarily two: verbal morphology is richer, and the morphosyntactic features involved are more generally agreed upon than are those operating in case systems.

(5)		SG	PL
	I	amō	amāmus
	2	amās	amātis
	3	amat	amant

Note that the six forms fall into three groups according to their treatment of the historic theme vowel: it appears long *(ā)* in the IPL, 2SG, and 2PL forms, short *(a)* in 3SG and 3PL, and as zero in ISG. If we wish to formulate bidirectional "strategies" for each of these six forms in relation to each other, we might decide for each pair to let X equal the maximum form they share, in which case X will have different values, depending on the pairs involved. Even when one and the same form is being related to other forms, as many as three values for X might be needed. Thus, using the format of Ford, Singh, and Martohardjono,[6] in the RECIPROCAL STRATEGIES relating the second person plural with each of the singular forms, X could have three different values, as shown in (6–8).

(6)	/Xō/ ISG	↔	/Xātis/ 2PL	(X = am)
	amō		amātis	

(7)	/Xs/ 2SG	↔	/Xtis/ 2PL	(X = amā)
	amās		amātis	

(8)	/Xt/ 3SG	↔	/Xatis/ 2PL	(X = ama)[7]
	amat		amaatis	

On the other hand, if we were to consider these six forms as a group, we might decide to let X equal the maximum form they all share. In this case, X would have the value it has in (6), coterminous with the lexeme root, and (7) and (8) would be modified as follows:

(7')	/Xās/ 2SG	↔	/Xātis/ 2PL	(X = am)
	amās		amātis	

(8')	/Xat/ 3SG	↔	/Xātis/ 2PL	(X = am)
	amat		amātis	

This latter alternative simplifies strategizing by keeping the value of X constant and not solely dependent on each reciprocal strategy, and thus provides some additional justification for admitting the paradigm to theoretical status. These forms are not just six randomly associated forms. Because the reciprocal strategies each is involved in all need to refer to the interlocking grid of grammatical categories (person and number in this instance) that unite them in a common semantic array of morphosyntactic features, when they are used to modify the form of one and the same lexeme, we are

---

6. Simplified somewhat for our purposes here. These bidirectional strategies in the spirit of Ford, Singh, and Martohardjono, which do not presuppose the existence of paradigms, we will be referring to as "reciprocal strategies" when necessary to distinguish them from "paradigmatic strategies," which do.

7. X takes this value only if long vowels are treated as geminate (see Aronoff 1992). Otherwise, X has the same value as in (6), and the strategy reads as /Xat/ ↔ /Xātis/.

dealing with the PARADIGM of its inflections—"the ways in which this particular lexeme changes its shape as its morphosyntactic features are changed."

Thus we are able to respond to the burden of proof put on us by Ford, Singh, and Martohardjono (1997) (see note 4) by pointing out that paradigms exist for sets of forms like the six we are considering here most fundamentally because of the way in which the forms are inherently interrelated in semantic space. Separate reciprocal strategies written in the format of Ford, Singh, and Martohardjono (1997) will acknowledge this interrelationship by citing points on the grid in each instance. The existence of paradigms on this basis makes possible more economic and efficient strategizing—the extending of a pattern of forms from one lexeme to another—in several ways. First, as shown in (6–8, 7'–8') above, they permit more economic and efficient strategizing by providing an arena within which there can be a common value for X. Second, it may also be that they permit a reduction in the number of separate reciprocal strategies required within this arena. It is to the exploration of this possibility that we now turn.

Let us examine in greater detail the reciprocal strategies involved among the six forms of (5). Thus far we have looked at only three of the total of 15 separate strategies that would be needed to relate each of these six forms to each other bidirectionally—the number would be double that if they were unidirectional![8] Fifteen seems a large number for relating just six forms, but for the full Latin paradigm of 90 finite verb-forms,[9] the number becomes overwhelming: 4,005 separate reciprocal strategies are needed to relate each form to every other form within a single conjugation. One way in which these numbers might be reduced would be to stipulate that within a paradigm, individual strategies should relate only forms that differ by a single morphosyntactic feature, as for example in (2), where –plural and +plural are related. (A series of strategies would then be required for relating forms that differ by more than one feature.)

Applying this restriction to the strategies among the six members separated by features of person and number, their total is reduced from fifteen to seven (assuming for the moment that the features involved are ±speaker, ±addressee, and ±plural).[10] Starting from 3SG (the least marked: –speaker, –addressee, –plural), three strategies relate each of these values to their + counterpart, forming respectively 1SG, 2SG, and 3PL, as in (9).

---

8. The number of ordered (unidirectional) pairs is equal to $n(n-1)$, and the unordered (bidirectional) pairs to $n(n-1)/2$. Thus, as the size of a paradigm expands, the number of strategies expands quadratically. (I have my University of Hawai'i colleague Thomas Ramsey to thank for help with these statements.)

9. Indicative and subjunctive, but not including imperative, or the various forms derived on the "t-stem" (see Aronoff 1992, 1994). Thus perfect, future perfect, and pluperfect passives, accomplished periphrastically using the perfect passive participle and a form of *sum*, are among those excluded.

10. This is the order in which I list the feature values in (9) and (10) below: speaker, addressee, plural.

(9)   +,−,−   1SG
                /Xō/
                 ↕
                /Xat/
        −,−,−   3SG   /Xat/ ↔ /Xant/   3PL   −,−,+
                /Xat/
                 ↕
                /Xās/
        −,+,−   2SG

1PL and 2PL could be tied in either by strategies with 3PL, or with 1SG and 2SG, respectively, or by both, yielding the full array of seven reciprocal strategies relating the six paradigm members that differ by just one feature, as in (10).[11]

(10)   +,−,−   1SG   /Xō/ ↔ /Xāmus/   1PL   +,−,+
                /Xō/   /Xāmus/
                 ↕   ↕
                /Xat/   /Xant/
        −,−,−   3SG   /Xat/ ↔ /Xant/   3PL   −,−,+
                /Xat/   /Xant/
                 ↕   ↕
                /Xās/   /Xātis/
        −,+,−   2SG   /Xās/ ↔ /Xātis/   2PL   −,+,+

Note that there are no reciprocal strategies relating any points in the paradigm other than the nearest neighbors in this arrangement. There are none along any of the diagonals, and there are none relating first and second person forms directly. At most, a sequence of three is necessary to relate 1SG and 2PL, or 2SG and 1PL. A "serial strategy" required to span one of these extremes may not be as complex as would appear at first glance, involving just two intermediate stages, as, for example, those for relating 1SG and 2PL given in (11) and (12).

(11)   /Xō/   ↔   /Xat/   ↔   /Xās/   ↔   /Xātis/
       amō   amat   amās   amātis
       1SG   3SG   2SG   2PL

(12)   /Xō/   ↔   /Xāmus/   ↔   /Xant/   ↔   /Xātis/
       amō   amāmus   amant   amātis
       1SG   1PL   3PL   2PL

Nevertheless, scenarios of this sort seem clumsy and far-fetched to those of us who have learned to recite these inflections in a different order, such as 1SG, 2SG, 3SG, 1PL,

---

11. Note that the total of seven could be reduced to five if one chose only one of the two alternatives for relating 1PL and 2PL, but at the cost of increasing the number of serial strategies required to interrelate certain combinations.

2PL, 3PL, yet who would probably not have to run through the sequence to get from one point to another. Before we dismiss serial strategies such as (11) and (12) summarily, though, it might be a good idea to remind ourselves of the basic question we are addressing, which concerns ultimately how those who are native speakers are able to cope with the high degree of morphological complexity of inflected fusional languages such as this. We are considering here an alternative to building each form up, layer by layer, upon its root, a process that is already notorious for its complexity in this type of language. The alternative we are considering starts with a ready-made form and "bends" it by analogies with other known forms. The assumption is that the inflections of several common verbs (the "exemplary paradigms") are already known, and that the "strategies" we are considering here are formalisms that underlie the analogizing process used to arrive at the inflections of less well-known verbs that also belong to the same paradigm—the same conjugation, in this instance. To avoid all sequences of strategies like those of (11) and (12), a total of 15 reciprocal strategies would be required to produce any of six forms from any of six starting-points. This number could be reduced to as few as five, if some sequences were tolerated, even when there is only one feature of difference.[12] To avoid any sequences when a single feature is involved, a minimum of seven reciprocal strategies are required—still fewer than half of the full 15.

In considering whether certain reciprocal strategies are so rarely used that sequences might be tolerated, a practical step would be to ask what the more frequent starting places might be, as well as the more frequent ending points. The answers to these questions are to be found in discourse. What new form is one likely to want to produce or identify in comparison with a form one has just uttered or heard? No doubt one of the first things to be concluded, if such a study were undertaken, is that to have first and second persons related only via third, as (11) requires, will never do. The deictic switches between "you" and "I" are much more basic and direct. Even if the analysis with ±speaker and ±addressee should ultimately prove to have validity at some level, it is hardly adhered to while strategizing in face-to-face situations. The reciprocal strategies relating first and second person directly, in all combinations of singular and plural, need to be added to the seven already provided in (11): "You-all say this, but I say that. . . ." Similarly, those relating first singular and third plural, first plural and third singular, etc.: "I did such-and-such, but they did so-and-so . . ." (cf. [12]). We would have to conclude that no economies are likely to be achieved short of the full 15 reciprocal strategies required to interrelate all the person-number combinations directly, with the persons treated equipollently rather than privatively.

To keep perspective, we should remember that we have reached this conclusion within just one small corner of Latin verbal inflection—albeit an intensive corner—

---

12. In a scheme that required only five reciprocal strategies, the three plural terms would not be interrelated by strategies; each would be related only to its singular counterpart. Alternatively, 1PL and 2PL would each relate to 3PL, but only the latter would relate to its singular counterpart. It should be made clear that I am not seriously proposing this as a workable alternative, but only as a theoretical possibility if one wished to keep the number of reciprocal strategie as low as possible.

the six person-number combinations of the nonperfect present active indicative,[13] and that this is but one of 15 parallel corners for the various tense-aspect-voice-mood combinations that make up the total of 90 inflections referred to in note 9. Whether economies might be possible in the strategies that tie together these 15 corners (each with its six members) should remain an open question. For the moment, we have determined that none of the 15 reciprocal strategies required to interrelate the six members in this particular corner can clearly be dispensed with in the interests of economy, especially in view of the price that would need to be paid in serial reciprocal strategies as replacements for any single reciprocal strategies dispensed with. Before looking further to see whether we might reduce the number of reciprocal strategies required—which do not assume or depend in any way on the existence of paradigms—we should explore the question as to whether a different sort of strategy is possible within a paradigm, that is, whether conflation of reciprocal strategies might be possible.

2.4 Paradigms as Rules in Parallel

Conflation of reciprocal strategies does in fact appear possible within a paradigm. If the six forms of (5) as members of the same paradigm do share a single value for X in all their reciprocal strategies, only the variable portion of each singulary term need be given in a conflated strategy, together with the combination of morphosyntactic features that each term represents. The X-constant need be given only once to establish its position in relation to the variable portions (although we repeat it in our examples for clarity of presentation). A conflated strategy, as in (13), can bear close resemblance to a traditional paradigm such as (5), with the understanding that its cells are linked together in radial fashion, as signified by the asterisk at the center of (13). The radial connection permits instantaneous switching among terms to create reciprocal strategies of all combinations as needed, accomplished by a "CPU" (central processing unit). We will call such conflated strategies PARADIGMATIC STRATEGIES.

(13)   PARADIGMATIC STRATEGY
        CONJ I: NONPERF PRES ACT IND

	SG		PL
1	Xō		Xāmus
2	Xās	*	Xātis
3	Xat		Xant

This array is thus a notational variant of the fifteen reciprocal strategies discussed in the preceding section. It exists because it fits the set of verbs that are the members of its class. Similar strategies exist in each of the other 14 corners of the overall paradigm, as for example those in (14–17).

---

13. Remembering also that we have examined the inflections of a verb from only one of the four conjugations. The perfect portions of the paradigm—of third conjugation verbs especially—present special problems that are not treated in detail in this chapter (see note 17).

(14) PARADIGMATIC STRATEGY
CONJ I: NONPERF PRES ACT SUBJ

	SG		PL
1	Xem		Xēmus
2	Xēs	*	Xētis
3	Xet		Xent

(15) PARADIGMATIC STRATEGY
CONJ I: NONPERF FUT ACT IND

	SG		PL
1	Xābō		Xābimus
2	Xābis	*	Xābitis
3	Xābit		Xābunt

(16) PARADIGMATIC STRATEGY
CONJ I: NONPERF PAST ACT IND

	SG		PL
1	Xābam		Xābāmus
2	Xābās	*	Xābātis
3	Xābat		Xābant

(17) PARADIGMATIC STRATEGY
CONJ I: PERF PRES ACT IND

	SG		PL
1	Xāvī		Xāvimus
2	Xāvistī	*	Xāvistis
3	Xāvit		Xāvērunt

Paradigmatic strategies (14–17) show the relations among inflections in selected corners while the features of tense, aspect, voice, and mood are being held constant. Their counterparts, which vary the latter while holding the former constant, are illustrated by (18–20).[14]

Three additional paradigmatic strategies like (18–20) are needed for the plural persons, and ten additional strategies like (13–17) for the other tense-aspect-voice-mood combinations. Using the analogies of "corridors" and "corners," any of the six corridor strategies (like 18–20) can be used to access any of the 15 corners (like 13–17). One uses corridors to switch tense-aspect-voice-mood combinations, and corners to switch person-number combinations. Although this may seem an arbitrary division of the six morphosyntactic categories into two sectors, there would seem to be some justification within discourse structure, as has already been noted with respect to the close deictic relation between first and second persons, and the many discourse-based occasions to vary person and number together. Note also in (21) below that person

---

14. It will be noted that I view the Latin "perfect" as an aspect, and the "imperfect" as a tense that I label here as "past." The perfect passive quadrant in (18–20) is vacant of forms, because this combination is accomplished periphrastically, not morphologically. The asterisk, while not centrally located within the overall array, still serves as a reminder that all 15 cells are to be viewed as radially interconnected. The arrows mark instances of syncretism.

(18) Paradigmatic Strategy, first conjugation, 1SG

CONJ I		−PERF		+PERF	
1SG		IND	SUBJ	IND	SUBJ
ACTIVE	FUT	Xābō	⇐	Xāverō	⇐
	PRES	Xō	Xem	Xāvī	Xāverim
	PAST	Xābam	Xārem	Xāveram	Xāvissem
PASSIVE	FUT	Xābor	⇐		
	PRES	Xor	Xer		*
	PAST	Xābar	Xārer		

(19) Paradigmatic Strategy, first conjugation, 2SG

CONJ I		−PERF		+PERF	
2SG		IND	SUBJ	IND	SUBJ
ACTIVE	FUT	Xābis	⇐	Xāveris	⇐
	PRES	Xās	Xēs	Xāvistī	Xāverīs
	PAST	Xābās	Xārēs	Xāverās	Xāvissēs
PASSIVE	FUT	Xāberis	⇐		
	PRES	Xāris	Xēris		*
	PAST	Xābāris	Xārēris		

(20) Paradigmatic Strategy, first conjugation, 3SG

CONJ I		−PERF		+PERF	
3SG		IND	SUBJ	IND	SUBJ
ACTIVE	FUT	Xābit	⇐	Xāverit	⇐⇩
	PRES	Xat	Xet	Xāvit	Xāverit
	PAST	Xābat	Xāret	Xāverat	Xāvisset
PASSIVE	FUT	Xābitur	⇐		
	PRES	Xātur	Xētur		*
	PAST	Xābātur	Xārētur		

and number are especially tightly fused in form, with the seeming plural exponent undergoing full suppletion from person to person: -us in 1st person, -ti- in second, and -n- in third.[15] No other combination of morphosyntactic features is quite so tightly intertwined by their exponents in the Latin verbal paradigm. Person and number do indeed need to covary because of the way in which they are now fused within this system—that is, for internal reasons as well as the external demands of discourse. One can go a step further and say that they have probably become as tightly fused as they are because they have so often been kept together in discourse.

---

15. The exponents of person and number are even more tightly fused in the singular, usually into a single segment.

(21)		SG	PL
	I	-m	-m-u̲s̲
	2	-s	-t̲i̲-s
	3	-t	-n̲-t

The various corridors would also seem to be fairly heavily traveled in discourse: the occasions are many, for example, to shift tense, aspect, voice, and/or mood, while holding 3SG or any of the other person-number combinations constant. To vary both, use of both corridor and corner would be required, in either order.[16]

2.5 Summary of the Case for Paradigms

Of the two strongest arguments for the existence of paradigms, one is semantic and the other is formal. (i) Put most simply, the semantic argument says that inflections belong inherently in a paradigmatic array when their meanings are most clearly stated with reference to an intersecting grid of morphosyntactic categories. (ii) The formal argument notes that whereas inflections could vary independently, they in fact do not. They vary interdependently in a way that adheres to the Paradigm Economy Principle (Carstairs 1987).

In sections 2.3–4, we have noted several additional arguments for the existence of paradigms. (iii) The use of reciprocal strategies (Ford, Singh, and Martohardjono 1997) to relate inflections is simplified if the value of X can be held constant within a frame such as the paradigm—X being the root portion of a lexeme that does not vary from inflection to inflection—and if it is not permitted to vary according to the two items being related. (iv) Furthermore, a constant X makes it possible for strategies to be conflated and interconnected radially, while being placed in an arrangement dictated by the morphosyntactic properties involved, thereby bearing close resemblance to traditional paradigms. This is the basis for the title of 2.4, "paradigms as rules in parallel": when paradigms are understood as conflations of multiple reciprocal strategies they are in effect "rules in parallel"—rules of an analogizing sort. (v) Whereas over 4,000 reciprocal strategies would be required to interrelate the 90 inflections of a Latin verb, only a score of these "paradigmatic strategies" are needed to accomplish the same end if the six morphosyntactic categories involved can be separated into two sectors: those involved in the corners, and those involved in the corridors. (vi) The X in common among the members of a paradigm brings us a bonus: insight into why it is that morphological solutions prevail over phonological solutions in those situations where they can be seen as being in contention. This is the subject of the last major section of this chapter.

---

16. It is, of course, possible that language users switch directly among all combinations of features without division into subareas (our corners and corridors), even in such large paradigms. (Even more herculean tasks are sometimes imputed to speaker-hearers by those who would build up complex forms from their roots.) There may also be differences among users in the pathways used, and with respect to those that are especially well-trodden.

2.6 When Morphological Classes Are Created

Inflection does not necessarily require that there be different morphological classes, that is, noun declensions, verb conjugations, and so forth. In many languages, inflections fall into a single paradigm. This seems to be true especially of languages said to be agglutinative. It is only when we have determined the extent of X for each verb and we are left with differing residues for different verbs (as with our Latin verbal data), that we find it necessary to group the verbs into different conjugations according to these residues, as in (22).

(22)   VALUES FOR X AND RESIDUES FOR REPRESENTATIVE VERBS OF EACH OF THE FOUR LATIN CONJUGATIONS

X	RESIDUE	X	RESIDUE	X	RESIDUE	X	RESIDUE	
am	ō	doc	eō	trah	ō	aud	iō	1SG
am	ās	doc	ēs	trah	is	aud	īs	2SG
am	at	doc	et	trah	it	aud	it	3SG
am	āmus	doc	ēmus	trah	imus	aud	īmus	1PL
am	ātis	doc	ētis	trah	itis	aud	ītis	2PL
am	ant	doc	ent	trah	unt	aud	iunt	3PL
'love'	1ST CONJ.	'teach'	2ND CONJ.	'drag'	3RD CONJ.	'hear'	4TH CONJ.	

There is no way we can obtain identical residues, for this would require enlarging X to a point where it no longer had a constant value, such as occurs when we obtain the residues of (21).[17]

Traditional grammars exemplifying the WP model of description sometimes list the forms in (22) as whole words, but use a different typeface for the residue portions, and present them as being exemplary paradigms for their respective conjugations. This may seem contradictory: why should a whole-word approach indicate what appears to be an internal boundary using a different typeface? As noted in 1.1, this is an implicit seam that reflects what we hold constant and what we vary in modeling other words on these words. It is determined by how the whole words relate to each other, and is in no sense a determinant. Highlighting it with typefaces is evidently intended as a pedagogical aid, but is probably of dubious value that is limited to work with written materials. It is but one of a number of features in traditional grammars that go beyond the WP approach and reflect the alternative approach discussed in 1.1, that of building com-

---

17. As students in my Spring 1998 morphology class discovered while writing full paradigmatic strategies for all four verbs in (22), verbs with irregular perfect stems (primarily those of the third conjugation) present another complication in that they require a variant for X (call it X') in strategies involving the perfect aspect. Thus, for *trahō* 'drag', X = trah, X' = trāx (trācs). That this cannot be solved simply by shrinking X (to *tr* in this instance) is shown by verbs that reduplicate, such as *caedō* 'cut', where X = *caed,* X' = *cecīd,* etc. The use of X' variants makes it possible to treat deeper and less regular root mutations separately from—and without proliferating—the main strategies, which deal with the more superficial but regular features of the inflections.

plex forms up from their roots. Grammars that include these non-WP bits and pieces are catering to our quite human fascination with reflective analysis.

2.7 Phonological Solutions as Alternatives

Both of the alternative models discussed by Hockett (1954) are morpheme-based, and one of these, the Item and Process (IP) model, attempts to reduce variants to more abstract basic forms that can be combined and adjusted by process rules. In early generative grammar, such solutions were viewed as being phonological, under a broadened view of phonology, and as preferred, in part because they are often able to eliminate morphological classes by building indicators of class membership into the phonetic substance of the more abstract base forms. At the same time, it has been no secret that such base forms often turn out to be closer to those reconstructed for earlier stages of a language than are the contemporary surface forms—that the establishment of well-motivated base forms constitutes an exercise in internal reconstruction.

As an example of the use of this type of analysis in order to eliminate morphological classes, Pearson (1977b:60) includes the following exercise under "Latin phonology (advanced), Reexamine the data . . . [roughly equivalent to that of (22)—BWB]. Describe the four verb conjugations so that all can be merged into a single undifferentiated class. To do this, it will be necessary to assign the theme vowel for each conjugation to the verb stem, set up a single set of affixes for all four conjugations [identical to those in (21)[18]—BWB], and posit a series of process rules to make the necessary phonological adjustments." The same workbook also includes an equivalent exercise for Spanish (Pearson 1977b:7), in which students are asked to "devise at least two different statements that will account for the present tense formation. . . ." The data include present indicative paradigms and infinitives for two verbs in each of the three Spanish conjugations. A conjugational statement given in the instructor's key requires—in addition to the stem for each verb, a knowledge of their conjugation membership and a total of 21 different suffixes, seven for each conjugation. In this solution, the infinitives end in -ar, -er, and -ir. A phonological alternative is presented as the preferred solution because it requires only seven suffixes and four rules, with no conjugation membership. In this solution, all infinitives end in -r.

Morphological and phonological solutions have been pitted against each other in a number of other contexts. As just one example, many Oceanic languages have developed longer and shorter verb forms through loss of final consonants—longer forms in contexts in which the consonant was protected by an affix and preserved, and shorter where it was lost. A morphological solution finds that reanalysis has occurred, and that original stem-final consonants are now part of suffixes that are added. To derive longer forms, it is now necessary to group shorter verb forms into morphological classes according to the consonant in the suffix to be added. The alternative phonological solution would have for each verb a single abstract form that includes the final consonant and that underlies both longer and shorter forms, with a rule of final consonant deletion

---

18. With the exception of 1s, which is -ō for the NONPERF PRES ACT IND.

to derive the shorter ones. This is documented for Māori in Hohepa (1967) and Hale (1968, 1973), where there is increasing evidence that the morphological solution is more in accord with speakers' intuitions and with continuing developments in the language.

The same would appear to be true for Latin and the various Romance languages. Spanish infinitives, for example, are commonly viewed as ending in -V*r*, reflecting a morphological analysis in which those in *-ar* are becoming predominant. How can one account for the staying power of morphological solutions, in spite of the seeming economy offered by phonological solutions?

2.8 Why Morphological Solutions Prevail over Phonological Solutions

The answer to this question was hinted at at the close of 2.5. It is because, whenever possible, the whole-word analogizing of morphology uses paradigmatic strategies that presuppose a common and constant X that is defined paradigmatically. When the residue left by this X-constant varies for different groups of verbs, morphological classes exist, and the X-constant coincides with the stem called for by the morphological solution. Thus, these classes and the X-constant depend on each other, just as both depend, in turn, on whole-word strategizing.

REFERENCES

Aronoff, Mark. 1976. *Word formation in generative grammar.* Cambridge, Mass.: MIT Press.
————. 1992. Stems in Latin verbal morphology. In *Morphology now*, ed. by Mark Aronoff, 5–32. Albany: State University of New York Press.
————. 1994. *Morphology by itself: Stems and inflectional classes.* Cambridge, Mass.: MIT Press.
Bender, Byron W. 1998. The sign gravitates to the word. In *Productivity and creativity: Studies in general and descriptive linguistics in honor of E. M. Uhlenbeck*, ed. by Mark Janse with the assistance of An Verlinden, 15–25. Berlin: Mouton de Gruyter.
Carstairs, Andrew. 1987. *Allomorphy in inflexion.* Beckenham: Croom Helm.
Comrie, Bernard. 1981. *Language universals and linguistic typology.* Chicago: University of Chicago Press.
Ford, Alan, Rajendra Singh, and Gita Martohardjono. 1997. *Pace Panini: Toward a word-based morphology.* American University Studies 13, Linguistics, vol. 34. New York: Peter Lang Publishing.
Hale, Kenneth L. 1968. Review of Hohepa 1967. *Journal of the Polynesian Society* 77:83–99.
————. 1973. Deep-surface canonical disparities in relation to analysis and change: An Australian example. In *Current trends in linguistics,* vol. 11, ed. by Thomas A. Sebeok, 401–458. The Hague: Mouton.
Hockett, Charles F. 1954. Two models of grammatical description. *Word* 10:210–231. (Reprinted in *Readings in Linguistics,* ed. by Martin Joos, 386–399. New York: American Council of Learned Societies.)
Hohepa, Patrick W. 1967. *A profile generative grammar of Maori.* Indiana University Publications in Anthropology and Linguistics, Memoir 20. *International Journal of American Linguistics.* Baltimore: Waverly Press.
Matthews, P. H. 1974. *Morphology: An introduction to the theory of word-structure.* Cambridge: Cambridge University Press.
————. 1991. *Morphology.* 2d ed. Cambridge: Cambridge University Press.
Pearson, Bruce L. 1977a. *Introduction to linguistic concepts.* New York: Knopf.

————. 1977b. *Workbook in linguistic concepts.* New York: Knopf.
Robins, R. H. 1959. In defence of WP. *Transactions of the Philological Society* 1959:116–144. Reprinted in Diversions of Bloomsbury: Selected writings on linguistics, by R. H. Robins, 47–77. Amsterdam: North-Holland.
Spencer, Andrew. 1991. Morphological theory. Oxford: Basil Blackwell.
Starosta, Stanley. 1991. No more boundaries, no more phemes. Paper presented to the Tuesday Seminar, Department of Linguistics, University of Hawai'i, May 5.
————. 1995. Word formation. Notes from class at University of Hawai'i, Spring semester.
Uhlenbeck, E. M. 1953. The study of word classes in Javanese. *Lingua* 3:322–354.
————. 1992. General linguistics and the study of morphological processes. *Oceanic Linguistics* 31:1–11.
Wurzel, Wolfgang. 1989. *Inflectional morphology and naturalness.* Dordrecht: Reidel.

3: SOURCES OF PROTO-OCEANIC INITIAL PRENASALIZATION: THE VIEW FROM OUTSIDE OCEANIC

Lawrence A. Reid
University of Hawai'i

1. Introduction

One of the foremost contributions of Stan Starosta to the study of language change has been his insistence on motivated analyses of the structures under consideration within a constrained linguistic theory in order to provide sets of comparably analyzed data. His version of dependency grammar, Lexicase, provides such a framework for the study of syntactic change. In this paper, I attempt to show, using Lexicase analyses of the relevant structures, that certain syntactic changes in early stages of Austronesian languages, specifically the development of determiners from earlier demonstrative nouns, were accompanied by varying patterns of loss of the prepositions that had earlier functioned as either case markers of their following noun phrases, or had connected them to following (nominal) relative clauses. In some languages, the preposition was lost. In others, it became fused with the demonstrative noun, while in others it became a proclitic to what was earlier the head of the following relative clause. It was the latter process, still in operation in some of the Minahasan languages, that ultimately led in Proto-Oceanic to the development of a set of initial prenasalized consonants.

These patterns of grammaticalization are not at all uncommon in languages of the world, and have been discussed widely in the literature (e.g., Hopper and Traugott 1993; Heine, Claudi, and Hünnemeyer 1991), but usually without the benefit of an explicit grammatical framework within which to determine the lexical categories of the forms under consideration. Heine, Claudi, and Hünnemeyer in fact call for a "framework of linguistic descriptions that is not confined to static, discrete units such as word classes or constituent types but rather includes dynamic entities such as chains of grammaticalization" (1991:231–233). Such a framework, they believe, is required because grammaticalization chains are continua, with fuzzy, nondiscrete boundaries between the stages. Such analyses imply "hybrid" forms, such as "part verb, part preposition," or "part demonstrative, part determiner," for example, at intermediate stages of the grammaticalization process. Starosta's version of dependency grammar rejects intermediate, hybrid forms that are neither one thing nor the other. It permits a constrained set of only seven lexical categories (noun, verb, adjective, adverb, pre/postposition, conjunction, and sentence particle), which are claimed to have universal properties. Specific forms cannot be a member of more than one lexical category at the same time. The theory, however, does allow for sets of homophonous, derivationally related forms. It is the availability to children of this process of "zero-derivation" that provides them with one of the mechanisms by which they can reinterpret syntactically ambiguous structures in the process of acquiring their language.

30

Changes in lexical category often result in changes in dependency relations. Thus a noun that is the head of its construction may end up as a determiner, a form that can only be a dependent. Heine, Claudi, and Hünnemeyer (1991:220) note also that "dependency forms a parameter that is of immediate relevance to our discussion. Whether a given entity governs or is governed by another entity is likely to determine its fate in the process of metaphorical use and of grammaticalization." Lexicase provides us with a dependency grammar framework within which to describe such grammaticalization processes.

Proto-Oceanic (POC), as reconstructed by various scholars beginning with Dempwolff (1969) and continuing through to the recent work of Ross (1988), has been shown to contain a contrasting series of obstruents, reconstructed by Ross (1988:93) as *p/b, *t/d, and *k/g, the voiced members of which had their origins in prenasalized versions of their voiceless counterparts, *mp, *nt, and *ŋk. In addition, there were two other contrasting pairs: *r/dr, traditionally symbolized as *d/nd, and *c/j, the second member of which developed from prenasalized *ns. These prenasalized consonants could occur only in syllable-initial positions, either word initially or word medially.

Ross's (1988) work provides us with the most thorough overview to date of this so-called "oral grade–nasal grade" phenomenon, including an excellent discussion of what have been termed the "cross-over" reflexes, where some Oceanic languages show an oral-grade reflex of a form reconstructed with a prenasalized obstruent, while others show a nasal-grade reflex of a form reconstructed with a nonprenasalized obstruent. He also provides a full discussion of what he refers to as the "fortis grade–lenis grade" developments in many Oceanic languages that have contributed to confusion in understanding the reflexes of the oral grade–nasal grade phenomenon. In the course of his discussion, Ross (1988:39–45) suggests that one of the explanations for cross-over had its origins in Pre-POC and possibly even earlier. He suggests that there were alternating forms attributable to morphophonemic variation that were "quite possibly fossilized" by POC times.

In this essay, I review earlier proposals for the origins of prenasalization in Proto-Oceanic, and examine Ross's explanation for the variation associated with prenasalization. I suggest that the explanations that Ross and others have provided do not adequately account for the variation, at least in some of the items they discuss. I attempt to show that it is more likely that the development of at least some of the prenasalized obstruents in word-initial positions[1] in Proto-Oceanic was the result of a process of grammaticalization in the structure of the noun phrase. Under this process, a nasal that at an earlier stage formed the final consonant of a determiner marking a definite noun became prefixed to definite nouns. Contexts in which indefinite nouns were

---

1. The origin of prenasalized obstruents in word medial position has been discussed fairly fully in the literature, beginning at least with Dempwolff (1969:96–115) who posits a "facultative" or optionally occurring nasal for reconstructed forms whose reflexes were ambiguous, some languages having doublets with and without the nasal, with other languages showing either one or the other. The situation is somewhat analogous to the Oceanic cross-over phenomenon. At any rate, the sources of POc medial prenasalized consonants appear to be quite different from those in initial position and will not be further discussed in this paper.

required would not have allowed nouns having initial prenasalized obstruents to occur in them. After definiteness became marked by some grammatical feature other than prenasalization, this alternation between prenasalized definite nouns and nonprenasalized indefinite nouns appears to have persisted as free variation between the two classes of forms, providing a reasonable explanation for the variation in POc that could have resulted in cross-over.

2. Earlier Views of the Sources of POc Prenasalization

Although there are numerous publications examining the development of nasal-oral grade consonants in Oceanic languages (see, for example, Lynch 1975, Grace 1990) very few go beyond the level of Proto-Oceanic to try to account for their development in Proto-Oceanic. There have been two opposing views expressed regarding the nature of prenasalization in Western Austronesian languages. One treats prenasalization as a process associated with certain verbal affixes that, along with those affixes, is reconstructable to their parent language. The other treats prenasalization as originally a phonetic feature that only after the breakup of POc acquired grammatical significance in Western Austronesian languages.

2.1 Prenasalization as a Phonological Process

Dempwolff (1969:30–33), in his discussion of the phonologies of his "Indonesian" languages, Tagalog, Toba-Batak, and Javanese, notes that all three languages share a phonological process affecting the initial segments of words. He labels the processes NASAL ACCRETION for cases in which a homorganic nasal preceded the initial segment, and NASAL SUBSTITUTION for those cases in which a (homorganic) nasal replaced the initial segment. He further notes that these processes are associated with certain prefixes in Tagalog (*ma-*, *na-*, and *pa-*) and in Toba-Batak (*ma-* and *pa-*), but occur without an associated prefix in Javanese. He labels both processes PRENASAL-IZATION and reconstructed the process to the parent language, Proto-Extra-Formosan (PEF).[2] In his summary (Dempwolff 1969:124), he states that the original grammatical function of prenasalization (and also the grammatical function of the nasal increment responsible for word medial nasal-obstruent clusters) remains a problem.

2.2 Prenasalization as a Phonetic Feature

Milner (1965), apparently working with the assumption that Oceanic (the "Eastern Austronesian" of his title) and Western Austronesian are coordinate branches of the Austronesian family, reexamines the nature and distribution of prenasalization in the two branches. Although Milner notes that in the Western languages from which Dempwolff cited data, nasal accretion has "morphophonemic" or grammatical functions only in initial position, he claims that in medial position it is only a "phonetically distinct feature", without apparent grammatical function. In Oceanic, moreover, prenasalization

---

2. Dempwolff's UIN (Urindonesisch).

had no grammatical function; it was a phonetically distinct feature initially as well as medially. He notes also the presence in Hova (an isolated "peripheral" language spoken in Madagascar) of initial prenasalized consonants "in contexts [that] cannot be accounted for by reference to the regular morphophonemic processes [that] Hova shares with other Western Austronesian languages" (Milner 1965:428). Prenasalization in initial position in Hova then, as in Oceanic languages, is also a phonetically distinct feature rather than a morphophonemic process. He concludes that "it is therefore reasonable to suppose that at a very early date throughout the entire Austronesian region there was (as there still is frequently in contemporary Oceanic languages) the possibility of a phonemic distinction between prenasalized voiced stops in free distribution and nonprenasalized voiceless stops, also in free distribution" (Milner 1965:428). Initial prenasalized obstruents then, according to Milner, should be reconstructed to PAN. Their shift in the "central" Indonesian languages, from phonetically unrestricted distribution to grammatically conditioned distribution, was an innovation in those languages.

With the more commonly accepted subgrouping hypotheses that today place Proto-Oceanic in a far more subordinate position in the Austronesian family tree, Milner's arguments no longer carry much weight. There is no evidence from earlier stages of Austronesian that the grammatical functions associated with prenasalization in the Western languages were subsequent developments from earlier phonemically distinct prenasalized voiced stops.

3. The Prefix *maN- in Pre-Oceanic

Dempwolff's view that prenasalization of initial obstruents developed out of the accretion of the final nasal element of the prefixes *maN- and *paN- provided the foundation for Ross's explanation for some of the forms that appear to constitute evidence for the cross-over phenomenon in Oceanic languages. Ross notes (1988:41) that, in Western Austronesian languages, addition of the prefixes to stem-initial voiceless obstruents always results in homorganic nasal substitution, whereas their addition to stem-initial voiced obstruents results in either a nasal-obstruent sequence or homorganic nasal substitution. It is to this (predictable) alternation, Ross claims, that at least some cases of cross-over can be attributed.

The forms in question (cited from Ross 1988) are:

Gela, West Guadalcanal, Talise *mabulu* 'rotten'. Ross suggests that this form reflects Pre-POC *mampuRu(k) (< PAN (?) *maN- + buRuk).[3]

Bilbil, Takia *madid*; Manam *madidi*, Kaiep *marir*, Kairiru *meřiř*, Ulau-Suain *madid*; Tolai, Duke of York *madiriŋ*; and Halia (Haku) *maririŋ* 'cold'. Ross suggests that these forms reflect Pre-POC *mandindiŋ 'cold' (< *mandiŋdiŋ < PAN (?) *maN- + diŋdiŋ 'cold').

---

3. Regular devoicing of PAN voiced bilabial (and velar) stops is one of the defining features of POC.

Bariai *mad-madid;* Kilenge, Maleu *mari;* Tami *moji;* Sio *madi;* Mangap *meder;* Mari, Adzera *munti;* Wampur *monti;* Sukurum *mundi;* Dangal *mundik;* Wampar *mondeŋ;* Sirak *mindiŋ;* Yalu *midiŋ(te)* 'stand'. Ross suggests that all of these forms reflect POC *mandiʀi 'stand' (< *mandiʀi < PAN (?) *maN- + diʀi 'stand').

According to Ross (1988), each of the above forms results from prenasalization through the prefixing of *maN-. The difficulty with this explanation is that none of the meanings 'rotten', 'cold', and 'stand' satisfy what appear to be the semantic conditions for the appearance of *maN-. Ross states (as does Dempwolff 1969) that the original meaning of *maN- is unclear. However, in none of the Western Austronesian languages that I have looked at is this a prefix that typically occurs on descriptive terms such as 'rotten' and 'cold'. Nor does it occur on bodily-action verbs such as 'stand'. The former stems typically require a reflex of the stative prefix PAN *ma-, and in fact a number of other Western Melanesian languages cited by Ross still retain a frozen *ma-* prefix on them (without evidence of prenasalization). There is no evidence that either Pre-POC or POC had alternate forms of the stative prefix, either *ma- or *maN-. It is only in the Southeast Solomonic subgroup that the apparently prenasalized forms occur. Thus a local explanation is probably better than appealing to variation in Pre-POC. Either some now unrecognizable analogical process in the immediate parent of the languages introduced a prenasalized stop, or perhaps the medial *-b-* can be attributed to an intervocalic lenition process that has not gone all the way to *-v-*.

Prenasalization of the initial segment of the stem for 'cold' is more likely to be the result of a spread of the medial prenasalization to the first segment of the reduplicated syllable, than of the affixation of *maN-*. Thus PMP *ma-diŋ.diŋ > *ma-di.ndiŋ > *ma-ndi.ndiŋ > POC *ma.ndi.ndiŋ.

Bodily-action verbs such as 'stand' are typically affixed with a reflex of PAN *mu-/ -um-. The prefix *maN-, at least in PMP if not in PEF, derived verbs with distributive action. They could be intransitive, such as 'cognate object' verbs, whose derivational source was nouns of the sort that could be gathered or collected, such as 'wood', 'taro', and so forth. But more commonly they were antipassive (pseudo-transitive) verbs having an actor subject and an implied, if not always expressed, additional NP translatable as an indefinite but specific object.[4] The source of such verbs was grammatically transitive forms requiring a definite Patient. The semantics of 'stand' do not fit either type. However, I have no alternate explanation for the data Ross provides.

Other forms that Ross (1988:41–42) attributes to *N- substitution are also questionable. In particular, the bilabial nasal substitution in forms that are reflexes of PAN *buni 'hide' (> Pre-POC *muni) with intransitive interpretation (i.e., 's.o. hides'— another bodily-action verb) are far more likely to be the result of affixation with PAN *mu-/-um- than with *N-, because this type of nasal substitution involving labial obstruents (e.g., *b-um-uni > *muni) is widespread in Western Austronesian languages (including Formosan) and is independent of the nasal substitution processes

---

4. Such NPs, although indefinite, could not refer to generic, or nonspecific, entities of the sort that could only occur as the notional object of grammatically intransitive verbs derived with the *-um-* infix.

associated with *maN- that developed at a much later stage in the family, and affected obstruents at all points of articulation.

4. Proto-Oceanic Phrase Structure

Despite the comments in section 3, it is likely that at least some of the forms reconstructable to POc with prenasalized stem-initial consonants had their origin in verbs derived with a *maN- prefix, as suggested by Dempwolff and also by Ross. Some POc nouns with prenasalized stem-initial consonants could also have resulted from affixation of the *paN- instrumental prefix, which brought about similar morphophonemic changes to stem-initial obstruents. However, these two affixes alone cannot account for the majority of Pre-POc forms reconstructed with prenasalized stem-initial obstruents, such as *mpanij 'wing', *mpune 'pigeon', *mpoʀok 'pig' (each of which also has at least one Oceanic reflex without a prenasalized initial consonant); *mpuaq 'areca nut', *mpoɲi 'night', *ndaʀaq 'blood', *ndanum 'water', *ŋkudu 'thunder', *ŋkapu 'mist', etc. (cited from Ross 1988:34–39). To account for forms such as these, it is necessary to consider the structure of the noun phrase from which the POc noun phrase developed.

4.1 Overview

In section 4.2, I outline relevant aspects of the structure of the Proto-Oceanic noun phrase as discussed by Ross (1988:96–100). In section 5, I present evidence for corresponding structures from Western Austronesian languages, which appear to be ancestral to the Oceanic structures. Finally, in section 6, I show how, in at least some western Austronesian languages, processes of phonologization have resulted in prenasalization of initial obstruents, and I draw the conclusion that the same processes probably brought about prenasalization of stem-initial consonants in Pre-POc.

4.2 The Proto-Oceanic Noun Phrase

Ross (1988) compares the noun phrase structures of the languages of the Central-Malayo-Polynesian and South Halmahera-West New Guinea subgroups with Oceanic languages, and concludes that, whereas the former languages have innovated certain aspects of the noun phrase, Proto-Oceanic has retained structures that are widespread in other Western Austronesian areas. In particular, he notes the retention in Proto-Oceanic of preposed articles that mark the head noun as common or personal, and the order of possessed-possessor where a possessor occurs in the noun phrase. Crowley (1985) reconstructs for Proto-Oceanic a "Type III" noun-marking system, with "two classes of common nouns, determined by whether the noun is marked by a reflex of *na/a, or by zero, in most (but not all) syntactically unmarked environments" (Crowley 1985:173–174, 176). He considers the form *na/a to be "actually a marker of a specific or a definite noun phrase (or something semantically close to either of these functions" (Crowley 1985:176–177). He also looks for Oceanic evidence to account for the alternate forms *na and *a, but is unable to determine whether, in Proto-Oceanic, they were separate morphemes with related but separate

meanings, or whether they were variants of a single morpheme. I claim that, at least in Proto-Extra-Formosan, they were phonologically conditioned allomorphs. Expressed in a dependency grammar stemma, the structure of Pre-POC *na/a *mpoʀok 'the pig' would appear as in (1).

(1)

```
        ┌──────────────┐
      ┌─┘          *mpoʀok              'the pig'
   *na/a             [N]
   [Det]
   [–prsn]
```
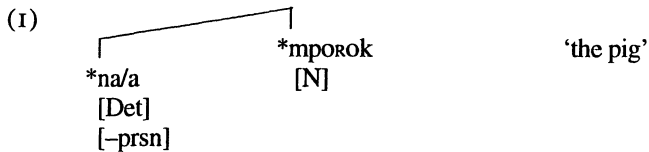
5. The Proto–Extra Formosan Noun Phrase

Western Austronesian languages show a wide variety of forms that appear to be determiners preposed to head nouns.[5] Although the forms vary considerably, the distinctions they mark frequently include those reconstructed for POC, that is, common ([-prsn]) versus personal ([+prsn]). Nominative noun phrases in PAN, although not case-marked and possibly not containing a determiner (Starosta 1993), probably followed a case-marking preposition in PEF, the ancestral language of the Philippines and other non-Formosan languages. The form of the nominative case-marking preposition before common nouns was probably *i, still retained as the initial formant of nominative pronouns and/or demonstratives in languages such as Ilokano and Tagalog, and as a nominative noun phrase marker in some environments in Pangasinan and other languages (Reid 1978, 1979). A dependency stemma of the exocentric[6] prepositional phrase *i balay 'the house' in PEF would be as in (2).

Because nominative complements usually required head nouns to be interpreted as definite,[7] demonstrative nouns ('this one', 'that one') and personal pronouns would have frequently appeared in the head-noun position following the preposition, as in (3a, b). This construction, however, was ambiguous (because *i was also a locative case-marking preposition), and probably lost its case marking function fairly early, but not before the preposition had become cliticized to demonstratives and pronouns that followed, where it remained in some languages as a distinguishing nominative (and/or predicative) formant, as in (4a, b). In other languages, however, even these

---

5. Much of the discussion in this section, although differing in a number of details, was presented in an informal way in Reid 1978. That study provides more extensive Philippine language evidence for the forms cited.

6. Exocentric constructions are shown in lexicase dependency grammar stemmas with horizontal rather than slanting lines joining the head of the construction and its dependent.

7. That not all nominative complements were interpreted as definite is clear from structures with existential verbs, as in the following Bontok sentence, which still retains a reflex of *i as a postclitic to the existential verb wáda 'there is', here interpreted as a possessive: Wadʔa-y ábuŋ-da 'They have a house. (lit. There is a house of theirs.)' (Reid 1979:14). Mamanwa similarly maintains a reflex of *i as a postclitic to the negative existential verb wara 'there is none' before an indefinite nominative NP: Wara-y daro. 'There is no plow.' (Svelmoe and Svelmoe 1974:61).

forms became unmarked and could appear as plain nouns, without prepositions or determiners, in positions where they could be interpreted as either Nominative or predicate, as in (5a, b).

(2)

  *i   'the house'
  [P]   *balay
  [Nom]   [N]
        [−prsn]

(3a)

  *i   'this one'
  [P]   *ni
  [Nom]   [N]
        [+dmns]

(3b)

  *i   'I'
  [P]   *akən
  [Nom]   [N]
        [+prnn]

(4a)   *i.ni   >   Tagalog *ini* 'this one'
    [N]
    [+dmns]
    [Nom]

(4b)   *i.akən   >   Ivatan *yakən* 'I'
    [N]
    [+prnn]
    [Nom]

(5a)   *na   >   Bontok *ná* 'this one'
    [N]   e.g., *ʔiláʔəm ná* 'Look at this.'
    [+dmns]

(5b)   *akən   >   Cotabato Manobo *akən* 'I'
    [N]
    [+prnn]

In PEF, head nouns, including demonstratives but possibly not personal pronouns, could be modified by a following relative clause, the predicate of which could be either a noun or a verb. A relative clause was a reduced clause (missing its Nominative constituent) and formed part of an exocentric construction, headed by a preposition, the form commonly referred to in the literature as "the ligature." The form of this preposition in PEF was probably the phonologically conditioned *na/a, with *na occurring following stems ending in a vowel, and *a occurring following stems ending in a consonant.[8] The

stemmas in (6) and (7) illustrate the structure of (nominal) relative clauses dependent on common noun and demonstrative noun heads, respectively.

(6) 'the big house' (lit. 'the house which is a big one')

```
┌─────┬─────────┐
*i             ┌─────────┐
[P]    *balay   ┌─────────┐
[Nom]  [N]     *a         │
       [–prsn]  [P]        *dakəl
                [+rltv]    [N]
                           [prdc]
```

(7) 'that big one' (lit. 'that one which is a big one')

```
┌─────┬─────────┐
*i             ┌─────────┐
[P]    *ná      ┌─────────┐
[Nom]  [N]     *na        │
       [+dmns]  [P]        *dakəl
                [+rltv]    [N]
                           [prdc]
```
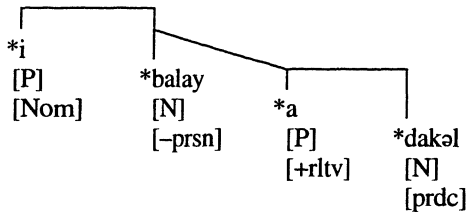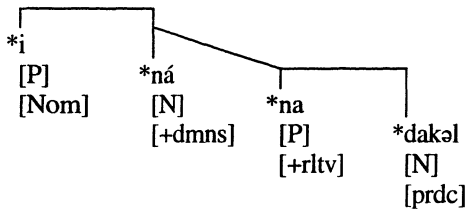
In some of the northern Philippine languages, such as those belonging to the Central Cordilleran subgroup (including Bontok, Kankanay, and Ifugao), prepositions became postclitics to the preceding noun, with subsequent loss of the final vowel, as in (8). Now, however, BON, KNK, IFG *nan* is no longer a sequence of noun plus preposition. It has become grammaticalized as a single morpheme functioning as a definite determiner, as in (9). In Bontok (and Kankanay), the demonstrative *ná* has replaced the earlier close-to-speaker form *ni, and now means 'this one', as in (5a), and a new preposition introducing relative clauses, *ay*, has developed, as in (10). This preposition, in normal speech, is also phonologically attached to the preceding noun.

The process of reduction of a morphological sequence of demonstrative noun plus preposition to a determiner shown in (9) continues in Bontok, with casual speech forms of (10) occurring as (11), in which the sequence *ná* + *-ay* occurs as a single morpheme *náʔay*, often reduced to *nay*, and constituting part of the complex determiner *nannay*.

---

8. This reconstruction is different from that given in Reid 1979, where the forms were *ŋa/a. The latter reconstruction was implicit in the work of Wolff (1967:72–24), and is cited by Zorc (1977:230): "The shape of the markers with final *n*, . . . which are probably cognate with forms [that] have final *ng* in other languages, indicates a change of *ng* to *n* under certain conditions. . . . [W]hat the conditions are is not clear." Blust (1974:7) also reconstructed a "linker" of the same shape for PAN. He stated, "Apart from the *ni phrase, at least one other feature of organization transcending the level of the word can be reconstructed and assigned with equal confidence to Proto-Austronesian. Thus the use of a linker *ng(a) to connect two numerals in a multiplicative relationship is attested in a number of widespread Austronesian languages." Foley (1976) reconstructed a ligature *ña (with a variant *-ng occurring after vowels) for PAN. The evidence for the reconstruction given in this study is presented in Reid 1983.

(8) Pre-Bontok 'that big one' (lit. 'that one which is a big one')

```
      ┌──────────┬──────────┐
     *ná        *-n
     [N]        [P]        *dakdakəl⁹
     [+dmns]    [+rltv]     [N]
                            [prdc]
                            [−plrl]
```

(9) Bontok 'the big one'

```
      ┌──────────┐
     ┌       dakdakəl
    nan       [N]
    [Det]     [−plrl]
    [+dfnt]
```
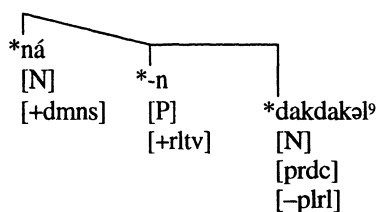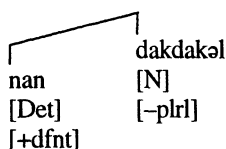
(10) Bontok *nan náʔay dakdakə́lay ʔábuŋ* 'this big house'
      (lit. 'the this one which is a big one which is a house')

```
      ┌──────────┬──────────┬──────────────────┐
     ┌          ná         ┌
    nan         [N]         -ay        ┌──────────┬──────────┐
    [Det]       [+dmns]     [P]       dakdakə́l   ┌
    [+dfnt]                 [+rltv]    [N]        -ay
                                       [prdc]     [P]
                                       [−plrl]    [+rltv]    ʔábuŋ
                                                             [N]
                                                             [prdc]
```
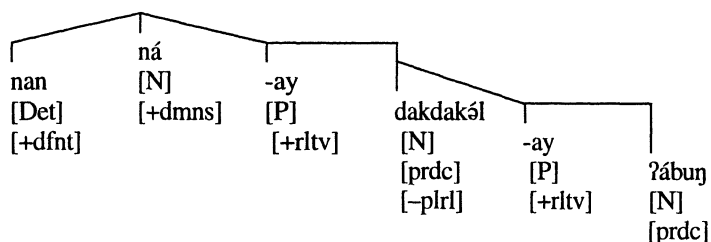
(11) Bontok *nannay dakdakə́lay ʔábuŋ* 'this big house'
      (lit. 'this big one which is a house')

```
      ┌──────────┬──────────────────┐
     ┌       dakdakə́l   ┌──────────┬──────────┐
    nannay      [N]         -ay
    [Det]       [−plrl]     [P]       ʔábuŋ
    [+dfnt]                 [+rltv]    [N]
                                       [prdc]
```
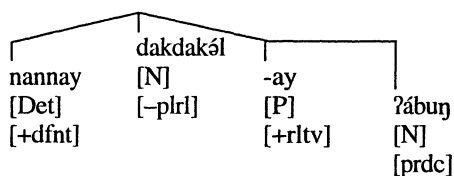
Precisely the same kinds of changes have operated to produce the well-known *aŋ* and *naŋ* determiners found in Tagalog, except that there was an innovation in the form of the relative preposition (ligature) *na. The postconsonantal variant *a was lost, and *na was generalized to all positions. Subsequently, the initial nasal of *na became a velar in postvocalic position (12).[10] The relative preposition became a postclitic to the preceding noun, losing its final vowel (13), although a relic of the earlier stage

---

9. Many Bontok "adjectival nouns," such as *dakəl* 'big', are now obligatorily marked for number, thus *dakdakəl* 'big one', but *danakkəl* 'big ones'.

10. This was an innovation, the cause of which is still obscure.

remains frozen on the Tagalog plural determiner *maŋa*. Nominative noun phrases lost their case-marking, and the demonstrative noun plus preposition sequence *a-ŋ became a morphologically simple determiner *aŋ* (14).

(12) Pre-Tagalog 'that big one' (lit. 'that one which is big')

```
  ┌─────────┬──────────────
 *i         *a ─────────┐
 [P]        [N]         *ŋa ──────────┐
 [Nom]      [+dmns]     [P]           *malaki¹¹
                        [+rltv]       [V]
                                      [+sttv]
```

(13) Pre-Tagalog 'that big one' (lit. 'that one which is big')

```
  ┌─────────┬──────────────
 *i         *a ─────────┐
 [P]        [N]         *-ŋ ──────────┐
 [Nom]      [+dmns]     [P]           *malaki
                        [+rltv]       [V]
                                      [+sttv]
```
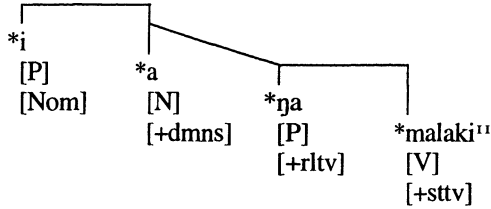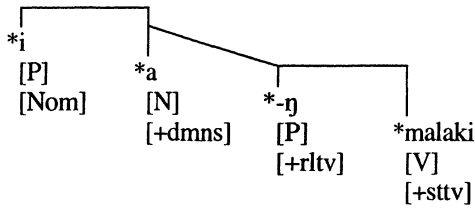
(14) Tagalog 'the big one'

```
  ┌────────────┐
  │         malaki
 aŋ        [N]
 [Det]
 [+dfnt]
```
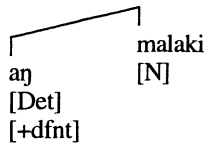
The stage represented in (13) is still present in Mamanwa in structures such as (15) and (16), in which a demonstrative noun occurs.

(15) Mamanwa (Miller and Miller 1976:33):¹²

Iniŋ	maŋa	bagʔoŋ	lodzoʔ	mahait	di.
this	plural	new	bolo	sharp	already

'These new bolo-knives are sharp.'

---

11. The analysis of *malaki* 'big' as a stative ([+sttv]) verb in Tagalog, rather than as a noun as in the Bontok examples above, is motivated by the presence in Tagalog of the stative verbal prefix *ma-* in the form. Following a Determiner, however, it is analyzed as a deverbal noun, as in (14).

12. Phonemic representation and analysis in this and the following examples are mine.

```
                                        mahait
                                        [V]        di
i-                                      [+sttv]    [Adv]
[P]          ni
[Nom]        [N]
             [+dmns]    -ŋ
                        [P]            bagʔo
                        [+rltv]  maŋa  [N]      -ŋ
                                 [Det] [prdc]   [P]       lodzoʔ
                                 [+plrl]        [+rltv]   [N]
                                                          [prdc]
```
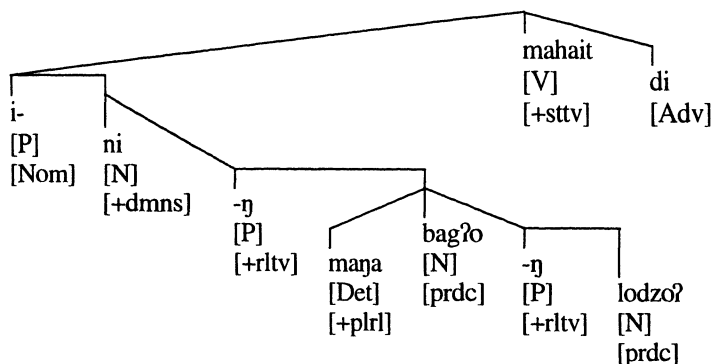
(16) Mamanwa (Miller and Miller 1976:33):
Iyaŋ isaŋ   ka   lodzoʔ kanao   di.
that one   Det   bolo   mine   already
'That one bolo is mine already.'

```
                                           kanao
                                           [N]       di
                                           [+prnn]   [Adv]
i                                          [prdc]
[P]       a
[Nom]     [N]      -ŋ
          [+dmns]  [P]     isa
                   [+rltv] [N]
                           [prdc]  -ŋ
                                   [P]      lodzoʔ
                                   [+rltv]  ka   [N]
                                            [Det] [prdc]
```
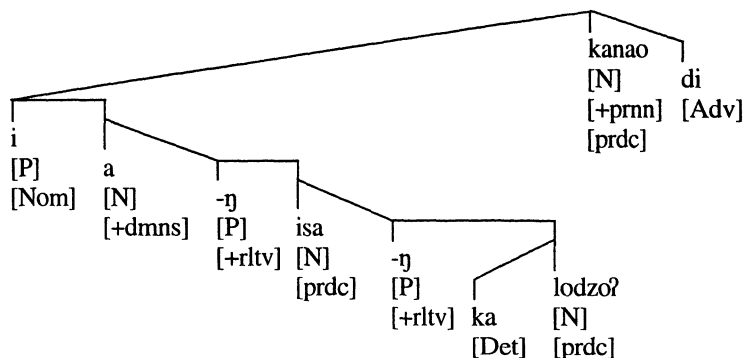
Nominative phrases without a demonstrative, however, show a pattern of reduction different from Tagalog. In Mamanwa, with loss of nominative case-marking and the development of the demonstrative as a determiner, the high front glide between the case marker and the demonstrative was retained as the initial consonant of the stem.

(17) Mamanwa (Miller and Miller 1976:33):
Inhasaʔ   di   ya kanaoŋ bagʔoŋ lodzoʔ.
sharpened already Det mine   new   bolo
'My new bolo is sharpened already.'

```
Inhasaʔ
[V]       di
          [Adv]         kanao
                   ya   [N]      -ŋ
                   [Det] [+prnn] [P]     bagʔo
                                 [+rltv] [N]    -ŋ
                                               [P]      lodzoʔ
                                               [+rltv]  [N]
```
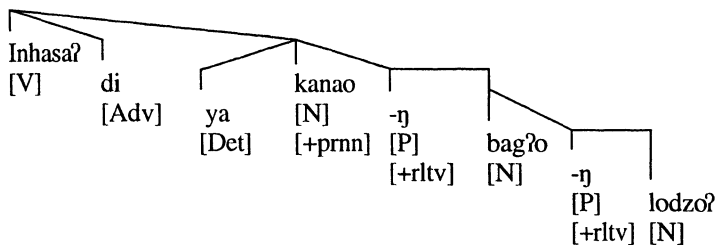
The relative preposition, which in Tagalog had become a postclitic and remained as a formant on the determiner, was lost in this position in Mamanwa. Thus the phonetic sequence [iyaŋ] (i + a + -ŋ) was reduced to ya, as in (17).

In Mansaka, both the earlier nominative case marker i and the velar nasal postvocalic variant of the relative preposition remain as frozen clitics on the definite determiner yaŋ, as in (18).

(18) Mansaka (Svelmoe and Svelmoe 1974:51):
yaŋ baray na madyaw
the house good
'the good house' (lit. 'the house which is good')
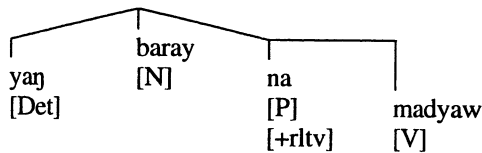
The above discussion has dealt only with some of the processes that resulted in the development of determiners in nominative common noun complements. Similar developments affected genitive case-marked complements, with what were originally genitive prepositions fusing in many languages with demonstratives to become genitive determiners, as in the case of the Tagalog genitive determiner naŋ.

6. Prenasalization in Proto-Minahasan

On the basis of data from the languages of Minahasa in Northeast Sulawesi, Sneddon (1978) reconstructs a series of nasal-obstruent clusters for Proto-Minahasan. The clusters could occur in both word-initial, as well as word-medial positions. Examples of forms with word-initial nasal-obstruent clusters are given in (19).

(19) Proto-Minahasan (Sneddon 1978:36, 77)

*mpelaʔ	'wound'	*nsuŋe	'horn'
*mbale	'house'	*ŋkaso	'rafters'
*ntali	'rope'	*ŋgio	'face'
*nduhi	'bone'		

Sneddon (1978:55) states that in initial position, the nasal was a morpheme separate from that of the following stop. He reconstructs it as *N-, a nasal that assimilates to the point of articulation of the following obstruent. The meaning of this morpheme is not clear, he says, because its functions in the various daughter languages differ. In Tombulu, Tonsea, and Tondano, the languages that constitute Proto–Northeast Minahasan, N- is a morpheme that signals the inanimate noun class. Sneddon notes (1978:36) that in Tombulu "when not preceded by a locative or instrumental preposition, inanimate nouns appear to always require a preceding particle, u indicating singularity and a indicating plurality," as in (20).

(20) Tombulu (Sneddon 1978:36)

u mbale   'a house'
a mbale   'houses'

However, in Tontemboan (which with Northeast Minahasan constitutes North Minahasan), the nasal is an indefinite noun marker, and follows one of a number of prepositions, as in (21).

(21) Tontemboan (Sneddon 1978:41)

N- + wale   →   mbale   'a house'
N- + rano   →   ndano   'water'

Although Sneddon states that the function of *N*- in Tonsawang (the language coordinate with North Minahasan) has not been fully determined, the examples cited in (22) show that it is neither a marker of indefinite nouns, as in Tontemboan, nor of inanimate nouns, as in the Northeast Minahasan languages. From these limited data, it appears to be a marker of definite, common nouns.

It seems probable, then, that Proto-Minahasan *N- is a continuation of the postvocalic variant of the nasal relative preposition, which has become a proclitic to the following noun rather than a postclitic to the preceding form, as happened in many Philippine languages. Syntactically, it is probably still a clitic determiner, rather than a prefix to the following noun, as in (22').

(22) Tonsawang (Sneddon 1978:57)

a N- + bale   →   a mbale   'to the house'
a N- + kedong   →   a kedong[13]   'to the child'

(22')

```
     ┌──────────────────────┐
     │              ┌───────┐
a                          bale
[P]          ┌──────      [N]
[Lcv]        m-           [-prsn]
             [Det]
```

If the Philippine languages that mark a distinction between personal ([+prsn]) and common ([-prsn]) categories are retentions of the PEF system, then Tonsawang is probably the most conservative of the Minahasan languages in this respect, and each of the other Minahasan languages has innovated. There are two other facts about the distribution of *N*- that are relevant. The first is that in all of the languages there are conditions under which the nasal is optionally or obligatorily absent. In Tonsea, for example, Sneddon (1978:22) notes that "[the nasal] is sometimes absent, for instance, when the noun is in an attributive position, e.g., *mbale,* sometimes *bale* 'house'." Typically, a predicate noun in an attributive structure is indefinite, and would not be preceded by a definite determiner (see, for example, [11] and [17], for instances of

---

13. Initial nasals have been lost before voiceless stops in Tonsawang. However, the fact that the form begins with a stop, rather than with a fricative is evidence that it derives from a nasalstop cluster (Sneddon 1978:57).

indefinite predicate nouns in Bontok and Mamanwa). Another position in which an indefinite noun would appear would be in the predicate position of nominal "descriptive" sentences. If Proto-Minahasan was a predicate-initial language, then unmarked nouns would probably have also appeared in this position. Sneddon (1978:52) states: "It is probable that in PNM [Proto-North-Minahasan] the prenasalised form of the word never began an utterance but that it always followed a particle. This is the case in TtB and probabl[y] also in PNE."

The second fact that is relevant is stated in the previous quote. It is that prenasalized forms probably all originally followed a "particle." One can assume from this that the particle being referred to must have been a definite determiner, having an origin, as in the other languages cited above, in an original demonstrative cum determiner.

7. Conclusion

I have attempted to show that prenasalization as it occurs in the Minahasan languages resulted from the stranding of what must have originally been a postclitic nasal that had its origins in the relative preposition, typically referred to as the ligature in Philippine languages. In these languages, the determiners of which they were probably originally a part have either been lost or have had their functions modified, stranding the nasal to become proclitic to the head noun of the construction. The function of the nasal, while originally agreeing with its head noun in terms of common versus personal, has now changed its meaning in various Minahasan languages to agree with the head noun in either animateness or definiteness.

Second, I have tried to show that there are grammatical conditions that require the presence of an indefinite noun, conditions under which the determiner would not normally have appeared. One such condition occurs in the initial predicate position of descriptive nominal clauses. An indefinite noun, without a determiner, would also have been required as the (predicate) head of nominal attributive structures (i.e., relative clauses).

In Pre-Oceanic, the same processes that we see operating in Minahasan languages must have also occurred. Nouns could occur, depending on their distribution, with a preceding nasal marking definiteness, or without the nasal, indicating indefiniteness. By Proto-Oceanic times, however, the marking of definiteness must have been accomplished by means other than the presence or absence of the nasal, probably by the presence or absence of the determiner *na/a alone, allowing the nasal to become interpreted phonologically as part of the immediately following obstruent. Once definiteness became marked by a feature other than prenasalization, what was once a grammatically conditioned alternation must have persisted for some time as free variation between the two classes of forms. Ultimately some of the prenasalized variants became dominant in some of the daughter languages of POc, while their nonprenasalized variants were lost. In other languages, it was the nonprenasalized variants that became dominant, while their prenasalized variants were lost, resulting in the situation that is described today as "cross-over."

REFERENCES

→ Blust, Robert. 1974. Proto-Austronesian syntax: The first step. *Oceanic Linguistics* 13:1–16.
—— → . 1991. The Greater Central Philippines hypothesis. *Oceanic Linguistics* 30:73–130.
→ Crowley, Terry. 1985. Common noun phrase marking in Proto-Oceanic. *Oceanic Linguistics* 24:135–193.
Dempwolff, Otto. 1969. *Vergleichende Lautlehre des Austronesischen Wortschatzes*, vol. 1, *Induktiver Aufbau einer indonesischen Ursprach*. Beihefte zur Zeitschrift für Eingeborenen-Sprachen 15. Berlin: Dietrich Reimer, 1934. Nendeln, Liechtenstein: Krauss Reprint.
Foley, William A. 1976. Comparative syntax in Austronesian. Ph.D. dissertation, University of California, Berkeley.
Grace, George W. 1990. "Consonant grade" in Oceanic languages. In *Pacific Island languages: Essays in honour of G. B. Milner*, ed. by J. H. C. S. Davidson, 41–50. London: School of Oriental and African Studies, University of London.
Heine, Bernd, Ulrike Claudi, and Friederike Hünnemeyer. 1991. *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.
Hopper, Paul J., and Elizabeth Closs Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
Lynch, John. 1975. Oral/nasal alternation and the realis/irrealis distinction in Oceanic languages. *Oceanic Linguistics* 14:87–99.
Miller, Jeanne, and Helen Miller. 1976. *Mamanwa grammar*. Language Data, Asian-Pacific Series, Number 3. Huntington Beach, California: Summer Institute of Linguistics.
Milner, G. B. 1965. Initial nasal clusters in Eastern and Western Austronesian. *Lingua* 14:416–430.
Reid, Lawrence A. 1978. Problems in the reconstruction of Proto-Philippine construction markers. In *Second International Conference on Austronesian Linguistics: Proceedings*, fasc. 1, *Western Austronesian*, ed. by S. A. Wurm and Lois Carrington, 33–66. Pacific Linguistics C-61. Canberra: Australian National University.
———. 1979. Evidence for Proto-Philippine nominative marking. *Philippine Journal of Linguistics* 10:1–20.
———. 1983. Ligatures and their implications for subgrouping. Paper presented to the April 14 meeting of the Austronesian Circle, Honolulu, Hawai'i.
Ross, Malcolm D. 1988. *Proto-Oceanic and the Austronesian languages of Western Melanesia*. Pacific Linguistics C–98. Canberra: Australian National University.
Sneddon, J. N. 1978. *Proto-Minahasan: Phonology, morphology, and wordlist*. Pacific Linguistics B-54. Canberra: Australian National University.
Starosta, Stanley. 1993. The case-marking system of Proto-Formosan. In *Pan-Asiatic Linguistics: Proceedings of the Third International Conference on Language and Linguistics*, vol. 3, 1207–1221. Bangkok, Thailand: Chulalungkorn University Printing House.
Svelmoe, Gordon, and Thelma Svelmoe. 1974. *Notes on Mansaka grammar*. Language Data, Asian-Pacific Series, Number 6. Huntington Beach, California: Summer Institute of Linguistics.
→ Wolff, John. 1967. History of the dialect of the Camotes Islands, Philippines, and the spread of Cebuano Bisayan. *Oceanic Linguistics* 6:63–79.
Zorc, R. David P. 1977. *The Bisayan dialects of the Philippines: Subgrouping and reconstruction*. Pacific Linguistics C-44. Canberra: Australian National University.

4: DEIXIS AND ANAPHORA
AND PRELINGUISTIC UNIVERSALS

Marybeth Clark
San Diego, California

1. Introduction[1]

The primacy of deictic concepts—the concern with the distinction between *here-and-now*/proximal concepts versus *there-and-then*/distal concepts—is reflected in language by explicit phonetic marking. Such marking is found in many languages throughout the world and includes the use of tonal distinctions, as in Hmong and Vietnamese, for example; vowel distinctions, as in Thai and many other unrelated languages; and consonant distinctions, as in English.

Thai and Lao have a three-way deictic marking system, dividing nonproximal into medial and distal. Many other languages also have such a three-way system, for example, Japanese. Fillmore (1982:53) states that "throughout the whole system of demonstrative words [in Japanese], distance categories are consistently signalled" by specific prefixed forms for proximal, medial, and distal. He also lists (48) Latin, Greek, and Spanish as having three-way systems. Diller (1989:5) states that "relative 'contrastive proximity' rather than absolute distance is important; thus one could hardly specify an exact objective measure for how the terms are applied." Nevertheless, the distinction between relatively proximal reference—"here, with me"—and relatively distal reference—"there, with others" and "out of sight"—is an important one, as is discussed further in section 3. The distinction is realized in language by the tendency to designate *there* with more highly marked phonetic signals, suggesting that deictic terms would in general be expected to carry the feature [± here], as the more predictable characteristic, rather than [± there].[2]

Such widespread recognition of the *here–there* distinction strongly suggests the existence of yet another linguistic universal, and one obviously based on fundamental concepts. The claim made here is that underlying the linguistic universality of such contrastive marking is the prelinguistic distinction between "present reference" behavior ("unmarked" behavior, acting fairly predictably in response to the immediately perceived environment: proximal reference) and displacement/projected behavior (more complex "marked" behavior, acting less predictably in reference to another place or time: distal reference). Such a prelinguistic basis for deictic paradigms in lan-

---

2. I am taking here the approach that items carry features according to unmarked (expected) characteristics rather than the current—and valid—consideration that features should be named by distinguishing marked characteristics.

guage is just one of a number of natural behavioral components that are reflected in patterns of language structure. (See Clark 1992:96 for other possibilities.)

The term "prelinguistic" here does not mean "proto-language" nor does it refer to any vocal communication that may be a precursor to human language. This essay is not concerned with the evolution of human language, but rather with the parallel in the cognitive processes that underlie both nonlanguage behaviors and deictic patterns in human language. Thus, "prelinguistic" refers to behaviors presumed to preexist language.

In section 2, I give a summary of the use in White Hmong of tonally-defined deictic classes and show in detail the Vietnamese tonally-defined deictic and anaphoric sets. In section 3, I mention other minimal-distinction means, such as vowels, used by other languages in dealing with *here* versus *there*. Finally, in section 4, I speculate on the relevance of these linguistic sets to "prelinguistic" behavioral patterns.

2. Deictic Contrasts Marked by Tone

Discussed here are the tonally-defined deictic paradigms of White Hmong and Vietnamese, in which a regular difference in tone marks the difference between *here–now* and *there–then* anaphoric designations. Willem de Reuse (pers. comm.) states that there is a parallel tone distinction in Western Apache in Arizona, that is, a distinction between *ai* 'that (visible)' with low tone and *áí* 'that (not visible)' in high tone, low tone being the unmarked tone. Visible may be perceived as being generally proximal while invisible is clearly distal.

2.1 White Hmong Tonally-defined Deixis

Ratliff (1992), in her discussion of tonal morphology in Hmong, gives parallel sets of tonally-defined deictic form classes in White Hmong. They involve a set of nonanaphoric relative location words that she calls denominal prepositions, and a set of corresponding anaphoric referents that she calls demonstrative nouns (Ratliff 1992:109, 112). She shows that the demonstrative nouns are derived from the denominal prepositions. Some examples from the two matched sets are given in table 2. Table 1 gives the written symbols used in this paper for White Hmong tones. Note in table 2 that all the basic location words are in the *-m* tone, while the derived demonstratives are in the *-d* tone, itself not a primary tone but one derived from the *-m* tone and one that is highly marked.[3]

Table 1. White Hmong Tones*

-b	high level	Ø	mid level	-m	low with final glottal
-j	high falling	-g	low falling breathy	-d	low rising (derived from
-v	mid rising	-s	low		-m tone)

* The Hmong Romanized Popular Alphabet uses consonant symbols word-finally for tones and VV for Vŋ, as ŋ is the only final consonant in the language.

---

3. For a full discussion of the probable tone development of the derived set of *-d*-tone demonstratives, see Ratliff 1992:112–120.

Table 2. Tonally-defined Deictic Sets in White Hmong*

RELATIVE LOCATION WORDS		DERIVED ANAPHORICS	
TOM 'over there'	mus tom khw 'go to the market'	TOD 'over there'	mus lawm tod 'go [far] over there'
NTAWM 'nearby'	nyob ntawm kev 'be on the trail'	NTAWD 'nearby there'	nyob ntawd 'be there (not far)'
PEM 'up'	nyob pem roob 'be up on the mountain'	PED 'up there'	nyob ped 'be up there'
SAUM 'above'	nyob saum ntoo 'be up in the tree'	SAUD 'there above'	nyob saud 'be there above'
NRAUM 'outside'	nyob nraum zoov 'be outside (in the forest)'	NRAUD 'there outside'	nyob sab nraud 'be outside, on the other side'
THAUM 'time when'	puag thaum kuv tseem yau 'when I was still young'	THAUD 'that time when'	puag txheej thaud 'in the old days'

---

* The information in this chart is found in Ratliff 1992:105–106, 113–114.

2.2 Vietnamese Tonally-defined Deixis and Anaphora

In Vietnamese, there is a series of sets distinguished by tonal alternations. One paradigm is comprised of *here-and-now* versus *there-and-then* demonstrative nouns. Another paradigm, primarily occurring in south and south-central Vietnam, uses a tone change to one particular tone for anaphoric reference in relative location words, kin terms as person reference, and a few other terms.

I will not go into the Vietnamese tone system other than to say that, in line with tone development in the area, Vietnamese at an earlier time had a set of three contour tones that later split into two registers when voiced initial consonants were lost. The tones in table 3 are classed according to the designations in Thompson (1984–85:141) and reflect the historical split of "second" tones from "first" tones. Words with final stops occur only with the short tones. (The glottals occur in northern dialects. The two-register split is complete across the board only in the northern dialects: the long tone has not separated in southern dialects; the central dialects, not shown here, reflect an even earlier stage, having only three or four tones.)

Table 3. Vietnamese Tone System

REGISTER	EVEN	SHORT	LONG
FIRST	a mid high, level	$á$ high rising	$\grave{a}$ mid falling/rising (South)
SECOND	$à$ low	$ạ$ low (with glottal)	$\tilde{a}$ mid rising (with glottal)

48

If we consider the more predictable tones—those used in borrowed words, for example—to be the less marked tones, then the even tones are less marked, with the high level tone being the least marked of all the tones. The high rising tone is unmarked for words with final stops but somewhat marked for words without final stops. In Vietnamese deixis, the *here–now*, new information terms have unmarked tones, that is, one of the even tones, while the *there–then* terms tend to have more marked tones, that is, one of the short tones. Third-person-designated kin and other terms use the long first-register tone, also more marked than the even tones.

2.2.1 Vietnamese deictic words. The paradigm of Vietnamese deictic terms is given in table 4, following Thompson (1984–85:142) in general. It shows the tone alternation between the *here-and-now* set in the first column and the *there-and-then* set in the second column. There is a third set (in the third column) that parallels the first two sets and is distinguished from the first set not by tone but by vowel. These words are indefinites, often used as question words. There are some deictic words in Vietnamese that appear to fit the vowel and tone patterns of these sets but are not paradigmatically complete. They are given here for speculation. The designation [N] refers to an obligatory noun occurring before or after the deictic word.

Table 4. Vietnamese Deictic Forms

PROXIMAL, NEW	DISTAL, ANAPHORIC	INDEFINITE
đây*	đấy/đó	đâu
'here'	'there'	'where(ver)'
bây	bấy	bao
'to this extent [N]'	'to that/such extent [N]'	'to what(ever) extent [N]'
này/nầy/nay	nấy/nọ/nãy/nớ	nào
'([N]) this'	'[N] that, (an)other'	'[N] which(ever)'
	ấy	ai
	'([N]) that (person/thing)'	'who(ever)'
vầy	vậy	(sao)†
'this way, thus'	'that way, (as) so'	'how(ever)'

---

* Diacritic marks with vowels in the following combinations represent distinctive Vietnamese vowels and do not refer to tones: *â, ă, ê, ô, ơ, ư.*

† Included by Thompson.

Nguyen (1992:128) distinguishes three positions for the place and reference deictic forms instead of the two given in table 4, as shown in table 5, where Nguyen's notation +Nom(inal) means an unbound morpheme, –Nom(inal) means that the word depends on the presence of another noun, and ±Nom(inal) indicates that the word occurs in either context.

Table 5. Deictic Positions According to Nguyen 1992

	PROXIMAL	MEDIAL	DISTAL	INDEFINITE
+NOM	đây 'place-this'	đấy 'place-that'		đâu 'place-what'
±NOM		đó '(place-)that'	kia '(place-)that/over there'	
−NOM	nầy 'this'	nấy 'that'	nọ 'that over there'	nào 'what'

Nguyen (1992:128) also suggests that the distinction [d-] versus [n-] correlates with *place* + *reference* versus *reference*, and *free* versus *bound*, but gives no other corresponding evidence. In fact, *nó* 'he, she, it' (pp.129, 132) is *free* and *person* + *reference*, separating it from other [n-] words.

My concern here is with tone alternation in *here–there* deictic contrasts and, although *kia* itself has interesting tonal alternations (see Nguyen's discussion, 1992:133–134), these alternations do not parallel the particular contrasts of interest in this study.

Naturally, some of these terms are in more common usage than others, while some are particularly adaptable to idiomatic usage. The combination of indefinites with *there* terms is particularly productive idiomatically, as shown in (1–4) and in the proverbs in (5–7). Other sentence examples follow.

(1)   Có **bao** nhiêu, xài **bấy** nhiêu.
         have however much spend that.extent much
         'Whatever you have you spend.' (Thompson 1984–85:147)

(2)   **Bao**.nhiêu người (thì) **bấy**.nhiêu ý.kiến.
         however.many person then that.many idea
         'However many people there are, there are that many different ideas.'

(3)   Nhà nấu món **nào** anh ăn món **nấy**!
         house cook dish which brother eat dish that
         'Whatever they cook, you eat it! (It's only polite.)' (Nguyen B. T. 1997:34)

(4)   Họ cho tôi công.việc **nào**, tôi nhận công.việc **nấy**.
         they give subordinate work which sub. receive work that
         'Whatever work they give me, I'll do it.'(Nguyen B. T. 1997:39)

(5)   Cha **nào** con **nấy**.
         father which child that
         'Like father, like son.'

(6)   Người **nào** giữ phận **nấy**.
         person which keep condition that
         'Everyone should keep to his proper place (in society) / attend to his proper duties.'

(7) Cha   mẹ   đặt **đâu** con ngồi **đấy.**
     father mother place where child sit   there
     'Whomever the parents select, the daughter marries.'

(8) **Đây** là nhà của   tôi, **đâu** là nhà của   ông?
     here be house belonging sub. where be house belonging grandfather
     'Here's my house, where's yours?' (Thompson 1984–85:254)

(9) Cô **ấy** rất thích làng, nhưng không muốn ở   **đấy** luôn.
     aunt that very like village but   not   want be.at there always
     'She likes the village very much, but she doesn't want to live there all
     the time.' (adapted from Thompson 1984–85:254)

(10) Trước khi   người Việt   đến **đây,**
     before moment person Vietnamese arrive here

     vùng cửa   sông **này,** . . .
     area door river this
     'Before the Viets came here to this estuary area, . . .' (Le 1986:13)

(11) Từ **đấy** cho-đến **nay**   Hội-An trở-thành . . .
     from there until   this(.time) Hoi-An become
     'From that time until now, Hoi An has become . . .' (Le 1986:13)

(12) (**Bây**   giờ chưa   biết.)
     this.extent hour not.yet know

     Làm xong việc, **bấy**   giờ sẽ   hay.
     do   finish work   that.extent hour will know
     '(At this time I don't know yet.) When I finish, then I'll know.'

(13) Sự **này** xảy   ra như **vầy:** anh   Bính
     affair this happen out like thus   older.brother Binh

     mệt lắm thì lái   xe   không coi.chừng.
     tired very then drive vehicle not   be.careful
     'It happened this way: Binh was very tired and drove carelessly.'
     [Southern] (Thompson 1984–85:148)

(14) Tôi   sẽ nói lại với cậu (**ấy**) như **vậy.**
     subordinate will speak back with uncle that as   so
     'I'll tell him that.' / 'I'll pass that message on to him.'

Nguyen (1992:133) points out that place can be used metaphorically to designate
person, as shown in (15), a use that emphasizes a distinction between First Person–
here and Second Person–there.

(15) **Đây**   đi chợ,   **đấy**   có đi không?
     this.place go market that.place have go not
     'I'm going to the market. What about you?'

2.2.2 Other tonally-distinguished anaphorics in Vietnamese. In south and south-central Vietnam, and perhaps of late, somewhat in the north, the rather marked long first-register tone (mid rising in the south) is used with kin terms for anaphoric reference. Kin terms are used in Vietnamese for first- and second-person reference, both within a family context and in extended use outside the family to maintain social coherence through reaffirmation of status and close/distant relationship. The same terms are used for third-person reference, usually with a modifying word meaning 'that', as in sentences (9) and (14) above, and in (16a) and (17a) below. It is possible that the use of the long first-register tone for third person developed from a contraction of the kin-term and the following high-rising tone of the determiner. The same tone distinction is used with some deictic terms, such as relative location, time, and extent nouns, in much the same way as in White Hmong. According to Thompson (1984–85:149), this anaphoric use of the mid-rising tone does not occur with forms having the high-rising tone nor with those ending in stops. Table 6 gives some of his list of kinship terms, one nonkin person term, and some deictic nouns that can occur with this tone alternation. Note the parallel with the White Hmong deictic set.

The sentences in (16) and (17) give the base noun + determiner in (a), and in (b) the tone alternation of the base word for anaphoric reference.

(16) a. Chị đã thấy **anh**   **ấy** ở đây lần.nào chưa?
        sister past see   brother that at here any.time not.yet
        'Have you ever seen him here?'

    b. Chị đã thấy **ảnh**   ở đây lần.nào chưa?
        sister past see   that.brother at here any.time not.yet

(17) a. Hôm qua tôi gặp cô   Thu ở bên Biên-Hòa.
        day   past sub. meet aunt Thu at side Bien-Hoa
        'I met Miss Thu in Bien Hoa yesterday.'

        Cô ấy nói sẽ ở   lại   **bên đó**
        aunt that say will be.at back side that

        với ông   (của)   **cô ấy.**
        with grandfather belonging aunt that
        'She said she would stay there with her grandfather.'

    b. **Cổ**   nói sẽ ở   lại   **bển**
        that.aunt say will be.at back there

        với ông   **cổ.**
        with grandfather that.aunt

(18) Ở **đằng**   **này** có tiệm ăn không?
      at   direction this have shop eat not
      'Is there a restaurant over this way?'

      Ở **đằng**   mới có.
      at   that.direction only have
      'No, only over there (place known).' (Thompson 1984-85:150)

(19) Tôi có **chừng** ấy, ông   cũng có   **chửng.**
      sub. have extent that grandfather also have that.extent
      'I have so much, you have just as much.' (Thompson 1984–85:150)

Table 6. Tonally-defined Third-person Reference

ông	ổng
'grandfather, gentleman'	'that grandfather, he'
bà	bả
'grandmother, lady'	'that grandmother, she'
cụ	củ
'great-grandparent, elder venerated person'	'that great-grandparent, . . .'
mẹ	mẻ
'mother'	'my mother, she'
anh	ảnh
'older brother, male equal'	'my older brother, that young man, he'
chị	chỉ
'older sister, female equal'	'my older sister, that female friend, she'
cậu	cẩu
'mother's brother'	'that uncle/young man, he'
cô	cỗ
'father's sister; young lady'	'that aunt/lady, she'
thằng (3rd person title)	thẳng
'inferior person (male)'	'that fellow'
bên	bển
'side, area'	'that side, there'
đằng	đẳng
'location (relative to)'	'there'
trong	trỏng
'inside, interior'	'inside (it)'
ngoài	ngoải
'outside, exterior'	'out there, outside (of it)'
trên	trển
'top, upper surface'	'on top (of it)'
hôm	hổm
'day'	'that day, then'
chừng	chửng
'certain amount, extent'	'to that extent, that much'

(20) Hôm qua tôi làm một cái   chuồng cho con   chim.
      day past sub. do   one thing cage   for animate bird

Bữa   nay nở   ở   trồng.
day   this that(.one) be.at inside(.of.it)
'Yesterday I made a cage for the bird. Today he's in it.' (Thompson
1984–85:150)

These person and place anaphoric words are patently nonproximal *there–and–then*
words because they refer exocentrically, that is, refer to other persons and places, to
content outside the immediate context of utterance or the environment of speaker and
addressee. Genevieve Escure (pers. comm.) makes the observation that anaphora
would be medial reference (nonproximal yet not quite distal), presumably because
anaphora refer to items that are easily recoverable, that is, not distant.

3. Proximal–Distal Marking by Other Phonetic Means

It is not only tonal distinctions that are used for proximal–distal contrast in deixis. It
is clear that phonetically defined space-and-time proximal–distal pairing is wide-
spread in languages. English uses initial consonant contrasts, but particularly preva-
lent is a front-to-back/lower vowel distinction.

3.1 Use of Vowel Contrasts

Many languages pair proximal and distal deixis by minimal vowel distinction. In his
discussion of Southern Thai deixis, Diller (1989:5) points out that in the sets of terms
for *here–this / there–that / (that) over there* (my simplification), the front/central/back
classification of vowel position codes the near-to-far distance continuum, dividing
nonproximal into medial and distal. A similar near-to-far continuum for vowels exists
in Lao, according to Carol Compton (pers. comm.), with 'be here / near / there / way
over there': *yuu phii / yuu hon / yuu phuun / yuu phuuuun*, with the far-far referent hav-
ing a long drawn-out vowel with particularly high tone.

Tanz (1971) claims that there is a universal tendency for a high front vowel in the
word for *here* and a vowel that is lower and/or farther back for *there*. She lists (Tanz
1971:268) languages of six widely-separated language families, that is, Semitic, Dravid-
ian, Malayan, Uto-Aztecan, Australian, and Finno-Ugric, which mark such a minimal
distinction between *here* and *there*. She further calls attention (274–276) to another inter-
esting proximal-distal vowel pairing, and that is in English irregular verbs. She lists about
60 verbs that have a high or mid-high front vowel in present tense and a further back and
often lower vowel in past tense. On the other hand, there are no past-tense verbs with a
high front vowel for which there is not also a high front vowel in the present-tense form.
A few examples of this present–past vowel distinction are given in table 7.

Table 7. English Verb Forms
in Present (Proximal) and Past (Distal) Tenses

PRESENT	PAST	PRESENT	PAST
sit	sat	bring	brought
ring	rang	teach	taught
sting	stung	see	saw

3.2 Initial Consonant Contrast in English Deixis

English has a set of deictic terms in which *here* and *there* terms are distinguished by initial consonants, as shown in table 8. These sets, like Vietnamese, also include a correlating set of indefinites. Some of the words in these sets are rather archaic and some are paradigmatically incomplete, but they show the pattern of intent.

Table 8. English Deictic Forms

PROXIMAL	DISTAL	INDEFINITE
here	there	where
'at this place'	'at that place'	'at which(ever) place'
hither	thither	whither
'toward this place'	'toward that place'	'to what(ever) place'
hence	thence	whence
'from this place/time'	'from that place/time'	'from what place/source'
(this, at?)	that	what
'this one'	'that one'	'what(ever) one'
	then	when
('at this time')	'at that time'	'at which(ever) time'

Fillmore (1982:47) claims that the *hence–thence* set is obsolete. These terms may indeed be archaic and perhaps seldom heard outside the written language, but since they can be used with the expectation of being fully understood, they are obviously not obsolete. Note the following use of *whence* in a recent American newspaper item (Jocasta Blashfield in "Dining Out: the French Gourmet Too," *La Jolla Light*, 3 June 1993).

(21) . . . the ever-bustling bakery whence many a successful party has been catered.

4. Prelinguistic Universals

Such extensive use of minimal phonetic distinctions to pair and contrast proximal and distal designations highlights the linguistic universality of such pairing, pairing that emphasizes the vital distinction between *here* and *there*. It further suggests that this linguistic universality stems from a more basic prelinguistic universality.

In using the term "prelinguistic" I do not refer to proto-language or any form of vocalization that may or may not be a precursor to human language. In short, this essay is not concerned with the evolutionary development of human language, what is and what is not language, nor whether nonhuman animals have language.

The concepts of deixis, of *here versus there* reference, are based on cognitive processes having to do with differentiation and interrelation. Such cognitive processes are the very foundation of not only language but of all nonlanguage behavior in which humans and other animals relate to their environment and each other (Della Volpe 1995:80ff; Lenneberg 1967:333, 336, 372ff; Lieberman 1984:1f, 331).

Thus, the term "prelinguistic universals" refers to those nonlanguage behaviors that reflect *here versus there* concepts and that seem to parallel to some degree the deictic paradigms in human language.

Cognitive function is a more basic and primary process than language (Lenneberg 1967:374), and the claim here is that the nonlanguage *here–there* distinction realized in responsive behavior forms the basis for the *here versus there* realization in language. It is not unreasonable to make this claim. As Lieberman (1984:1–2) states, "human linguistic ability is based on rather general neural mechanisms that structure the cognitive behavior of human beings as well as that of other animals. The evolution of human language is therefore analogous to the evolution of behavioral patterns that are 'unique' to other species." Lenneberg (1967: 333), in stating that concepts are modes of ordering or dealing with sensory data, goes on to say that "although this process is not peculiar to man (because it essentially results from the mode of operation of a mechanism that can only respond in limited ways to a wide variety of inputs), man has developed the behavioral peculiarity of attaching words to certain types of concept formation."

4.1 Present Reference and Projected Reference

In attempting to show parallels between deixis in language and deictic notions in the nonlanguage behavior of humans and other animals, it is easy enough to demonstrate the distinction between conceptual *here* and *there* in nonlanguage behaviors. Those behaviors that can be said to underly the proximal–distal paradigms are, on the one hand, the behaviors of response with present reference to an immediately observable environment and, on the other hand, displacement behaviors, that is, behaviors involving reference to a distant, past, or future environment—what might be called *projected reference.*

The distinction between proximal and distal is a crucial one. "Proximal" behavior can be considered to be unmarked behavior: a natural response to the immediately perceived environment, a response that is as a rule predictable or one of a number of predictable responses. It is significantly less complex than "distal" behavior in that it makes little reference other than memory outside the immediate environment. In linguistic terms, the most exact proximal environment involves First Person (usually in exchange with Second Person) with respect to *this place at this time*, either actually or metaphorically.

"Distal" behavior, on the other hand, is much more complex, much more "marked" in terms of predictability. Naturally, there is a continuum of complexity from very simple "proximal" behavior through medial distal to very complex "far distal" behavior, and, naturally, there is overlap with questionable areas of proximity, as implied in the quote from Diller 1989 in the Introduction. However, "distal" behavior is not just more complex than "proximal" behavior; there is another dimension of complexity that distinguishes distal from proximal. Distal conceptualization requires projected reference, conceptualization in absentia, so to speak. It requires what Frans de Waal calls "strategic intelligence," the fact that "a rational choice is based on an estimate of the consequences" (de Waal 1982:187). That is, the distinction is the process of separating another time, in this case the future, or point of reference from the present but maintaining the link between. The complexity involved in distal reference makes of it a more marked reference and, here again, if we are to assign binary features to nonlanguage behavior, time–space in nonlanguage behavior can be considered to be [± here] rather than [± there].

As for pairing, it is not so easy, in fact very problematical, to show pairing phenomena between particular sets of proximal versus distal nonlanguage behaviors. That is, the phonetic pairing in Vietnamese and English, for example, between *here* and *there* words is not so clearly matched in nonlanguage behavior, and, in fact, may not exist. In the following examples, I have paired what seem to be related behaviors, but perhaps such pairing is really not relevant. Language, with its more precise rules and paradigmatic nature, its explicit formation of words, and its tendency to play with concepts, can make very specific recognition of the contrasts in the underlying concepts of deixis. Nonlanguage behavior in general follows more implicit patterns.

4.2 Human Nonlanguage Behavior in Deixis

Following are some examples of proximal–distal contrasts in human behavior.

(a) A person picking up a spoon, fork, or chopsticks with the goal of feeding oneself from an available food supply, versus a person going shopping for groceries with the purpose of preparing a meal (medial distal?), or accumulating money in order to be able to purchase food—certainly more distal.

(b) A person walking into a room, seeing a book and picking it up, and sitting down to read, versus a person going to the library with the intention of reading (medial distal), or going to the library to check out a book to take away to read or sending away by mail for a book that the person wants to read.

(c) A person picking up a ringing telephone and answering, versus a person looking up a number in the telephone directory and dialing the number with the expectation of talking with someone on the other end of the line (medial distal), or a person compiling a list of people and their telephone numbers with the intent of calling each person to persuade them to participate in a business venture, and drawing up a convincing argument.

57

4.3 Chimp and Wolf Behavior

Examples of proximal versus distal behavior can also be found among other animal groups. To avoid as much as possible a confusion with instinctive behavior or human-influenced behavior, I will take my examples of nonhuman animal behavior from wild or captive undomesticated animals who, like humans, live in social groups, and most of the examples will be of social interaction within such groups. Specifically, these examples will be of chimpanzees, drawing primarily from de Waal 1982, and North American wolves, drawing primarily from Lopez 1978.[4]

Following are examples of behavior in chimpanzees that show the distinction between here/ present/ proximal and there/ projected/ distal.

(a) A chimp walking over to a bunch of leaves (purposely dropped by another chimp in a tree) in order to eat them, versus one chimp coaxing another to climb a tree to break off and throw down branches for the first chimp to eat leaves.

(b) A displaying male picking up a stone to make a more impressive display, versus male A spending time with each of four influential females (grooming, playing with the children) just prior to challenging male B. Is it "an attempt to get them to remain impartial during the subsequent clash?" asks de Waal (1982:99).

(c) Chimp A sitting down beside and grooming chimp B to reassure or comfort B after B has been involved in a conflict, versus a female making a fuss over a male to recruit him to attack, without her help, another female with whom she has a complaint (medial distal), or a female, wishing to assist the two quarreling males A and B to make peace for the general harmony of the group, grooming male A and then bit-by-bit persuading A to go with her over to male B to groom B.

(d) A third-ranking male ingratiating himself with the current alpha male for the immediate benefits of the association, versus a third-ranking male who, instead of supporting the alpha chimp who is being threatened by the beta chimp, correctly predicts

---

4. Many observations have been made of chimpanzee behavior. Zoologist Franz de Waal is among the foremost observers of the behavior of captive chimpanzees. His fascinating book *Chimpanzee Politics*, which I have used here, is the result of a scientific team study over a period of years of a chimpanzee colony at Arnhem Zoo in Holland in what was at the time a remarkably natural environment. For the past 20 plus years he has been at the Yerkes Primate Research Center in Atlanta, Georgia, where he has concerned himself, among other things, with the evolution of behavior in primates and has dealt with this theme in *Peacemaking among primates*, Harvard University Press, 1989, and *Good natured: The origins of right and wrong in humans and other animals*, Harvard University Press, 1996.

I have chosen to take examples of wolf behavior from Barry Lopez's *Of Wolves and Men* because it is so comprehensive, although in the twenty years since its publication much more data on wolves have been collected. Lopez refers to the work of many other naturalists; gives many examples of wolf behavior from his own and others' observations, reports that are consistent with more recent observations; includes attitudes and beliefs of Native Americans, especially but not exclusively those of northern regions; and covers everything from a detailed description of howl harmonic frequencies to the extensive mythology about wolves. Besides, it is beautifully written! Reference has also been made to naturalists R. D. Lawrence, especially *In praise of wolves*, Henry Holt and Company (1986) and Ballantine Books (1988); Farley Mowat; and L. David Mech.

that he should throw his weight in with the beta chimp, who eventually wins out as alpha male. It should be understood that the beta chimp did not come out on top solely because of the support of the third ranker; to be alpha male it is necessary to have the support of the influential females and finally the whole group. Therefore, the prediction of the third ranker involved a shrewd guess as to the dynamics of the whole group.

Because chimp behavior and thinking is closer to our own than is wolf behavior, our guesses as to what is going on are presumably more often on target. Furthermore, opportunities for researchers to observe chimpanzee social behavior have been extensive. Although the last twenty years have seen growing interest in and observations of wolves, it still holds that guesses as to what motivates wolves are often simply guesses. From observed consistencies in their behavior, it is obvious that far more than we can see is going on and that wolves have perceptions beyond our own abilities, and ways of relating these perceptions to govern their behavior.

The wolf behaviors given here as examples of proximal–distal contrasts are well-observed behaviors and reflect the wolf's ability to project.

(e) A wolf trotting along a woods trail and suddenly stopping upon hearing a small underground movement, then pouncing and digging for the unfortunate mouse who made the movement, versus a wolf attempting to make a secret cache of a kill. Lopez (1978:63) relates an instance when the naturalist Adolph Murie once trailed a wolf who had just killed a Dall lamb and carried off a part of it. The wolf went off across the snow, at one point backtracking 15 yards, then jumping off the trail eight feet to the side and wandering in several loops. He then crossed some wet tundra, stepping deliberately in small puddles; coming to a creek, he waded in the creek 15 yards downstream before coming out on the same side, then 300 yards further down crossed the stream and finally buried his cache beside a tree in some woods. Unfortunately, this elaborate attempt to cover his tracks was unsuccessful: Not only was Murie able to track the wolf, but a fox came onto the trail ahead of Murie and, by the time Murie reached the cache, the fox had raided it! This attempt on the part of the wolf would be a medial-distal projection: The wolf correctly predicted that he might be trailed by a potential raider and so sought to confuse the trail and thwart the raider; the fox, however, was just as smart as well as persistent.

More distal projection is involved in a wolf pack controlling the game within its territory. According to Lopez (1978:55), wolves have been known to "practice a kind of fallow-field farming by not killing deer in certain parts of their territories for four or five years, letting the prey population recover there."

(f) Pack members joining in a group howl, each successive joiner selecting a different harmonic, versus two wolves a mile apart exchanging howls and then trotting off obliquely to meet at a tangential third point (medial distal), or a pack with their noses together communicating (possibly planning or simply bonding?) before setting out to ambush a herd of caribou, the execution of which involves some of the pack chasing or stampeding the caribou in the direction of a crossing where other members of the pack are waiting to intercept the caribou.

5. Conclusion

It is apparent from the paradigmatic tonal alternations in Vietnamese and Hmong, the vowel alternations in Thai and other languages, the consonant alternations in English, and other linguistic devices that there is a universal need to both pair and contrast the deictic notions of *here–now* and *there–then*. It is customary in linguistics to call these two references proximal and distal, respectively. Some languages further divide nonproximal into medial and distal.

Three claims are made here, the first being that this linguistic universal tendency to pair and contrast proximal and distal deixis derives from a very basic prelinguistic universal that is reflected in animal behavior, both human and nonhuman. The prelinguistic behaviors that underlie the proximal–distal paradigms are (a) the present reference behaviors of response to an immediately observed environment and (b) displacement/ projected reference behaviors, that is, behaviors involving reference to a distant, past, or future environment.

The second claim is that the distinction between proximal and nonproximal is a crucial one. It is the distinction between reference to the visible, immediately understandable, and accessible environment, and reference to that which is distant in space and/or time and requires conceptual projection. In terms of the future, distal reference requires "strategic intelligence" (de Waal 1982:187), making a rational choice based on an estimate of the consequences, with the extent of strategy relative to the extent of distance. A rough parallel might be made with the near-to-far vowel continuum in Lao, with the extent of the vowel reach to the vocal back being relative to the extent of perceptual distance.

The third claim, dependent on the second claim, is that distal reference is more highly marked than proximal reference. In language, *there* notions consistently have more highly marked phonetic signals. The greater conceptual complexity of distal or projected reference in nonlanguage behavior as compared to proximal or present reference at once claims it to be a more highly marked behavior. If we label the features of items and concepts according to predictability, then deictic notions would carry the feature [± here] rather than [± there] because *here* is the unmarked characteristic involving greater predictability.

REFERENCES

Clark, Marybeth. 1992. Markedness in prelinguistic universals. In *The language game: Papers in memory of Donald C. Laycock*, ed. by Tom Dutton, Malcolm Ross, and Darrell Tryon, 95–109. Pacific Linguistics C-110. Canberra: Australian National University.

Della Volpe, Angela. 1995. Genetic classification and the historical method. In *The biology of language*, ed. by Stanislaw Puppel, 73–98. Amsterdam/Philadelphia: John Benjamins Publishing Co.

Diller, Anthony. 1989. Southern Thai deixis. *Papers in South-East Asian linguistics*, No. 11: *South-East Asian syntax*, ed. by David Bradley, 1–14. Pacific Linguistics A-77. Canberra: Australian National University.

Fillmore, Charles J. 1982. Towards a descriptive framework for spatial deixis. In *Speech, place, and action, Studies in deixis and related topics*, ed. by Robert J. Jarvella and Wolfgang Klein, 31–59. Chichester: John Wiley and Sons Ltd.

Le Van Hao. 1986. *Hoi An: Su hoi sinh ...* (Hoi An: The revival of an old town). Danang, Vietnam: DaNang Publishing House.

Lenneberg, Eric H. 1967. *Biological foundations of language*. New York: John Wiley & Sons, Inc.

Lieberman, Philip. 1984. *The biology and evolution of language*. Cambridge, Mass.: Harvard University Press.

Lopez, Barry Holstun. 1978. *Of wolves and men*. New York: Charles Scribner's Sons.

Nguyen Bich Thuan. 1997. *Contemporary Vietnamese: An intermediate text*. Southeast Asian Language Text Series, ed. by Edwin Zehner. DeKalb: Northern Illinois University Center for Southeast Asian Studies.

Nguyen Phu Phong. 1992. Vietnamese demonstratives revisited. *Mon-Khmer Studies* 20, *for Laurence C. Thompson on his sixty-fifth birthday*, 127–136.

Ratliff, Martha. 1992. *Meaningful tone: A study of tonal morphology in compounds, form classes, and expressive phrases in White Hmong*. Special Report No. 27. Center for Southeast Asian Studies Monograph Series on Southeast Asia. DeKalb: Northern Illinois University.

Tanz, Christine. 1971. Sound symbolism in words relating to proximity and distance. *Language and Speech* 14.3:266–276.

Thompson, Laurence C. 1984–85. *A Vietnamese reference grammar, Mon-Khmer Studies* 13–14, ed. by Stephen O'Harrow. Honolulu: University of Hawai'i Press. Reprint of 1965, *A Vietnamese grammar*, Seattle: University of Washington Press.

de Waal, Frans. 1982. *Chimpanzee politics: Power and sex among apes*. Counterpoint, London: Unwin Paperbacks.

5: THE EMERGING PARTICLE *POKO* IN KOREAN: A GRAMMATICALIZATION

In-Seok Yang
Hankuk University of Foreign Studies

1. Introduction[1]

Korean has two types of particles, namely, case markers and delimiters. Case markers are suffixed to NPs, whereas delimiters are suffixed not only to NPs but also to other categories such as adverbs and connectors. Case markers and delimiters interact with each other when suffixed to NPs (see Yang 1972). Both types of particles have long been recognized, and they generally refuse any morphological analysis. Interestingly enough, the structural sequence of *poko*, which consists of the verb *po* 'see' and the connector *ko* 'and', is emerging as a new particle. This particle behaves like a case marker, and no existing case marker properly overlaps with it.

The emerging particle *poko* has functions such as (i) sequence, (ii) target/goal, (iii) affection, (iv) beneficiary, (v) embedded-subject, and (vi) requestee. These uses may be viewed as forming a scale ranging from "concrete/literal" at the one end to "abstract/metaphorical" at the other, embracing a notion that is lacking in any of the existing case markers in Korean. This scale reflects the sense extension of the verb *po* 'see'. The emergence of this new case marker is viewed as a natural consequence of the general principle that states that functions must somehow be formally accommodated. The constellation of functions associated with this particular marker is given the mnemonic acronym "STABER" in this essay (see [9] below), and its formal marker may be referred to as "the STABER marker."

A cognitively interesting question is why it is that out of the five senses—seeing, hearing, smelling, tasting, touching—only the verb *po* 'see' qualifies for the emerging particle. My assumption is that "seeing" is the unmarked/representative member of all the senses. This assumption may be supported by the English saying "Seeing is believing," which is expressed in some Oriental cultures as "One seeing is better than one hundred hearing(s)."

The emerging particle *poko* in Korean is a good example of "grammaticalization" (see Hopper and Traugott 1993). This essay is organized as follows: section 2 discusses whether the structural sequence *poko* behaves like case markers or delimiters; section 3 deals with the sense extension of *poko*, which makes up the 'concrete/literal to abstract/metaphorical' scale; section 4 discusses whether existing case markers may be used alternatively with the emerging particle *poko*; section 5 discusses why no sense other than *po* 'see' makes up the "STABER" marker; and section 6 deals with grammaticalization and lexicalization of this marker. Section 7 summarizes the main points of this study.

---

2. Preliminaries

In Korean, more than one verb may make up compound verbs by using verbal connectors. There are two types of verbal connectors—*ko* and *A*. Hence, compound verbs include (i) *ko* compound verbs, (ii) *A* compound verbs, and (iii) combinations of *ko* and *A* when more than two verbs are compounded (see Sohn 1986:211–219, Yang 1993:335–357).[2] When two unit verbs are involved, the compound verbs have the following forms:

(1) a. V₁   -*ko*-   V₂   (*ko* compound verb)

    b. V₁   -*A*-   V₂   (*A* compound verb)

The emerging particle *poko* is the sequence of *po* in the V₁-slot and the connector *ko*, and V₂ is empty. Instead of a verb in the V₂-slot, some discourse-relevant expression follows. Thus, this marker has the following form:

(2) V₁ (=*po*) -*ko*

To determine whether the particle *poko* is a case marker or a delimiter, we have to turn to its distribution. Case markers are suffixed only to NPs, whereas delimiters are suffixed not only to NPs but also to other categories such as adverbs and connectors, both sentential and verbal. When a case marker and a delimiter are both suffixed to NPs, NP + case marker + delimiter, some case markers are obligatorily deleted, others are optionally deleted, and still others coexist with delimiters (see Yang 1972:32–115).

Delimiters may be suffixed to an adverb, but case markers may not. Since *poko* may not be suffixed to adverbs, this emerging particle is not a delimiter but a case marker, as illustrated in (3).

(3) a. John-i   Hankukə-lil   swipke   paeu-in-ta
        John-NOM   Korean-ACC   easily   learn-PRES-PRT
        'John learns Korean easily.'

    b. John-i   Hankukə-lil   swipke-nin/to/man   paeu-in-ta
        John-NOM   Korean-ACC   easily   DEL(IMITER)   learn-PRES-PRT
        'John learns Korean easily (plus the meaning of each delimiter)'[3]

    c. *John-i   Hankukə-lil   swipke-poko   paeu-in-ta

---

2. When more than two unit verbs are compounded, a variety of combinations is possible: V₁-*ko*-V₂-*ko*-V₃, V₁-*A*-V₂-*A*-V₃, V₁-*ko*-V₂-*A*-V₃, V₁-*A*-V₂-*ko*-V₃, etc. For example:

    na-nin i   sakwa-lil John-kwa nanu-A-mək-A-po-ko-siph-A-ci -   in-   ta
    I-DEL this apple-ACC John-with divide eat   try   desire become PRES PRT
    'I begin to [get the] feel[ing that I should] share this apple with John.'

Here, five unit verbs are compounded. The connector *A* is a morphophoneme that becomes [a] or [ə] according to Korean vowel harmony. This also applies to the morphophoneme *A* of the past tense *Ass*. Vowel harmony is becoming loose in contemporary Korean.

    Abbreviations used include ACC, accusative; CON, connector; DEL, delimiter; HON, subject honorific; NOM, nominative; PRES, present; PRT, statement-ending particle.

3. For detailed semantic and pragmatic analyses of delimiters, see Yang 1993:221–271.

Example (3b) shows the permissibility of attaching a delimiter to an adverb, while (3c) shows the ungrammaticality of suffixing *poko* to the adverb. This suggests that *poko* is not a delimiter, and thus far, it is enough to identify the emerging particle as a case marker.

One might argue that *poko* behaves like delimiters, because it may occur after the verbal connector *A*, like delimiters, citing (4) as an example.

(4) Kilsoo-ka   Niagara phokpho-lɨl ka-A-po-ko   nolla-Ass-ta
     Kilsoo-NOM Niagara falls-ACC   go-and-see-and surprised-PAST-PRT
     'Kilsoo was surprised by going and seeing Niagara Falls.'

Here the compound verb structure is 'V_1 (= *ka*) + CONNECTOR (= *A*) + V_2 (= *po*) + CONNECTOR (= *ko*)', which means 'go and see and'. The sequence *poko* is part of the whole compound verb, and *poko* in this context has only its literal meaning, nothing metaphorical. Hence, sentence (4) does not hinder my claim that *poko* is a case marker. This claim is clarified in the following sections. The next question to deal with is what kind of case marker *poko* is. This question is discussed in the next section.

3. Sense Extension

In the state of pregrammaticalization, *poko* behaves like other verbs in the V_1-slot, plus the connector.

(5) a. Kilsoo-ka   mək-ko-no-   in-ta
        Kilsoo-NOM   eat-and-not.work PRES-PRT
        'Kilsoo eats and does not work.' (Kilsoo is idle.)

     b. Kilsoo-ka   nol-ko-mək-   nin-ta
        Kilsoo-NOM   not.work-and-eat- PRES-PRT
        'Kilsoo does not work and eats.' (Kilsoo lives without working.)

(6) a. Kilsoo-ka   mək-ko-po-   in-ta
        Kilsoo-NOM   eat-and-see   PRES-PRT
        'Kilsoo eats (something) first and then sees (what happens).'

     b. Kilsoo-ka   po-ko-mək-   nin-ta
        Kilsoo-NOM   see-and-eat   PRES-PRT
        'Kilsoo eats (something) after considering the situation.'

We see, through (5) and (6), that *poko* has its original literal (not metaphorical) use: 'see and'. Interestingly enough, sense extension takes place with the "STABER" marker. Consider the following:

(7) a. uli-ka   salam-il   han-kaci-poko yəl-kaci-lɨl   al-suiss-ta
        we-NOM person-ACC one-thing   ten-thing-ACC know-can-PRT
        'We can infer a person's wholeness by seeing one of his behaviors.'

     b. jangsakkun-in o   li-poko sip li   ka-in-ta[4]
        merchant-DEL five profit   ten distance.unit go-PRES-PRT
        'Merchants go ten miles, looking for five percent profit.'

    c. sənsaengnim-i kɨ-ɨi apəci-poko atəl-ɨl   yongsəhaecu-Ass-ta[5]
       teacher-NOM his father   son-ACC forgive-PAST-PRT
       'The teacher forgave the son, considering his father.'

From the meaning of (7a), 'seeing' does not necessarily connote some physical obser-
vation. It may also mean 'considering', 'hearing', 'experiencing', 'touching', or the
like. Hence, in this sentence, sense extension has taken place. In (7b), the words 'ten'
and 'five' are compared in order to show that even a little profit triggers merchants to
go a long way. Here, 'seeing' means not 'looking at' but 'looking for', which shows
that sense extension has also taken place. In (7c), 'seeing' means 'considering'—a
mental state and not necessarily any physical observation. In all these sentences above,
the verb *po* 'see' is used metaphorically. There is a tendency for sense extension to go
from the concrete/literal state to the abstract/metaphorical state. This is a good exam-
ple of grammaticalization, which is discussed in more detail in section 6.

    The emerging particle *poko* has several uses: (7a) is a case of 'sequence' use; (7b),
'target/goal' use; and (7c), 'affection' use (which may become 'influence' if affection is
strong). The sense extension of *poko* goes further, as can be observed in (8).

   (8) a. Kilsoo-ka   uli nuna-poko uli cip- e   kkoch-ɨl   ponaeo-Ass-ta[6]
       Kilsoo-NOM my elder.sister my house-to flower-ACC send-PAST-PRT
       'Kilsoo sent flowers to my house for my elder sister.'

      b. Kilsoo-ka   Yongphali-poko totuknom i   la-ko   hae-Ass-ta
       Kilsoo-NOM Yongphali   thief   be PRT-quote say/do-PAST-PRT
       'Kilsoo said that Yongphali was a thief.'

      c. Kilsoo-ka   na-poko Yonghee-lɨl towatal-la-ko yochənghae-Ass-ta
       Kilsoo-NOM me   Yonghee-ACC help-PRT-quote request-PAST-PRT
       'Kilsoo asked me to help Yonghee.'

In (8a), *poko* does not signify any physical state, but only a beneficial state. The trans-
lation 'for my sister' may mean either the ultimate goal or affection.[7] At any rate, no
physical observation of 'seeing' is involved. (8b) means that Kilsoo said that Yong-
phali was a thief. Here, *poko* may mean a physical observation (i.e., deictically) or a
mental reference (i.e., nondeictically). (8b) is appropriate even for the case where
deixis is not relevent, as for example in letters. (8c) means that Kilsoo asked me to
help Yonghee. Here, *poko* may refer to a physical observation or may be used non-
deictically. Suppose Kilsoo sent me a letter, asking me to help Yonghee. This situa-
tion is appropriately expressed by (8c). The uses of *poko* in (8) are beneficiary in (8a),

---

4. *Li* is a homonym; it refers to (i) distance mileage, which is roughly equivalent to 400
   meters, or (ii) a profit rate of one percent. Thus sentence (7b) is an example of a pun. It
   implies that merchants make great efforts for even a small profit.

5. Sentence (7c) has an alternative, *apəci-lo po-A-(sə)*. In either case, the verb *po* is involved.

6. For the alternation between *naɨi /nae* 'my' and *uli* 'our', see Yang 1993:359–387.

7. In English, "for somebody" in the frame of sentence (8a) means the ultimate goal, e.g.,
   "John sent me a gift for my sister" means 'John sent a gift to my sister via me.'

embedded-subject in (8b), and requestee in (8c). All of these uses are far from the physical observation of 'seeing'.

These uses of *poko* illustrated above may not be exhaustive, but they cover the major ones. Thus, the major typology of the sense extension of this emerging particle, the STABER marker, is as follows:

(9) a. Sequence
b. Target/Goal
c. Affection/Influence
d. Beneficiary
e. Embedded-subject
f. Requestee

These uses of *poko* may be divided into two subtypes: "sequence use" and "others." In affirmative sentences, literal meanings of *po* require visible objects, while mental meanings of *po* do not. In other words, objects must be inside the deictic scope of seeing for the literal use of *po*, and objects may be outside the deictic scope of seeing for the metaphorical use. This is a natural consequence of the sense extension of *po* 'see', which goes "from a concrete/literal to an abstract/ metaphorical" state.

What should be noted at this point is the distinction between the verb *po* 'see' and the rest of the verbs for the five senses—seeing, hearing, smelling, tasting and touching. In the compound verb frame 'V$_1$-ko-$\emptyset$', verbs for all senses other than seeing in the V$_1$-slot have only literal meanings, not metaphorical ones. Hence, they cannot participate in the emerging particle. The reason for this is given in section 5.

Since the verb *po* of the emerging particle *poko* is a transitive verb, a natural question is whether the accusative marker *lil/il*, which marks objects, can be affixed to the *poko*-hosting NP, that is, NP-*lil-poko*. For this purpose, I will use examples already cited in (7) and (8).

(10) a. SEQUENCE USE
uli-ka   salam-il   han-kaci-lil-poko yəl-kaci-lil   al-suiss-ta
we-NOM person-ACC one-thing-ACC   ten-thing-ACC know-can-PRT
'We can infer a person's wholeness by seeing one of his behaviors.'

b. TARGET/GOAL USE
jangsakkun-in o   li-lil-poko sip li   ka-in-ta
merchant-DEL   five profit   ten distance.unit go-PRES-PRT
'Merchants go ten miles, looking for five percent profit.'

c. AFFECTION USE
sənsaengnim-i ki-ii apəci-lil-poko atil-ill   yongsəhaecu-Ass-ta
teachter -NOM his   father   son-ACC forgive-PAST-PRT
'The teacher forgave the son, considering his father.'

d. BENEFICIARY USE
?Kilsoo-ka   uli nuna-lil-poko uli cip-e   kkoch-lil   ponaeo-Ass-ta
Kilsoo-NOM my elder-sister   my house-to flower-ACC send-PAST-PRT
'Kilsoo sent flowers to my house for my elder sister.'

 e. EMBEDDED-SUBJECT USE
  *Kilsoo-ka Yongphali-lɨl-poko totuknom i la-ko hae-Ass-ta
  Kilsoo-NOMYongphali-ACC thief be PRT-quote say-PAST-PRT
  'Kilsoo said that Yongphali is a thief.'

 f. REQUESTEE USE
  *Kilsoo-ka na-lɨl-poko Yonghee-lɨl towatal-la ko
  Kilsoo-NOM me Yonghee-ACC help-PRT quote
  yochənghae-Ass-ta
  request-PAST-PRT
  'Kilsoo asked me to help Yonghee.'

In (10a), *poko* has the sequence use, in (10b), the target/goal use, and in (10c), the affection use. In all three, the accusative marker *lɨl*, which marks an object NP, may be affixed to the *poko*-hosting NP. On the other hand, the *poko*-hosting NP with the accusative marker is marginal in the beneficiary use of (10d), and is not acceptable in the embedded-subject use of (10e) and the requestee use of (10f). What is important here is that the sentence versions without the accusative marker sound a little more natural than the sentence versions with the accusative marker. This implies that the *poko*-hosting NP has a more important role as the *poko*-host than as the object NP of the verb *po* 'see' as part of *poko*.

 In the history of English, case marking was dropped (except for some pronouns) with the emergence of a fixed word order. This tendency may also take place eventually in other languages. In Korean, for instance, the accusative marker tends to be omitted especially when the verb and its object are adjacent. Under this circumstance, the object and the accusative marker are predictable. If something is predictable, it is redundant. If something is redundant, it may come to be an unnecessary burden that naturally disappears. In contemporary spoken Korean, the omission of the accusative marker is quite common (see Yang 1996, forthcoming).

 I will digress a little here. Some specific cases of compound verbs provide an interesting phenomenon for the omission of the accusative marker. Consider (11).

 (11) a. nae-ka sakwa-ka mək-ko-siph-ta
    I-NOM apple-NOM eat-CON-desire-PRT
    'I am desirous of eating apples.' (I want to eat apples.)

   b. nae-ka sakwa-lɨl mək-ko-siph-ta
    I-NOM apple-ACC eat-CON-desire-PRT
    [same translation as (a)]

   c. na-nɨn/to/man sakwa-ka/lɨl mək-ko-siph-ta
    I-DEL   apple-NOM/ACC eat-CON-desire-PRT

   d. nae-ka/nɨn sakwa-∅ mək-ko-siph-ta
    I-NOM/DEL

The subject NP is marked with the nominative marker *ka/i* or with any delimiter (*nɨn/ɨn, to, man*, etc.), depending on the context and the speaker's intention. The object NP is marked with the accusative marker *lɨl/ɨl* or with any delimiter. What concerns us

here in (11) is the object NP and its case marker. The nominative marker and the accusative marker alternate for the object NP in (11a, b, c). This alternation hinges on the structure of the compound verb *mək-ko-siph*, where *mək* in the V_1-slot means 'eat' (which requires an object), *ko* is the connector, and *siph* in the V_2-slot means 'be desirous of' (which requires the nominative marker for its object). When the speaker, though unconsciously, emphasizes V_1 (= *mək*), the object NP is marked with the accusative marker. On the other hand, when the speaker emphasizes V_2 (= *siph*), the object NP is marked with the nominative marker.

Korean has predicates of "self-judgment," mainly mental adjectives, such as *siph* 'be desirous of', *coh* 'be fond of', *silh* 'not be fond of', *yeppi* 'be pretty', and so forth (see Yang 1972: 159–174). For instance, the English sentence (12a) has several Korean counterparts, listed as (12b–e).

(12) a. I am fond of   kimchi
      NP_1 NP_2

    b. kimchi-ka   na-eke   coh-ta
      kimchi-NOM   I-for   fond-PRT
      NP_2 NP_1

    c. na-eke   kimchi-ka   coh-ta
      NP_1 NP_2

    d. nae-ka   kimchi-ka   coh-ta
      NP_1 NP_2

    e. na-nin   kimchi-ka   coh-ta
        -DEL
      NP_1 NP_2

(12b) is ambiguous due to the two meanings of the predicate *coh*. The sentence means (i) 'I am fond of kimchi', and (ii) 'kimchi is good for me'. We will be concerned only with the first sense, 'be fond of'. (12b) alternates with (12c), although both sound less natural. That is, they sound like logical paraphrases. The more natural one is (12d), where two NPs have the same nominative case marker. Unless the NPs are of a "macro-micro relation" (see Yang 1972: 41–58), case marker doubling/multiplying sounds less natural. Hence, the most natural version is (12e), where NP_1 has the delimiter *nin* and NP_2 keeps the nominative marker. Another acceptable alternative is for NP_1 to be marked nominative and NP_2 to have a delimiter, or for both NPs to have delimiters. This is a characteristic of predicates of self-judgment.

Returning now to (11), the nominative and accusative alternation for *sakwa* 'apple' is naturally explained. Both the nominative and the accusative versions, (11a) and (11b), for *sakwa* 'apple' sound natural. What is more natural, however, is the version without the case marker for *sakwa* 'apple' (11d). The naturalness of (11d) is two-fold, both from the naturalness of the omission of the accusative marker, and from the neutrality of the speaker's emphasis.

With regard to the compound verb of the form "V_1-*ko-siph*," it behaves as in (13).

(13) a. na-nin   Hawaii-e   ka-ko-siph-ta
        I-DEL   -goal   go-CON-desirous-PRT
        'I want to go to Hawaii.'
     b. na-nin   Hawaii-ka   ka-ko-siph-ta
     c. na-nin   Hawaii-lil   ka-ko-siph-ta
     d. na-nin   Hawaii-∅   ka-ko-siph-ta

Sentences (b) through (d) have the same meaning as sentence (a). The compound verb has the structure of *ka* 'go' in the V_1-slot and *siph* 'be desirous of' in the V_2-slot. When the speaker emphasizes, although unconsciously, *ka* 'go', (13a) results. When the speaker emphasizes *siph* 'desirous', (13b) results. (13c) occurs due to the accusative intrusion/replacement of some other case markers (see Yang 1972:95–115). Here also as in (11), the version without the case marker in (13d) sounds most natural for the same reason as previously stated.

Note that not all compound verbs behave like "V_1-*ko-siph*." For example, the existence verb *iss* in the V_2-slot of the "V_1-*ko-iss*" compound form does not behave like "V_1-ko-siph." Consider the following:

(14) a. Kilsoo-ka   chaek-il   ilk-ko-iss-ta
        Kilaoo-NOM   book-ACC   read-CON-exist-PRT
        'Kilsoo is reading a book.'

     b. *Kilsoo-ka   chaek-i   ilk-ko-iss-ta
           -NOM   -NOM

     c. Kilsoo-ka   chaek-∅   ilk-ko-iss-ta

In (14a), the object NP *chaek* 'book' is marked with the accusative marker, which is required by the transitive verb *ilk* 'read' in the V_1-slot. (14c) is grammatical due to the omissibility of the accusative marker. The problem is (14b), where the object NP *chaek* is marked with the nominative marker, which might be expected due to the existence verb *iss* in the V_2-slot, which is analogous to the verb *siph* 'be desirous of' in the V_2-slot of the "V_1-*ko-siph*" compound verb. Nevertheless, this analogy does not hold. That is, the case marker of the object NP *chaek* cannot be assigned based on the existence verb *iss* in the V_2-slot. The reason is that the existence verb *iss* in the "V_1-*ko-iss*" compound verb form has lost its independent state as a unit verb of compound verbs. The existence verb *iss* is "aspectualized" in the form of "V_1-*ko-iss*," which expresses the progressive aspect (see Yang 1993:317–334). The aspect system of Korean is the result of such grammaticalization (see section 6).

Now apart from the accusative marker and its alternation, let us consider the subject-honorific *si/isi*. Because verbs may get the subject-honorific *si* (*isi* after a consonant), another natural question is whether *po* of the emerging particle *poko* is verb-like in this respect. For this purpose, I use the already cited examples (7) and (8), or similar ones.

(15) a. SEQUENCE USE
         əlin-til-in   salam-il   han-kaci-posiko   yəl-kaci-lil   al-suiss-ta
         adult-PL-DEL person-ACC one-thing   ten-thing-ACC can-know-PRT
         'Adults can infer a person's wholeness by seeing one of his behaviors.'

b. TARGET/GOAL USE
jangsakkun-in o   li   posiko sip li   ka-in-ta
merchant-DEL five profit.rate   ten distance.unit go-PRES-PRT
'Merchants go ten miles, looking for five percent profit.'

c. AFFECTION USE
sənsaengnim-iki-ii apəci-posiko atil-il   yongsəhaecu-Ass-ta
teacher -NOM his   father   son-ACC forgive-PAST-PRT
'The teacher forgave the son, considering his father.'

d. BENEFICIARY USE
?Kilsoo-ssi-ka   uli nuna-   posiko uli cip-e   kkoch-il
Kilsoo-HON-NOM my elder.sister   our house-to flower-ACC

ponaeo-Ass-ta
send-PAST-PRT
'Kilsoo sent flowers to my house for my elder sister.'

e. EMBEDDED-SUBJECT USE
*Kilsoo-ssi-ka   Yongphali-posiko totuknom i   la-ko
Kilsoo-HON-NOM Yongphali   thief   be PRT-quote

hae-Ass-ta
do-PAST-PRT
'Kilsoo said that Yongphali was a thief.'

f. REQUESTEE USE
*Kilsoo-ssi-ka   na-posiko Yonghee-lil   towatalla-ko
Kilsoo-HON-NOM me   Yonghee-ACC help-quote

yochənghae-Ass-ta
request-PAST-PRT
'Kilsoo requested me to help Yonghee.'

When speakers want to express deference to the subject NP of a sentence, they attach the subject-honorific *si/isi* to the verb.[8] (15a–c) are all acceptable, (15d) is marginal, and (15e–f) are unacceptable. This difference suggests that the sequence, target, and affection uses of *poko* are not far from the literal meaning of *po* 'see', while the beneficiary, embedded-subject, and requestee uses are far. In other words, this test of the subject-honorific *si/isi* also shows that the verb *po* 'see' of the emerging particle *poko* has a metaphorical sense; sense extension has taken place in *poko*.

Thus, both the accusative test and the subject-honorific test support our assumption that the sequence, target, and affection uses/functions of the emerging particle *poko* are closer to the literal meaning of the verb *po*, whereas the beneficiary, embedded-subject, and requestee uses are closer to the metaphorical meaning of the verb *po*. This difference reflects the literal–metaphorical scale of table 1, which summarizes the

8. When a sentence has several verbs, e.g., compound verbs and others, the subject-honorific *si/isi* might be attached to every verb. However, such a repetition is redundant and sounds odd. The economical way of using *si/isi* in this instance is to attach it to only one verb, usually the final one.

observations discussed thus far. This dichotomy is supported by two syntactic tests, as we have seen, and corresponds to the degrees of grammaticalization of *poko;* the first three uses are in a state of flux, while the last three have been grammaticalized.

Table 1. Literal–Metaphorical Scale

USE	LITERAL/ METAPHORICAL	ACCUSATIVE	HONORIFIC	GRAMMATICAL- IZATION
a. Sequence		O.K.	O.K.	
b. Target	MORE LITERAL	O.K.	O.K.	IN FLUX
c. Affection		O.K.	O.K.	
d. Beneficiary		??	??	
e. Embedded- subject	MORE METAPHORICAL	NO	NO	ACCOMPLISHED
f. Requestee		NO	NO	

This literal–metaphorical scale may be reinforced by another syntactic test. The typical form of the sequence use of *poko* has 'V$_1$-*ko*-V$_2$', which may also be expressed by 'V$_1$-*ko-(na)sǝ*-V$_2$'. The latter form may be contracted to *po-ko-sǝ* with respect to *poko. Nasǝ* or *sǝ* means 'after doing something'. Hence, *pokosǝ* means 'after seeing something'. *Pokosǝ* may be employed for the sequence use of *poko,* but not for the other uses. This shows that the target and the affection uses of *poko* are rather far from the literal use of *po* 'see'; the target and the affection uses are also closer to the metaphorical side.

The literal-metaphorical scale is not an artifact, but a natural representation of the direction of sense extension, which goes from concrete/literal to abstract/metaphorical. This direction of sense extension seems to be universal. An example from English, say, the verb 'clinch' has gone through the following sense extension:

(16) a. to bend the protruding point of a nail
 b. to fasten together (even as in boxing)
 c. to make sure/settle

4. Alternation

This section discusses whether the emerging particle *poko* may alternate with existing case markers. For this purpose, a bit of background on the Korean case markers is given in table 2. The examples of (17) show us that the emerging particle *poko* may be used in alternation, in a limited way, with certain existing case markers.

(17) a. Kilsoo-ka   na-poko/*eke/*tǝlǝ   tomangchi-Ass-ta
 Kilsoo-NOM   me   run.away-PAST-PRT
 'Kilsoo ran away, seeing me.' (sequence use)

71

b. kae-ka nalakanin talk- poko/ʔhanthe/*tələ ttwiyəka-Ass-ta
dog-NOM flying chicken run-PAST-PRT
'The dog ran toward the flying chicken.' (target use)

c. sənsaengnim-iki-ii apəci-poko/*eke/*tələatil-il yongsəhaecu-Ass-ta
teacher-NOM his father son-ACC forgive-PAST-PRT
'The teacher forgave the son, considering his father.' (affection use)

d. Kilsoo-ka nae nuna-poko/*eke/*tələ uli cip-e sənmul-il
Kilsoo-NOM my elder.sister our house-to gift-ACC

ponaeo-Ass-ta
send-PAST-PRT
'Kilsoo sent a gift to my house for my elder sister.' (beneficiary use)

e. Kilsoo-ka Yongphali-poko/ka/lil/ʔeke/*tələ totuknom i la-ko
Kilsoo-NOM Yongphali thief be PRT-quote

hae-Ass-ta
do/say-PAST-PRT
'Kilsoo said that Yongphali is a thief.' (embedded-subject use)

f. Kilsoo-ka na-poko/eke/tələ Yonghee-lil towatal la-ko
Kilsoo-NOM me Yonghee-ACC help PRT-quote

yochənghae-Ass-ta
request-PAST-PRT
'Kilsoo requested me to help Yonghee.' (requestee use)

Table 2. Korean Case Markers

	BASIC FORM	ALTERNANTS			
		/C___	NP]HUMAN	NP]ANIMATE	NP]HONORIFIC
NOMINATIVE	ka	i			
ACCUSATIVE	lil	il			
INSTRUMENT	lo	ilo			
GOAL TOWARD	lo	ilo			
GOAL TERMINUS	e		eke*	hanthe	kke
CAUSER	e		eke	hanthe	kke
SOURCE	esə		ekesə	hanthesə	kkesə
GENITIVE	ii				
COMITATIVE	wa	kwa			
(ARCHAIC)	tələ				

* _e_, which has the function of goal terminus or causer, becomes _eke (< eke-e)_ if the particle-hosting NP is human; it becomes _hanthe (< hanthe-e)_ if the NP is animate; it becomes _kke (< kke-e)_ if the NP is honorific. These three forms, _eke, hanthe_, and _kke_, are usually called "dative" markers. Likewise, the source marker _esə_ becomes _ekesə, hanthesə_, or _kkesə_, respectively, depending on the nature of the NP, as with _e_ (see Yang 1972:32–115).

Among the six different uses, only three—the target, embedded-subject, and requestee—allow some existing case markers to some degree, as summarized in (18).

(18) a. Target: poko, ?hanthe
     b. Embedded-subject: poko, ka, lil, ?eke, ?tələ
     c. Requestee: poko, eke, tələ

In total, the alternative case markers of *poko* are: (i) *tələ*, (ii) *eke, hanthe*, (iii) *ka*, and (iv) *lil. Tələ* is marginal in the embedded-subject use but is acceptable in the requestee use. It may be significant to mention, however, that *tələ* sounds archaic and is vanishing. As pointed out in the note to table 2, *eke* and *hanthe* are the same as *e. E* in the target and embedded-subject uses is marginal. Then, the problematic cases boil down to (i) *eke* and *tələ* in the requestee use and (ii) *ka* and *lil* in the embedded-subject use. In the requestee use, it is most natural to use *poko* instead of *eke* or *tələ*. In the embedded-subject use, on the one hand, *poko*, *ka*, and *lil* are natural. Thus, the real problem is the nominative marker *ka* and the accusative marker *lil* in the embedded-subject use of *poko*.

In terms of transformational grammar, we can trace the source of the nominative marker *ka* and the accusative marker *lil* in the embedded-subject use. For example, (17e) has the following process:

(19) a. UNDERLYING
        Kilsoo-ka   [s Yongphali-ka totuknom i ta]- ko   malhae-Ass-ta
        Kilsoo-NOM Yongphali-NOM thief   be PRT quote say-PAST-PRT
        'Kilsoo said that Yongphali was a thief.'

     b. "I-TA" TO "I-LA" CHANGE
        Kilsoo-ka   [s Yongphali-ka totuknom i la]-ko   hae-Ass-ta
        Kilsoo-NOM Yongphali-NOM thief   be

     c. SUBJECT-TO-OBJECT RAISING
        Kilsoo-ka   Yongphali-lil totuknom i la-ko   hae-Ass-ta
        Kilsoo-NOM Yongphali-ACC thief   be quote

Note that the underlying structure (19a) has an embedded sentence, *Yongphali-ka totuknom i ta*, where the statement-ender (i.e., the last element of the modality component in case-grammar terms) is *ta*. When an equative or predicate nominal sentence is embedded as the quote-*ko* complement, the statement-ender *ta* must necessarily be changed to *la* when the embedded equative sentence is in the present tense (see Yang 1972: chapter 3). The verb *malha* is usually reduced to the pro-verb *ha*[9] in the quote-*ko* context. Thus,

---

9. The pro-verb *ha* 'do' is a verbalizer that turns nouns into their cognate verbs—denominal verbs. In this process, the noun must have some verbal property (see Suh 1975, Yang 1993: 389–422). There are two types of verbal property: inherent and process. For example, (i) *kongpu-ha* 'study-do' → 'study'. In (i), the noun 'study' has its inherent verbal property. (ii) *namu-ha* = 'tree-do' → 'make a tree into firewood'. In (ii), 'tree' has no inherent verbal property. Instead, in the process of making a tree into firewood, the verbal property is involved. The Japanese counterpart, 'NP-*su*' works in the same way. The verbalizer *ha* also turns mental adjectives into their cognate verbs, using the connector *A*: e.g., *coh* 'fond of' → *coh-A-ha* 'like'.

we use (19b), but not (19a). (19b), in turn, may optionally be changed to (19c), where the original embedded-subject NP *Yongphali* is marked with the accusative marker, which is usually affixed to the object NP. This shows that the embedded subject is raised to the matrix object position. Thus, in (17e), three case markers—*poko*, the nominative *ka*, and the accusative *lil*—are used alternatively for *Yongphali.*

    At this point in our discussion, an interesting thought arises. The existing case markers cannot properly and naturally mark the uses/functions under consideration, except for the embedded-subject use. This situation triggers the need for some case marker to arise. This formal gap is properly and naturally filled by the emerging particle *poko.* This is a natural consequence of the general principle that functions are to be mapped by forms one way or another. But we are left with the question as to why the emerging particle *poko* should invade the territory of the nominative and accusative markers in the embedded-subject use (see [19b, c]). This phenomenon might appear to undermine my claim that the particle *poko* is emerging in order to fill a formal gap. However, we can view it as the consequence of a spreading out of the emerging particle from the uses it was initially called upon to fill even into the embedded-subject use. Thus, the emerging particle *poko* is a beautiful example of grammaticalization, and it will eventually be settled as a case marker, as is discussed further in section 6.

5. Cognitive Support

A cognitively interesting question to address is why out of five senses, *po* 'see' is singled out to be the emerging case marker. To put the conclusion first, "seeing" is the most salient sense, at least in Korean culture. That is, "seeing" is the unmarked/representative perceptual and conceptual member of the five senses. There are various shreds of evidence that point in this direction. First, medical doctors check a patient's eyes first before anything else, because the eyes are the most sensitive sense organ. Second, a blind person is considered to be most miserable. Third, in Korean, the "hello" in a telephone call is expressed by *yo-po-seyo*, which means "look here please", intimating the significance of the sense of sight. Fourth, in Korean, it is said that learning is done by seeing, which is expressed as *po-paeu-ta* (< *po-ko-paeu-ta*), which means 'learn by seeing'. Finally, English has the saying, "Seeing is believing." This saying is expressed in Chinese, Japanese, and Korean as "One seeing is better than one hundred hearings." I think a similar saying exists in almost every culture in one form or another.

    One might argue that the choice of the verb *po* 'see' is accidental. If we adopt this simple and naive view, the story ends. We need to make the story intellectually interesting. In order to accommodate this need, I claim that the choice of the verb *po* 'see' for the emerging particle *poko* is not accidental; it is closely related to our cognition.

    When the verb *po* 'see' makes up a compound verb with the verbal connector A (V_1-A-po), *po* in the V_2-slot is usually translated into English as 'try', thus, 'try to V_1'. For example,

(20) a. mək-A-po 'try to eat (something)' : 'eat'
      b. ip-A-po 'try to put on (clothes)' : 'wear'
      c. ilk-A-po 'try to read (something)' : 'read'
      d. ka-A-po 'try to go (to some place)' : 'go'
      e. o-A-po 'try to come (here)' : 'come'

(21) po-A-po 'try to see (something)' : 'see'

In these examples, the verb *po* in the V_2-slot corresponds to English 'try'. Interestingly enough, in (21), the same verb *po* may occur in both the V_1- and the V_2- slots, and is translated 'see' in the V_1-slot, and 'try' in the V_2-slot.

Considering the data in (20) and (21), one might regard the two verbs *po* (Japanese *mi*) as homonyms, which would mean that *po* in the V_1-slot and *po* in the V_2-slot of (21) are not related in meaning. This homonymy view is not correct. The two *po*'s are semantically related, and *po* is a polysemous word.[10] The proper meaning of the compound verb in (21) is not (22b), but (22a).

(22) a. to have an experience of seeing (something)
      b. to try to see/seeing (something)

The confusion arises from the fact that the English verb "try" is a nonimplicative verb, according to Karttunen (1971). However, the Korean verb *po* (Japanese *mi*) is an implicative verb. As we know, the verb "try" subcategorizes for either an infinitive complement or a gerundive complement. If the infinitive complement is selected, the verb "try" behaves like a nonimplicative verb, whereas if the gerundive complement is selected, it behaves like an implicative verb. Dixon (1991:179) illustrates this semantic difference as follows:

(23) a. John tried to eat the cake.
      b. John tried eating the cake.

(24) a. John tried to travel by train.
      b. John tried traveling by train.

(25) a. John tried to catch the ball (but missed).
      b. John tried catching the ball (during the team's two-hour fielding practice).

Dixon says that the (a) sentences indicate that John wanted to engage in the activity, but may well not have been able to, whereas the (b) sentences indicate that John did engage in it for a sample period of time. If Dixon's judgment is correct, we can infer that the English verb 'try' is ambiguous between being implicative and nonimplicative.

Since the Korean verb *po* (Japanese *mi*) is necessarily an implicative verb, (21) has the following canonical form for its proper translation:

(26) a. po-A-po
      b. to have an experience of doing (something)

---

10. The verb *po* has another related sense, that of 'watch', as in *aeki-po* 'to baby-sit', *cip-po* 'to house-sit', etc.

(26a) has its proper meaning in (26b), which has two parts: 'having an experience' and 'doing (something)'. 'Doing (something)' represents any verb in the V_1-slot. 'Having an experience' represents the verb *po* 'see' in the V_2-slot. Note that 'having an experience' covers an experience of any kind/sense, which typically includes seeing, hearing, smelling, tasting, touching, and so on. All of these experiences are expressed by the verb *po* 'see'. This shows that seeing is the representative/unmarked member of all the senses. Thus, the compound verb *po-A-po* is not a semantically contradictory or redundant combination. The verb *po* in the V_1-slot represents the literal meaning of seeing, and the verb *po* in the V_2-slot represents the metaphorical meaning of seeing/experiencing.

Returning now to the emerging particle *poko*, the verb *po* here does not literally mean 'seeing'. Neither does it mean 'having an experience of doing (something)'. The verb *po* in the emerging particle *poko* is sense-extended far from its literal meaning and the first-step sense extension is 'having an experience'. The verb *po* has different senses, depending on the degree of sense extension.[11] This is represented in (27).

(27) Sense extension of *po* 'see'
    a. the literal meaning of 'seeing'
    b. having an experience of doing (something)
    c. uses/functions of the emerging particle

In support of the sense extension of the verb *po*, consider (28).

(28) a. Kilsoo-ka   holangi-poko tomangchi-Ass-ta.
       Kilsoo-NOM tiger   run.away-PAST-PRT
       'Kilsoo saw a tiger and ran away.'

    b. Kilsoo-ka   holangi-lil tha-poko yongki-lil   ət- Ass-ta.
       Kilsoo-NOM tiger-ACC ride   courage-ACC get-PAST-PRT
       'Kilsoo got courage, after having an experience of riding a tiger.'

    c. Kilsoo-ka   holangi-poko tampae-lil kkinh la- ko   hae-Ass-ta[12]
       Kilsoo-NOM tiger-ACC   tobacco-ACC quit PRT- quote do-PAST-PRT
       'Kilsoo asked the tiger to quit smoking.'

Poko has the literal meaning of 'seeing' in (28a), has the one-step extended meaning of 'having an experience of riding a tiger' in (28b), and has the further extended meaning of 'the metaphorical meaning of requestee' in (28c). In short, it is shown that the sense extension of the verb *po* 'see' has cognitive support.

---

11. I do not mean that other senses do not have extensions. For example, 'hearing/listening to' has the sense-extension that is universally idiomaticized as 'obey'. 'Smelling' has sense-extended as 'smelling bad', and 'tasting' translates into Korean as *mas-po* which means 'having an experience of eating'. *Mas-po* even means 'having an experience of doing (something)'.

12. Tigers are often cited in Korean folklore. One example says "Tigers used to smoke," which refers to primitive times. Another says "Tigers are having a wedding ceremony," which refers to that weather condition where the sun is shining while it is raining.

6. Grammaticalization

A currently attractive topic is "grammaticalization." This term was first used by Meillet (1912, 1958). More recently, Hopper and Traugott (1993:xv) have defined it as *"the process whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions,* and once grammaticalized, continue to develop new grammatical functions" [emphasis mine]. They say further (1993:1–2) that "'grammaticalization' as a term has two meanings: (i) as a term referring to a framework within which to account for language phenomena, it refers to that part of the study of language that *focuses on how grammatical forms and constructions arise, how they are used, and how they shape the language,* (ii) the term . . . also refers to the actual phenomena of language that the framework of grammaticalization seeks to address, most especially *the process whereby items become more grammatical through time"* [emphasis again mine]. The characteristics of "grammaticalization" are expressed in the italicized portions of these quotes, which cover both its historical and synchronic aspects. The salient feature of grammaticalization is the process that is expressed by the affixes *"-ize* or *-ization."* This process includes both the start and the terminus. Thus, grammaticalization means (i) that some form begins to serve some grammatical function, and (ii) that the form ends in being lexicalized. A case in point is the emerging case marker *poko* in Korean.

It is not known or traceable when *poko* began to serve such functions in the history of the Korean language. I have no solid statistics on the use of *poko* in written Korean. However, I am confident that it is as frequently used as the other particles in spoken Korean. Since language changes begin in the spoken form, this is not accidental.

At this point in the development of the Korean language, some Korean linguists might shake their heads as regards the present account. Although I am certain that the emerging case marker *poko* is a beautiful example of grammaticalization, it is fair to ask whether and when *poko* will more widely be judged or analyzed as such. My prediction that it eventually will is based on the following evidence. The uses/functions mentioned in table 1 cannot be performed by any existing case marker. This situation creates functional needs. Functional needs are somehow mapped onto forms in one way or another. This is a case of the universal saying that "Necessity is the mother of invention." Thus, certain forms have to be selected and used repeatedly to serve the functions mentioned in table 1 until they get settled. The settling down of such forms involves their lexicalization. Here again, a case in point is the emerging case marker *poko* in Korean.[13]

---

13. The Japanese counterpart *mite* does not seem to have emerged as a case marker. However, the Japanese counterpart will follow the course of the Korean *poko* because no existing case marker can cover the "STABER" functions dealt with in this essay.

7. Conclusions

Korean has two types of particles (i.e., dependent functional elements): case markers and delimiters. The emerging particle *poko*, which consists of the verb *po* 'see' and the connector *ko*, behaves like a case marker, and not like a delimiter. Case markers are suffixed only to NPs, whereas delimiters are suffixed not only to NPs but also to other categories such as adverbs and connectors. The emerging particle *poko* is suffixed only to NPs.

The emerging particle has undergone sense extension. *Poko* literally means (i) 'see and', (ii) 'looking for', and (iii) 'considering something'. These meanings correspond to their uses as 'sequence', 'target', and 'affection', respectively. The sense extension goes further. The emerging particle *poko* metaphorically has (iv) 'beneficiary', (v) 'embedded-subject', and (vi) 'requestee' functions. Thus, the sense extension of the emerging particle *poko* makes up a literal-metaphorical scale. The first three uses are closer to the literal side, while the last three are closer to the metaphorical side. This dichotomy has been shown to be supported by several syntactic tests, and it is a natural consequence of the sense extension of *poko*, which usually goes from concrete/literal to abstract/metaphorical.

Existing case markers may not happily substitute for the emerging particle *poko*, except in its embedded-subject use. In this function, *poko*, *ka* (nominative), and *lil* (accusative) can be used alternatively. The nominative and the accusative markers for the embedded-subject NP are not accidental. In the underlying structure, the embedded-subject is marked with the nominative marker. However, when the embedded-subject is raised to the matrix object position, it is marked with the accusative marker.

A cognitively interesting question is why only the verb *po* 'see' makes up the emerging particle *poko*. My assumption is that of the five senses the unmarked/representative member is "seeing" as evidenced by popular sayings in various cultures.

There may be some Korean linguists who might shake their heads at my claim that *poko* has emerged as a case marker. However, I am confident that shaking will change to nodding after reading the arguments in this essay.

REFERENCES

Bolinger, Dwight. 1977. *Meaning and form.* London: Longman.

Copestake, Ann, and Ted Briscoe. 1995. Sense-productive polysemy and sense extension. *Journal of Semantics* 12:15–67.

Dixon, R. M. W. 1991. *A new approach to English grammar on semantic principles.* Oxford: Clarendon Press.

Hopper, Paul J., and Elizabeth Closs Traugott. 1993. *Grammaticalization.* Cambridge: Cambridge University Press.

→ Karttunen, Lauri. 1971. Implicative verbs. *Language* 47:340–358.

Lakoff, George, and Mark Johnson. 1980. *Metaphors we live by.* Chicago: University of Chicago Press.

Martin, Samuel E. 1992. *A reference grammar on Korean: A complete guide to the grammar and history of the Korean language.* Rutland, Vermont: Charles E. Turtle Company.

Meillet, Antoine. 1912. Révolution des formes grammaticales. *Scientia* 12, no. 26. [reprinted in Meilett 1958].

———. 1958. *Linguistique historique et linguistique generale.* Paris: Champion.

Ortony, Andrew (ed.). 1979. *Metaphor and thought.* Cambridge: Cambridge University Press.

Sohn, Ho-min. 1986. *Linguistic expeditions.* Seoul: Hanshin.

Suh, Cheong-Soo. 1975. The grammar of the verb HA (in Korean). Ph.D. dissertation, Yonsei University.

Talmy, Leonard. 1985. Lexicalization patterns: Semantic structure in lexical forms. In *Language typology and syntactic description,* ed. by Timothy Shopen, vol. 3, 57–149. Cambridge: Cambridge Universiry Press.

Traugott, Elizabeth Closs, and Bernd Heine, eds. 1991. *Approaches to grammaticalization,* I and II. Amsterdam: John Benjamins.

Wierzbicka, Anna. 1988. *The semantics of grammar.* Amsterdam: John Benjamins.

Yang, In-Seok. 1972. *Korean syntax.* Seoul: Paekhapsa.

———. 1993. *Linguistic explorations: Syntax, semantics, and pragmatics.* Seoul: Hanshin.

———. 1996. Negative left-shifting in Korean (in Korean). *Language and Linguistics* 22. Hankuk University of Foreign Studies.

———. Forthcoming. Left-shifting phenomena in Korean.

6: POWER AND INTIMACY: A CONTRADICTION IN A THAI PERSONAL PRONOUN

Pranee Kullavanijaya
Chulalongkorn University

1. Personal Pronoun Systems in Tai

Languages in the Tai language family share a personal pronoun system that displays a three-way contrast in person, that is, a system that distinguishes the first, second, and third persons of personal pronouns proper.[1] However, when it comes to number, the Tai languages vary. Although probably most Tai languages show a two-way contrast in number, that is, a singular-plural distinction, some Tai languages display a three-way contrast. Tai Neua, a Tai dialect in the southwestern group now spoken in Dehong, Yunnan Province, China, is an example of a language with a three-way number contrast. As the table shows, there is inclusion or exclusion of the addressee(s) in the dual and plural numbers of the first person.

Table 1. Dehong Personal Pronouns Proper

	SINGULAR	DUAL	PLURAL	
1ST	kau 33	ha:ŋ 55 ha: 55	hau 55	INCLUSIVE
		ha:ŋ 55 xə 24	tu 33	EXCLUSIVE
2ND	matɯ 55	sɔŋ 24 xə 24	su 24	
3RD	man 55	sɔŋ 24 xa: 24	xau 24	
	xau 24			

The Bangkok Thai dialect, on the other hand, has only a two-way number contrast, as shown in table 2.[2] As can be seen in that table, there are variants in each box. These variants differ in, for example, gender, formality, and intimacy. Another point seen from the table is that some pronouns can occur in more than one box, for example, /khun 1/, /khaw 5/, /raw 1/. A form from one of these boxes can be assumed as the base form from which the other forms are derived. This study will focus on the characteristics of the four forms of /raw 1/ and their possible derivation.

---

1. In Bangkok Thai, kinship terms as well as some common nouns can be used as personal pronouns. To exclude these derived pronouns and include only the nonderived personal pronouns, such as /chan 24/ 'I', the term "personal pronoun proper" is used.

2. Table 2 does not give a complete list of personal pronouns proper in Bangkok Thai, because the complete list is not the focus of the paper. Only /raw 1/ will be discussed here.

    The numerals in the citation forms mark tones in Bangkok Thai, which has altogether 5 tones: no. 1 represents the mid-level tone or 33 ; no. 2 represents the low-level tone, 22; no. 3 represents high-falling tone, 42; no. 4 represents high tone, 423 ; no. 5 represents low-rising tone, 24.

Table 2. Bangkok Thai Personal Pronouns Proper

	SINGULAR	PLURAL
1ST	chan 5, dichan 5, phom 5, raw 1, kha:p 3 pha 4 caw 3, ku: 1	raw 1, phuak 3 raw 1
2ND	khun 1, thə: 1, tha:n 3, mɯɯŋ 1, raw 1	khun 1, thə: 1, raw 1, tha:n 3, mɯɯŋ 1
3RD	khaw 5, man 1	khaw 5, phuak 3 khaw 5, man 1

2. Occurrences of /raw 1/.

Table 2 shows that /raw 1/ may occur as the first person pronoun in plural number and in singular number, and as the second person pronoun in singular and plural numbers. Following are examples of /raw 1/ in these different occurrences.

First person, plural number

(1) ba:n 3   khun 1   som 5 chaj 1   yu: 2   maj 3   klaj 1
    house   Mr.   Somchai   is   not   far

raw 1   də:n 1   paj 1   di: 1   kwa: 2   naʔ 4
we   walk   go   good   more   FP [3]
'Khun Somchai's house is not far. We'd better walk there.'

(2) suʔ 2 da: 1   kap 2   dichan 4   hen 5   pho:ŋ4   tɔŋ 3   kan 3   wa: 2
    Suda   with   I   see   agree   agree   say

raw 1   khuan 1   də:n 1 tha:ŋ 1   paj 1   kɔ:n 1   khun 2
we   should   travel   go   before   you
'Suda and I agree that we should travel before you.'

Although both sentences (1) and (2) show /raw 1/ as the first person pronoun in plural number, there is a semantic distinction between the two instances. While /raw 1/ in sentence (1) includes the addressee, /raw 1/ in sentence (2) excludes the addressee. It can be said, therefore, that although inclusion and exclusion of the addressees can be recognized in context, there is no distinction in their form.

---

3. Abbreviations used include FP, final particle; Q, question word.

First person, singular number

(3) **raw 1**   ca? 2   pok 2 khrɔːŋ 1 phɛn 2 din 1   doːj 1   tham 1
       we   will   rule   country   by   justice
       'We will rule the country with justice.'
       (His Majesty King Bhumiphol's address to his people
       on the occasion of his ascension to the throne)

(4) Speaker A:
       thaː 3   khun 1   thuːk 2   khɔːŋ rɔːŋ 4   haj 3
       if   you   touch   ask   give

       ʔɔːk 2   khun 1   ca? 2   tham 1   yaːŋ 2 raj 1
       resign   you   will   do   how
       'If you are asked to resign, what would you do?'

       Speaker B:
       phom 5 waː 3   thaː 3 khon 1   thaŋ 4 laːj 5 maj 3   hen 5   waː 3
       I   say   if   man   all   not   see   say

       raw 1 tham 1 maj 3   chaj 3   phɯa 3 tua 1 **raw 1** ʔeːŋ 1
       we   do   not   right   for   body   we   self

       teː 2   phɯa 3 chaːt 3 **raw 1** kɔ 3   khoŋ 1 tɔŋ 3 paj 1
       but   for   nation   we   so   may   must   go
       'I think if people still do not see that we did everything not for ourselves
       but for the country, we'd have to resign.'

/raw 1/ as it occurs in sentence (3) has a restricted use. In an address to the public, this use of /raw 1/ is limited only to the King. Usually, the King and the royal family omit the pronoun for themselves. If a personal pronoun is used at all, /chan 5/ is heard more often than /raw 1/, which expresses power and—automatically—detachment from the addressees.

Example (4) presents an interesting use of /raw 1/ as a first person singular pronoun. As can be seen, Speaker B switches from /phom 5/ 'I' to /raw 1/ 'I' when he wants to make a statement that can be applied to anyone, not only to himself, that is, "If **one** does things for the country, not for oneself, and if people do not see it, then one should realize that it is useless to go on." It seems as though the speaker is trying to exclude himself from the environment, thus making his statement objective and general. In this usage then, /raw1/ carries the detachment feature as well as the inclusive feature in the sense that it is general and therefore applies to the addressee as well.

(5) (Female speaker to male addressee)
       thaː 3   khun 1 jaŋ 1   phɔː 1   miː 1   weː 1 laː 1 **raw 1**
       if   you   still   enough have   time   we

       jaːk 2   chuan 1 haj 3   maː 1   chuaj 3 ŋaːn 1 **raw 1**
       want   invite   give   come   help   work   we
       'If you still have some free hours, I would like to ask you to consider
       working with me.'

Example (5) reveals another interesting use of /raw ɪ/. In Thai culture, young girls and boys usually stay apart and are not close friends, unless they have a special relationship such as boyfriends or girlfriends. In some cases where the girl has male friends or vice versa, she/he will address them as /thəə ɪ/ 'you' and call herself/himself /raw ɪ/. In the case of acquaintances, in order to create a friendly atmosphere, the address term /khun ɪ/ 'you' is used with /raw ɪ/ as the term for the speaker. We can therefore say that /raw ɪ/ is an attempt to show a close relationship or a certain degree of intimacy.

(6) Speaker A: (Female speaker to intimate female addressee)
pha: 3 khlum ɪ laj 2   rə: 5,   no:n 3 ŋaj ɪ,
shawl   right?   there

phe:ŋ ɪ   na? 4   ca? 2   suɪ: 4   paj ɪ   fa:k 2   khraj ɪ
expensive FP   will   buy   go   give   who
'The shawls? Over there! They are quite expensive. Who are you going to buy it for?'

Speaker B:
raw ɪ   ca? 2   suɪ: 4   fa:k 2   mɛ: 4
we   will   buy   give   mother
'I will buy it for my mother.'

Speaker A:
tua ɪ   nia? 3   pen ɪ   lu:k 3   kə tan ɪ ju: ɪ
body   here   is   child   good

ciŋ ɪ ciŋ ɪ   na? 4   **raw ɪ**   wa: 3
indeed   FP   we   say
'You are such a good daughter, I think.'

Example (6) shows real intimacy. A female speaker uses /raw ɪ/ for herself and /tua ɪ/ to address her intimate friend, whether male or female.[4]

Second person, singular number

(7) Speaker A:
raw ɪ   maj 3   paj ɪ   duaj 3   ruɪ: 5
we   not   go   together Q

phu: 3 jaj 2   ?o:k 2 pa:k 2   lɛ:w 4
senior   speak   already
'Aren't you coming with us? The director himself mentioned that you should come.'

Speaker B:
paj ɪ   kha? 3   tha:n 3,   tɔ:n ɪ re:k 3
go   FP   you   first

---

4. /raw ɪ/ is replaced by /khaw 4/ (originally 'he') among some young speakers.

maj 3   saːp 3   ləːj 1   klua 1   waː 3
not   know   at all   afraid   say

caʔ 3   paj 1   keʔ 2 kaʔ 2
will   go   in other's way

'I am coming sir. I didn't know before and thought I would be in your way.'

Second person, plural number

(8) tiw 5   tuj 5   tɔj 3   sənuk 2 kan 1   yaj 2   siʔ 1   **raw 1**
    Tiu   Tui   Toy   have fun   big   FP   we

'Tiu, Tuy Toy : Having a very good time, aren't you?'

The last two examples show /raw 1/ as a second person pronoun, either singular or plural. This /raw 1/ is used by a speaker who is more senior or higher in position to a younger addressee. Thus, when it is misused with equals, the pronoun can cause ill feeling toward the speaker.

The various semantic usages of /raw 1/ above can be put into two main groups. The first group, as in sentences (1), (5), and (6), shows a feature of inclusion or intimacy. Sentences (2), (3), and (4) share the meaning of exclusion or detachment and form the second group. Intimacy and detachment are on opposite poles, as can be seen from figure 1. It is thus a question why /raw 1/ as a first person pronoun can have such contradictory features.

Figure 1. Usage of /raw 1/ as a first person singular pronoun

/raw 1/ (5)⌐   ⌐/raw 1/ (3)
INTIMATE   DETACHED
◄————————/raw 1/ (1) : /raw 1/ (2)————————►
/raw 1/ (6)⌐   └/raw 1/ (4)
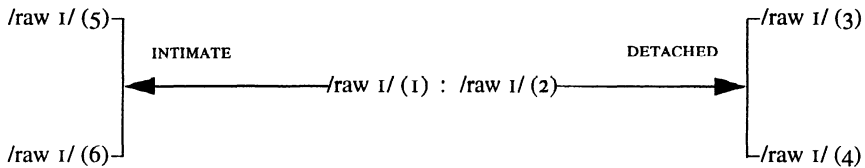
3. Proposed Explanation

The occurrences of /raw 1/ as a first person pronoun with a semantic feature 'intimate' on the one end and a 'detached' semantic feature on the other can be explained if we look at a personal pronoun system in the early Ayudhya period and a Tai Neua personal pronoun system at the present time.

A study of personal pronouns in a classical Thai epic "Phra Lore," believed to have been written about 500 years ago, reveals that for the first person pronoun, there are the forms given in table 3. Following the table are examples.[5]

_____

5. Only examples of /raw/ are given. Tones are not marked.

Table 3. First Person Pronouns in "Phra Lore"

SINGULAR	DUAL	PLURAL
raw	raː (INCLUSIVE)	raw (INCLUSIVE)
khaː	phɯa (EXCLUSIVE)	tuː (EXCLUSIVE)
kuː	khɯa kha, khɔj, phɯa khaː	
'I'	'we two'	'we'

raw (singular)

(9) อย่าช้าเราจะรบ

jaː	chaː	**raw**	caʔ	rop
not	slow	we	will	fight

(The king to his followers:) 'Be quick ! I want to fight the battle.'

(10) เรานี้เราเทพเจ้า จอมผา ไสร้นา

raw	niː	**raw**	theːp pha caw	cɔːm phaː	saj naː
we	this	we	deity	mountain	FP

เขาใส่สม-าเรา ปู่เจ้า

khaw	saj	somyaː	**raw**	puː caw
they	put	name	we	Puː Caw

'I am the deity of this mountain. They call me "Pu Caw."'

raw (plural)

(11) หลานเอยคอยพยายาม ฤๅรอดเราเลย

laːn	ʔəːj	khɔj	phayayaːm	rɯː	rɔːt	**raw**	ləːj
granddaughter	FP	slowly	try	Q	escape	we	FP

'My granddaughter, if you keep trying, (the hero) cannot escape from us.'

As can be seen from examples (9) and (10), /raw/ is used as a first-person pronoun by a royal figure or an equivalent. This usage thus indicates power. In example (11), /raw/ is a plural personal pronoun and as shown in table 3, it is also an inclusive form in contrast to /tu/, which is an exclusive pronoun. The use of /raw/ as a plural pronoun is probably the base form for the derived singular form /raw/.

Now if we look at the personal pronoun system of Tai Neua as shown in table 1 above, we can see that /hau 55/ 'inclusive first-person plural pronoun' can be compared with /raw/ as inclusive first-person plural pronoun in Phra Lore. There is similarly a distinction between /hau 55/ and /tu 33/ as inclusive and exclusive pronouns. Because in the Tai Neua system there is no occurrence of /hau 55/ as a singular pro-

noun, it is most probable that the usage of /raw/ as a singular first-person pronoun in the Thai epic Phra Lore is a derived form.

It is the features [+ plural] and [+inclusive] in the plural form of the first person pronoun both in Phra Lore and Tai Neua that are the keys to the contradictory usage of /raw ɪ/ in Bangkok Thai.

The semantic change from 'plurality' in the base form plural pronoun /raw/ is not unusual. Brown and Gilman (1970) cite the use of Royal 'I' or 'we' in English and point out that "plurality is a very old and ubiquitous metaphor for power." Thus, the occurrence of /raw ɪ/ in Bangkok Thai as a first-person pronoun in singular number (sentence [3]) with a [+ power] feature can now be explained. It can then be stated why the [+ detached] semantic feature has become a semantic feature of /raw ɪ/ in Bangkok Thai, as in example (3) : Power naturally puts a person in isolation. It is apparent here that when the pronoun has derived the [+ power, + detached] features, the [+ detached] feature normally excludes the feature [+ inclusive].

If we take a look at /raw ɪ/ in the Bangkok Thai example (6), however, it is clear that the [+ inclusive] feature is chosen when the plural /raw ɪ/ is transferred to the singular form /raw ɪ/. Intimacy implies inclusion.

The occurrence of /raw ɪ/ in (4) and (5) reveals traces of both plurality and inclusiveness simultaneously, although one feature or the other is more dominant in each instance. In (4) [+ detached], a faint trace of the "power" of plurality is dominant, and yet, as has been mentioned before, [+ detached] can bring in a paradoxical [+ inclusive] through objectivity and generality. In (5), [+ inclusive] is dominant, because the speaker intends to emphasize a closer relationship. At the same time, there is a faint trace of [+ power] in an attempt to show equality of sex by a female speaker. This strange combination of the [+ inclusive], [+ detached], and [+ power] features is also present in the use of /raw ɪ/ as a second pronoun in examples (7) and (8), where the speaker is more senior or higher in position than the younger addressee.

This study of /raw ɪ/ shows that the contradictory usage of an item in a present-day language may be explained through historical study and by comparison with a contemporary sister language. It also shows that semantic features, when chosen separately historically, can develop into contradictory features at a later stage.

REFERENCES

Anonymous. (1448-1533). *Phra Lore* (in Thai). Bangkok: Khurusapha Press.
Brown, Roger, and Albert Gilman. 1970. The pronouns of power and solidarity. In *Readings in the sociology of language*, ed. by Joshua A. Fishman, 252–275. The Hague: Mouton.
Kullavanijaya, Pranee. Forthcoming. A sketch of Tai Neua (in Thai).
Palakornkul, Angkab. 1972. A sociolinguistic study of pronominal strategy in Spoken Bangkok Thai. Ph.D. dissertation, University of Texas at Austin.

Part 2
Morphology, Syntax, Semantics

7: SOME ASPECTS OF PAZEH SYNTAX

Paul Jen-kuei Li
Institute of Linguistics, Academia Sinica

1. Introduction[1]

The Pazeh language became nearly extinct by the end of World War II. It ceased to be actively spoken, although it still existed in the memory of a few older people a few decades later. Aside from wordlists collected by Ferrell (1970), Tsuchida (1969a), and Li (1976), there is little information on its syntax. Both Tsuchida and I managed to collect only a few texts, still unpublished manuscripts. Li (1978) gives a preliminary description of its case-marking system. Then all our informants passed away. I thought Pazeh had become a dead language at that point, and that it would not be possible to get anything further of substance on the language from that time on.

Then in February 1997 I started to work with an 83-year-old woman named Jin-yu Pan, whose knowledge of Pazeh is more satisfactory than expected, especially at this stage. I have been checking through Tsuchida's relatively long wordlist and trying to work out a sketch of Pazeh grammar, including its aspect, focus, and case-marking system. I found a few new things. For one thing, I found a new set of locative pronominal forms, such as *yakuan* and *imuan*, which had been overlooked in my (Li 1978) earlier report. For another, I found Locative-focus verbs ending in *-an* and Instrumental-focus verbs beginning with *sa-* ~ *saa-*. Again, these features were not treated or even mentioned in any of the previous studies of the language.

An aim of this essay is to point out syntactic similarities and differences between Pazeh and the other Formosan languages, and to determine, if possible, to which Formosan language Pazeh is most closely related.

2. The Focus System of Pazeh

Like many other Formosan and Western Austronesian languages, Pazeh has four types of focus indicated by verb affixes: (1) Agent-focus (AF)[2] by *mu-, me-, mi-, m-, ma-* or zero; (2) Patient-focus (PF) by *-en* or *-un*, (3) Locative-focus (LF) by *-an*, and (4) Referential-focus (RF) by *sa-, saa-* or *si-*.

2.1 Agent-focus

Verbs with the AF affixes indicate an action or event that has taken place or is in the process of taking place. Different types of AF-marking affixes are illustrated in (1–6).

(1)   mu-baket rakihan ki   aba
      AF-beat child   NOM Dad
      'The father beat/is beating a child.'

(2)   yaku ka   ma-baza imisiw   ukuazixa mu-puzah
      I.NEU TOP AF-know he.NEU   yesterday AF-come
      'I know he came yesterday.'

(3)   mi-kita akim ba-baket   rakihan ki   abuk
      AF-see Akim PROG-beat child   NOM Abuk
      'Abuk saw Akim beating a child.'

(4)   aba ka   kasibat   rakihan mu-depex babizu
      Dad TOP AF-teach   child   AF-read book
      'The father teaches a child to read books.'

(5)   me-ken asay pai siw
      AF-eat what Q   you.NOM
      'What do you eat?'

(6)   m-ituku aku   lia
      AF-sit I.NOM ASP
      'I have already sat down.'

Verbs with the AF affixes appear not only in finite but also in nonfinite forms, as in sentence (4). Sequential verbs must agree in focus (indicated by italics), as in (7–10).

(7)   kuang a   udal ka,   rakihan *mu-kusa ma-hadas*
      not.have LIG rain TOP   child   AF-go   AF-play
      'When there is no rain, the child will go to play.'

(8)   *ma-baza   mu-kawas* pazih a   rahan siw
      AF-know   AF-speak Pazeh LIG   words you.NOM
      'Do you know how to speak Pazeh?'

(9)   *mu-repun me-ken* dadas
      AF-finish AF-eat potato
      'He finished eating sweet potatoes.'

---

2. Abbreviations as used in this essay are as follows: AF, Agent-focus; ANT, anticipating; ASP, aspect; BF, Beneficiary-focus; CAUS, causative; DIST, distant; EXCL, exclusive; FUT, future; GEN, genitive; IF, Instrumental-focus; INCEP, inceptive; INCL, inclusive; IMP, imperative; INVIS, invisible; IRR, irrealis; LF, Locative-focus; LIG, ligature; LOC, locative; NEUT, neutral; NOM, nominative; OBL, oblique; PERF, perfective; PF, Patient-focus; PL, plural; PROG, progressive; PROX, proximate; Q, question particle; RED, reduplication; RF, Referential-focus; SG, singular; TOP, topic.

(10) a. yaku *m-ukukusa mu-ruput* lia
 I.TOP AF-work AF-finish ASP
 'I have finished the work.'

 b. *ukukusa-en rupud-en* lia
 work-PF finish-PF ASP
 'The work has been finished.'

 c. *yaku m-ukukusa rupud-en lia
 AF PF

Pazeh verbs marked with *m-* appear even in irrealis. Such a phenomenon is similar to most Formosan languages such as Atayal (Huang 1995:191) and Seediq (Holmer 1996:60), and to other Western Austronesian languages, but different from Rukai (Li 1973:223). For example,

(11) ana mi-talam
 don't AF-run
 'Don't run!'

(12) usa m-ara dalum
 go AF-take water
 'Go take some water!'

(13) isiw m-apa rakihan ka, yaku m-apa'-ay rakihan
 you.NEUT AF-carry child TOP I.NEUT AF-carry-IRR child
 'If you carry a child on your back, I'll do so, too.'

(14) ma-isat baxala siw ka, m<a>i-kixi-ay nimisiw a ina
 AF-scold Baxala you.NOM TOP AF-PROG-tell-FUT his LIG Mom
 'If you scold Baxala, I'll tell his mother.'

Nevertheless, verbs in the main clause and embedded clause do not need to agree in focus. For example,

(15) aba mu-daxal dadas zazing-en
 Dad AF-dig potato throw.away-PF
 'The sweet potato Father dug up was thrown away.'

2.2 Patient-focus

Patient-focus is indicated by the suffix *-en* or *-un* (conditioned by the preceding vowel *u*, as in (24). It generally refers to a present or past action or event, but not to the future, as in several other Formosan languages such as Atayal (Huang 1995:45) and Seediq.

(16) naki a rakihan rapun-en ni hakezeng a saw
 I.GEN LIG child take.care-PF GEN old LIG person
 'My child is/was taken care of by an old person.'

(17) xaringa'-en naki
play-PF I.GEN
'I've played it.'

(18) baked-en ni sabung rakihan ka, ma-raxiw lia
beat-PF GEN Sabung child TOP AF-escape ASP
'The child beaten by Sabung has escaped.'

(19) baxa'-en rakihan lia ki paray
give-PF child ASP NOM money
'The money has been given to a child.'

(20) rakihan rabex ka iba'-en ni ina
child little TOP hold-PF GEN Mom
'The little child is/was held by the mother in her arms.'

(21) nuang xe'ed-en naki lia
cow tied-PF I.GEN ASP
'The cow has been tied by me.'

(22) ngazib-en wazu lia ki rakihan
bite-PF dog ASP NOM child
'The child has been bitten by a dog.'

(23) hakezeng a saw pa-'anid-en ki rakihan
old LIG person CAUS-cry-PF NOM child
'The old person made the child cry.'

(24) huruhur-un dua dini lai³ ki adadumud a luxud a isia
drag-PF there here ASP NOM one LIG deer LIG that
'That one deer was dragged over there and over here.'

(25) xakela'-en ka anu⁴ say
think-LF TOP for what
'What use is there if you just think about it?'

In most Formosan and Western Austronesian languages, the perfective form of the Patient-focus does not bear the focus suffix -en, but the infix -in-. Pazeh may bear either -in- (e.g., b<in>aket 'to have been beaten') or -en (e.g., baked-en 'to be beaten or to have been beaten') or both affixes together (e.g., b<in>aked-en 'to have been beaten') to indicate the perfective of the Patient-focus (see section 6).⁵

The future in PF constructions is indicated by reduplication of the verb stem or insertion of the infix-a- after the initial consonant of the verb stem (see section 5.3). A few examples showing a past and a future action or event in Patient-focus are given in (26–29) for comparison.

---

3. The form *lai* is a variant of *lia.*

4. The form *anu* is a variant of *aunu* 'for (the sake of)'.

5. I would like to thank Videa De Guzman (pers. comm.) for reminding me of this important difference between Pazeh and the other languages.

(26) a. xe'ed-en naki ki   nuang
     tie-PF   I.GEN NOM COW
     'The cow was tied by me.'

    b. xaa-xe'ed-en naki   ki   nuang
      RED-tie-PF   I.GEN NOM COW
      'The cow will be tied by me.'

(27) a. baked-en ni   sabung ki   rakihan
     beat-PF   GEN Sabung NOM child
     'The child was beaten by Sabung.'

    b. baa-baked-en ni   sabung ki   rakihan
      RED-beat-PF   GEN Sabung NOM child
      'The child will be beaten by Sabung.'

(28) a. rakihan rabex iba'-en   ni   ina
     child   baby   hold-PF GEN Mom
     'The baby was held by Mother.'

    b. rakihan rabex a-'iba'-en   ni   ina
      child   baby   RED-hold-PF GEN Mom
      'The baby will be held by Mother.'

(29) a. xaringa'-en naki
     play-PF   I.GEN
     'It was played by me.'

    b. x<a>aringa'-en naki
      FUT-play-PF   I.GEN
      'It will be played by me.'

    c. xa-xaringa'-en naki
      RED-play-PF   I.GEN
      'It will be played by me continuously.'

Note the difference between reduplication of the verb stem and insertion of-*a*- when both can occur with the same verb stem, as illustrated in (29b,c) above.

2.3 Locative-focus

As in the other Formosan languages, the Locative-focus is indicated by the verbal suffix -*an*. It normally refers to a past action or event, but not always. For example,

(30)   xutaxa'-an naki, ini mu-puzah siw
        wait-LF   I.GEN not AF-come you.NOM
        'I waited (for you), but you didn't come.'

(31)   imini ka   xizib-an lia
        these TOP cut-LF   ASP
        'These have been cut.'

As Locative-focus generally refers to a past action or event, there is often a combination of -an and -in- 'perfective' in verbs. For example,

(32) k<in>exed-an ni abua naki a rima
    PERF-cut-LF GEN Abua my GEN hand
    'My hand was cut by Abua.'

(33) x<in>e'ed-an ni awi lia ki wazu
    PERF-tie-LF GEN Awi ASP NOM dog
    'The dog was tied by Awi.'

(34) '<in>angid-an ka mairad-ay inang say
    PERF-cry-LF TOP alive-FUT again Q
    'Will crying help to make (him) live again?'

(35) x<in>alid-an nuang matla dini mu-razaw di daran
    PERF-lead-LF cow from here AF-pass LOC road
    'The cow was led from here to pass through the road.'

(36) x<in>u'ud-an naki ki xalam
    PERF-pull-LF by.me NOM vegetable
    'I've pulled the vegetables.'

(37) x<in>azing-an ni saw ki dadas
    PERF-throw-LF GEN person NOM potato
    'People threw away the sweet potatoes.'

2.4 Referential-focus

Referential-focus, including Instrumental-focus (IF) and Beneficiary-focus (BF), is indicated by si- in most Formosan languages, by ši- in Thao, by sa- in Amis, by saa- in Pazeh and Saaroa, by 'is- in Bunun, and by ti- in Kavalan (Li 1996b:73–76). Compare the following sentences showing the different focus forms of the same verb.

(38) a. mu-xe'et nuang ki yaku
    AF-tie cow NOM I.NEUT
    'I tied a cow.'

  b. xe'ed-en naki lia ki nuang
    tie-PF I.GEN ASP NOM cow
    'The cow has been tied by me.'

  c. saa-xe'et nuang ki kahuy
    IF-tie cow NOM tree
    'A cow was tied to the tree.'

Of all the Pazeh data I have collected, very few examples contain RF main verbs, as in (39).

(39) saa-talek alaw ki bulayan
    IF-cook fish NOM pan
    'Fish was cooked in the pan.'

Verbs with the prefix *saa-* indicating IF or BF appear more often in embedded verbs, as in:

(40)　kalu imini kabat saa-bazu syatu　ki　　mamais a　　saw
　　　use　this　board IF-wash clothes NOM female LIG person
　　　'The woman used this board to wash clothes.'

(41)　ina　　kalu bulayan mu-talek alaw saa-pakan rakihan
　　　Mom use　pan　　AF-cook fish IF-feed　　child
　　　'Mother used a pan to cook fish to feed a child.'

(42)　ina　　mu-sais syatu　anu sabung　saa-pasyatu
　　　Mom AF-sew clothes for　Sabung IF-wear.clothes
　　　'Mother sewed clothes for Sabung to wear.'

My Pazeh informant tends to use the verbs *kalu* and *aunu* to indicate Instrument and Beneficiary, respectively, as in the following examples. (This may be due to the influence of Southern Min, the language of her daily use.)

(43)　ina　　kalu saapa[6] m-apa　　rakihan
　　　Mom use　strap AF-carry child
　　　'Mother uses straps to carry a child on her back.'

(44)　aba mu-tahan paray　aunu rakihan
　　　Dad AF-make money for　child
　　　'Father made money for the child.'

I found only one example for a verb with the prefix *si-* indicating Instrumental-focus:

(45)　si-te'eng wazu ni　　rakihan ki　　batu
　　　IF-throw　dog　GEN child　　NOM stone
　　　'The stone was used to throw at a dog by the child.'

The difference between *saa-* and *si-* is not clear. According to my informant, *saa-te'eng* would sound better than *si-te'eng* in (45) above, and the latter is rarely used. However, there is a form *si-te'eng-en* 'to have been thrown'.

3. Case Markers

Pazeh has the following four case markers:

　　　ki　Nominative　　　　　　　ni　Genitive
　　　di　Locative　　　　　　　　u　Oblique

Most of these case markers are optional. Unlike many other Formosan (see Li 1997a) and Western Austronesian languages, Pazeh does not distinguish between personal names and common nouns in terms of case marking. The nominative marker *ki* may

---

6. As pointed out by Tsuchida (pers. comm. 1998), the form *saapa* is derived from *saa-apa*. One of the vowels is deleted, as Pazeh does not allow sequences of three identical vowels.

precede either a common noun or a personal name. It may even occur before a nominative personal pronoun, as in (38a) above. The genitive marker *ni* may also precede a common noun or a personal name. Like most other Formosan languages, Pazeh's genitive marker *ni* indicates not only possession but also the agent in a non-Agent focus construction. The locative marker *di* indicates a location or even a goal. The oblique marker *u* indicates an object or anything other than subject, possession, or location. These markers are illustrated in (46–54).

(46) aba   ka   kasibat rakihan mu-depex babizu
      Dad   TOP   teach   child   AF-read book
      'Father teaches a child to read books.'

(47) aba' a   paray ka   baya-en di   rakihan lia
      Dad LIG   money TOP   give-PF LOC   child   ASP
      'Father's money was given to a child.'

(48) mu-puzah ki   awi, yaku   (ka)   kaa-ken sumay
      AF-come NOM   Awi   I.NEUT TOP   RED-eat rice
      'When Awi came, I was eating rice.'

(49) ma-baza   mu-languy ki   rakihan
      AF-know   AF-swim   NOM   child
      'The child knows how to swim.'

(50) tahayak lia   ki   aba, duyla m-idem
      tired   ASP   NOM   Dad   go   AF-sleep
      'When Father was tired, he went to bed.'

(51) imini a   syatu ka   ni   rakihan
      this LIG   clothes TOP GEN   child
      'These clothes are the child's / belong to the child.'

(52) pa-daux   inusat ni   kalagu ki   atun
      CAUS-drink   wine   GEN   Kalagu NOM   Atun
      'Kalagu made Atun drink.'

(53) ma-bazu syatu   di   xuma ki   mamais
      AF-wash clothes LOC   house NOM   female
      'The woman washes clothes at home.'

(54) isiw ka   taa-tahay u   baruzak
      you TOP   PROG-kill OBL   pig
      'Are you(SG) killing a pig?'

In addition to the four case markers listed above, there is a topic marker *ka* and a ligature *a* in Pazeh. As usual, the topic appears in clause-initial position. The topic marker *ka* is optional, but there is always a pause after the topic. The ligature *a* simply links two nouns. While the nominative marker *ki* indicates definiteness of the following noun, the ligature *a* indicates a subordinate relation between the two nouns. Compare the two utterances in (55).

(55) a. ni   taruat ki   babizu
      GEN Taruat NOM book
      'The book is Taruat's.'

    b. ni   taruat a   babizu
      GEN Taruat LIG book
      'Taruat's book'

4. Personal Pronouns

Pazeh has four sets of personal pronouns, as shown in table 1. Most of the pronominal forms in Pazeh are derived from PAN; there is only a minor change in the second person singular form: *Su > siw.[7] All the pronominal forms are free. There is only one set of short forms, the nominative, which occurs freely in various positions in a sentence. However, it appears more frequently in the second position and right after the main verb (see examples below). Aside from the short nominative forms, a prefix ta- (an abbreviation of the first person inclusive genitive nita) attaches to the Patient-focus form of the main verb in a polite request, e.g., ta-kan-i 'Let's eat!' , ta-daux-i 'Let's drink!' The third

Table 1. Personal Pronouns

		NEUTRAL	NOMINATIVE	GENITIVE	LOCATIVE
1SG		yaku	aku	naki	yakuan, yakunan
2SG		isiw	siw	nisiw	isiwan
	PROX	imini	mini	nimini	iminiyan
3SG	DIST	imisiw	misiw	nimisiw	misiwan
	INVIS	isia	sia	nisia	isiaan
1INCL		ita	ta	nita (ta-)	itaan
1EXCL		yami	ami	nyam(i)	yamian, yaminan
2PL		imu	mu	nimu	imuan
	PROX	yamini	amini	naamini	yaminiyan
3PL	DIST	yamisiw	amisiw	naamisiw	yamisiwan
	INVIS	yasia	asia	naasia	yasiaan

person pronouns are derived from demonstratives, as in Rukai (Li 1973:75). There are three subsets for the third person. They distinguish between 'visible' and 'invisible', and 'proximate' and 'distant' for the visible. Plurality of the third-person forms is indicated by the prefix a-, similar to Amis and Puyuma (Li 1997a).

---

7. The form imisiw 'he, she' shows up as imisu in formal ritual songs of Pazeh. We may infer that the change *Su > siw took place fairly recently.

97

4.1 The Neutral and Nominative Pronouns

The set of neutral forms has various functions, namely, topic, subject, and object. These neutral forms may alternate with the nominative (short) forms.

(56)   yaku   ka   mu-ngazip rumut
         I.NEUT TOP AF-bite   meat
         'I bit meat.'

(57)   pazih yaku ~ aku
         Pazeh I.NEUT ~ NOM
         'I am a Pazeh.'

(58)   mu-ngazip yaku   ki   wazu
         AF-bite   I.NEUT NOM dog
         'The dog bit me.'

(59)   m-ituku isiw ~ siw
         AF-sit   you.NEUT ~ NOM
         'You (SG) are sitting.'

Further examples of the short nominative forms are as follows:

(60)   mu-daux aku   dalum
         AF-drink I.NOM water
         'I drink water.'

(61)   mu-suzuk naki a   paray siw
         AF-hide I.GEN LIG money you.NOM
         'Did you hide my money?'

(62)   ma-baza mu-tu'ul   lalawa siw
         AF-know AF-weave cloth   you.NOM
         'Do you know how to weave cloth?'

Notice that the short form *siw* occurs in sentence-final position in (61) and (62). That indicates that it is not bound to the preceding verb or noun predicate. It can be treated as a clitic. There is greater restriction for the short forms than for the long ones. Except for *ta-* as mentioned above, the short forms may not occur in sentence- or clause-initial position, as in (63b):

(63) a. yaku ki-kita   wazu.
          I.TOP PROG-look dog
          'I am looking at a dog'

      b. *aku ki-kita wazu

As a matter of fact, the long and short forms of the same personal pronoun may cooccur: the former as the topic and the latter as the subject of a clause, as in (64a).

(64) a. yaku ma-baza **isiw**   (*ka) uhuni a   dali (ka) nahani **siw**
          I.TOP AF-know you.TOP   now LIG day TOP come   you.NOM
          'I know that you will come today.'

    b. *yaku ma-baza **siw** uhuni a dali nahani **isiw**
       I.TOP AF-know you.TOP now LIG day come you.NOM
       'I know that you will come tomorrow.'

As stated above, the short form of a pronoun may not occur in the topic or preverbal position. If the positions of the long and short forms of a pronoun are interchanged, the result is an ungrammatical sentence, as in (64b).

    In an interrogative sentence, the interrogative marker *pai* appears right after an interrogative word and is followed by a noun phrase, including a pronoun or a noun and a demonstrative. For example,

(65) a. m-usay asay pai isiw
          AF-go where Q you.NEUT
          'Where are you going?'

       b. m-usay asay pai ki awi
          AF-go where Q NOM Awi
          'Where is Awi going?'

(66)   hapet mu-'asay pai isiw
          want AF-what Q you.NEUT
          'What do you want?'

(67)   ima pai imisiw
          who Q he.NEUT
          'Who is he?'

(68)   asay pai ini
          what Q this
          'What is this?'

4.2 The Genitive Pronouns

Genitive pronouns are usually short and are attached to the preceding verb or noun in many Formosan languages, such as Atayal, Bunun (Li 1997b), Rukai (Li 1973, 1996a), and Kavalan (Li 1996b). However, there are no such short genitive forms in Pazeh. Moreover, perhaps due to the influence of Taiwanese (Southern Min), the genitive forms usually precede the nouns they modify, with the ligature *a* in between, as in (69–73).

(69)   usa m-ara dalum aunu nisiw a ina
          go AF-take water for you.GEN LIG Mom
          'Go and get some water for your mother.'

(70)   naki a 'yah m<a>w-puzah-ay lia
          I.GEN LIG elder AF-PROG-come-IRR ASP
          'My elder sister will come soon.'

(71)   imini a alaw ka nimu (a) k<in>umux
          this LIG fish TOP you.GEN LIG PERF-catch
          'These[8] fish were caught by you(PL).'

(72) imini a   tadaw ka nimisiw (a) z<in>upay
      this   LIG knife TOP he.GEN LIG PERF-sharpen
      'This knife was sharpened by him.'

(73) ayan ka   nyam   (a) p<in>aturay.
      Ayan TOP we.GEN LIG PERF-sing
      'Ayan is what we sang.'

The ligature *a* is optional between a genitive and a verb, as in (71–73). Note, however, that the genitive pronouns are the agents of the following PF verbs. The genitive forms may also occur after the verbs they modify, and in this case, *a* does not occur. For example, (74) is an alternative statement for (71).

(74) imini alaw ka kinumux nimu

Like most other Formosan languages, a genitive form may also indicate the agent in a non-Agent-focus construction. In this case, it follows the main verb, as in (75–77).

(75) x<in>u'ud-an naki ki   xalam
      PERF-pull-LF I.GEN NOM vegetable
      'I pulled the vegetables.'

(76) ta-baked-aw (nita)   ki   balan
      we-beat-FUT we.GEN   NOM   cat
      'Let's beat the cat!'

(77) ta-kan-i   (nita)   ki   alaw
      we-eat-IMP   we.GEN   NOM   fish
      'Let's eat the fish!'

It is interesting to note that both genitive forms *ta-* and *nita* may cooccur in the same sentence, as in (76) and (77) above. The long form is optional, as indicated by the parentheses. The form *ta-*, as mentioned earlier, is the only short genitive form among the personal pronouns. It occurs as a preverbal (PF) person marker.

The long genitive form may also function as predicate, as in (78).

(78) imini rakihan ka   naki
      this child   TOP I.GEN
      'This child is mine.'

4.3 The Locative Pronouns

Like most other Formosan languages, Pazeh has a set of locative pronouns with the suffix *-an*. This was overlooked in my earlier report for Pazeh (Li 1978). A locative pronoun is optionally preceded by the locative marker *di*, as in (79–82).

---

8. Although the form *imini* indicates singular, the phrase *imini a alaw* may mean 'this plate/ basket of fish' and indicates a plurality of fish.

(79)   naki   a   aba mausay mikiliw saw   puzah (di) yamian
            I.GEN LIG Dad will.go call   people come LOC we.LOC
            'My father is going to ask people to come to our house.'

(80)   alu   (di) yakuan malaleng
            come LOC I.LOC   live
            'Come to stay with me.'

(81)   yaku mausay (di) isiwan   mituku
            I.TOP will.go LOC you.LOC sit
            'I'll go and sit in your house = I shall visit you.'

(82)   imisiw mausay (di) ita'an   mituku
            he.TOP will.go LOC we.LOC sit
            'He will go to our house to sit = He will visit us.'

Two locative pronouns have free variants: *yakuan ~ yakunan* 'I, locative' and *yamian ~ yaminan* 'we (EXCL), locative'.

5. Aspect

In terms of form, there are three major types of aspect markers in Pazeh: (1) affixation, *-in-* 'perfective', *-a-* 'progressive', *-ay* 'AF, irrealis', *-aw* 'PF, irrealis'; (2) reduplication of the first syllable of the verb stem to indicate 'progressive'; and (3) the grammatical particle *lia* 'already'.

5.1 The Perfective *-in-*

Like the other Austronesian languages, the perfective aspect in Pazeh is indicated by the infix *-in-* inserted after the initial consonant of the verb stem. It indicates an action or event that has taken place, as in (83–87).

(83)   m<in>e-ken siw   sumay lia
            AF-PERF-eat you.NOM rice   ASP
            'Have you eaten rice?'

(84)   awi ka p<in>a-kan   durun ni   tata
            Awi TOP CAUS-PERF-eat-PF bran   GEN step-mom
            'Awi has been fed with rice bran by the stepmother.'

(85)   mamah ka uzay nisiw   p<in>arisan   aku
            elder   TOP not   you.GEN PERF-born-PF I.NOM
            '(My) elder (brother) said I had not been borne by you.'

(86)   nisiw   b<in>axa   a   arim ka   riak a   ka-kan-en.
            you.GEN PERF-give-PF LIG peach TOP good LIG RED-eat-PF
            'The peach given by you is good to eat.'

(87)   x<in>areb-an[9] ki   xuma
            PERF-burn-LF NOM house
            'The house has been burned.'

101

The perfective marker *-in-* cooccurs with the AF verbs (as in [83]), the PF verbs (as in [84–86]), and the LF verbs (as in [87]). However, there is no combination of *-in-* and *-en*, as exemplified in the verb forms above. In other words, the PF affix is zero[10] when the perfective infix *-in-* appears, just as in most other Formosan and Western Austronesian languages (see Ross 1995).

Although Agent-focus markers such as *mu-* may indicate a past event, their combination with the perfective marker *-in-* ensures that an event has actually taken place. Compare the pairs of sentences in (88–89).

(88) a. aba   ka   paxarihan mu-puzah   dini
     Dad TOP forget   AF-come   here
     'Father forgot to come here.'

   b. aba   ka   paxarihan m<in>u-puzah dini
      Dad TOP forget   AF-PERF-come here
      'Father forgot that he had come here (before).'

(89) a. atun   ka   paxarihan mu-baket   rakihan
     Atun TOP forget   AF-beat   child
     'Atun forgot to beat a child.'

   b. atun   ka   paxarihan m<in>u-baket rakihan
      Atun TOP forgot   AF-PERF-beat child
      'Atun forgot that he had beaten a child.'

5.2 The Inception Marker *lia*

The inception marker *lia* 'already', with a grammatical function similar to Mandarin *le*, indicates an action or event that has started or is about to take place. For example,

(90)   mangit lia   ki   rakihan
     AF-cry ASP NOM child
     'The child has cried.'

(91)   yaku m-uku-kusa   mu-ruput lia
     I.TOP AF-RED-work AF-finish ASP
     'I have finished working.'

(92)   mausay lia   aku
     will.go ASP I.NOM
     'I shall go. / I am about to leave.'

The marker cooccurs with various forms of the verb, including AF, PF, LF, and even verb forms that indicate the future, as in these examples.

---

9. As observed by Videa De Guzman (pers. comm.), this verb is marked LF when semantically it is an object that has been burned. In Tagalog, there are verbs that also take *-an*, but are considered PF, such as *labh-an* 'launder', *punas-an* 'wipe off (a surface)', *tikm-an* 'taste s.t.'.

10. As noted above, my informant occasionally produced a form with both *-in-* and *-en*, such as *b<in>aked-en* 'to have been beaten' (see section 2.2).

5.3 The Progressive Marker

Progressive aspect is indicated by reduplication of the first syllable of the verb stem, with the vowel replaced by *a* and lengthened.[11] For example (all the following verbs are in AF),

(93)   yaku kaa-ken dukul
          I.TOP RED-eat taro
          'I am eating taro.'

(94)   yaku daa-daux dalum
          I.TOP RED-drink water
          'I am drinking water.'

(95)   ina   ka baa-bazu syatu
          Mom TOP RED-wash clothes
          'Mother is washing clothes.'

Another way of indicating the progressive aspect is through the insertion of *-a-* after the initial consonant of the verb stem, as in (96–97).

(96)   m-a-idem   lia ki   rakihan
          AF-PROG-sleep ASP NOM child
          'The child has been sleeping.'

(97)   abuk t<a>umala abua m<a>a-turay
          Abuk PROG-hear Abua AF-PROG-sing
          'Abuk is listening to Abua singing.'

5.4 The Irrealis Markers *-ay* and *-aw*

Both irrealis suffixes *-ay* and *-aw* indicate the future. While the former is affixed to Agent-focus verbs, the latter cooccurs with Patient-focus ones (see section 6, and see also Ross 1995:739).

(98)   m-apa'-ay   rakihan ki   kayu
          AF-carry-IRR child   NOM Kayu
          'Kayu will carry a child on his back.'

(99)   parazem a   isiw,   ta-kita'-aw
          new.year LIG you.NOM we-see-IRR
          'During your new year, let's see each other.'

---

11. Tsuchida (pers. comm.) has pointed out to me that "the progressive marker is a long redu-
plication CVV- of the first vowel of the stem if the stem is disyllabic and takes *mu-* or *me-* in
AF neutral, e.g., *mu-baxa* 'give', *baa-baxa* 'is giving'. It is infix *-a-* if the stem [that] takes
mu- or *me-* consists of more than three syllables, or with the stems [that] do not take *mu-* or
me- in AF neutral, e.g., *mu-dakedak* 'thresh', *d<a>akedak* 'is threshing'; *paharas* 'lie
down', *p<a>aharas* 'is lying down'. But there are some exceptions."

(100)   yamisiw dadua nahani-ay
           they   all   come-IRR
           'They shall all come.'

(101)   '<in>angid-an ka   mairad-ay inang say
           PERF-cry-LF   TOP alive-IRR   again   Q
           'Will crying make him live again?'

5.5 Various Combinations of the Aspect Markers

It is possible to combine the aspect markers in different ways. One of them is a combination of reduplication, marking progressive, and the irrealis marker *-ay*, as in (102–104).

(102)   sarawan xe-xe'ed-ay   lia ki   wazu
           tomorrow PROG-fasten-IRR ASP NOM dog
           'The dog will be fastened tomorrow.'

(103)   yaku ka   ba-baked-ay   ma-sakaw me-ken alaw a   balan
           I   TOP PROG-beat-IRR AF-steal   AF-eat   fish   LIG   cat
           'I shall beat a cat that steals fish to eat.'

(104)   ina   ka   ma-baza   isiw   ka-kan-ay   alaw
           Mom TOP   AF-know you   PROG-eat-IRR fish
           'Mother knows that you will eat fish.'

Another combination is that of insertion of the infix *-a-* 'progressive' plus the irrealis marker *-ay*, e.g., *m-a-usay > m-a-usa-ay* 'to be going, to be leaving'.

(105)   m-a-usa-ay   lia aku
           AF-PROG-go-IRR   ASP I.NOM
           'I shall be leaving soon.'

Double marking of a verb stem with the progressive aspect, such as reduplication plus an affix (*-a-* or *-ay*), carries a more emphatic meaning in Pazeh. It is not so emphatic when there is only one of the markers. It is not clear if it is possible to mark all verbs in this combined way. This requires further investigation.

6. Focus and Aspect

In discussing PAN verbal morphology, Ross (1995) uses the terms "indicative" and "nonindicative" to separate the two major types of mood in Austronesian languages. Following the tradition of Formosan linguistics, I shall use the terms "realis" and "irrealis" to stand for the dichotomy of mood, as in table 2 below. Some empty slots for LF and RF, especially those in the irrealis mood, can be filled when more information becomes available.

Correlation of the focus system and the aspect system in Pazeh can be exemplified with the verb root *baket* 'to beat', as shown in table 2 and in (106–119) (see Zeitoun et al. 1996).

Table 2. Mood, Focus, and Aspect

		AGENT	PATIENT	LOCATIVE	REFERENTIAL
REALIS					
	NEUT	mu-baket	baked-en	baked-an	saa-baket
	PERF	m\<in>u-baket	b\<in>aket	b\<in>aked-an	s\<in>u-baket
	PROG	ba-baket	ba-baked-en	ba-baked-an	saa-ba-baket
IRREALIS					
	FUT	ba-baked-ay	ba-baked-en	ba-baked-ay	
	ANT	mu-baked-ay	baked-aw		
	IMP	baket	baked-i		

(106) mu-baket rakihan ki   saw
      AF-beat   child   NOM person
      'The person beat a child.'

(107) baked-en saw   ki   rakihan
      beat-PF   person NOM child
      'The child was beaten by a person.'

(108) baked-an saw   ki   rakihan
      beat-LF   person NOM   child
      'The child was beaten by a person.'

(109) a. saa-baket rakihan ki   patakan
       IF-beat   child   NOM bamboo
       'A child was beaten with the bamboo.'

    b. saa-ba-baket rakihan ki   patakan
       IF-PROG-beat child   NOM bamboo
       'The child will be beaten with the bamboo.'

(110) m\<in>u-baket rakihan ki   aba
      AF-PERF-beat child   NOM Dad
      'The father has beaten a child.'

(111) b\<in>aket saw   ki   rakihan
      PERF-beat person NOM   child
      'The child has been beaten by a person.'

(112) b\<in>aked-an saw   ki   rakihan
      PERF-beat-LF   person NOM   child
      'The child has been beaten by a person (and bruises are visible).'

(113) s\<in>u-baket rakihan ki   patakan
      IF-PERF-beat child   NOM bamboo
      'The bamboo has been used to beat a child.'

(114) baa-baket   rakihan ki   saw
       PROG-beat.AF child   NOM person
       'The person is beating a child.'

(115) ba-baked-en   nisiw   a   aba ki   rakihan
       PROG-beat-PF you.GEN LIG Dad NOM child
       'The child will be beaten by your father.'

(116) ba-baked-an   nisiw   a   aba ki   rakihan
       PROG-beat-LF you.GEN LIG Dad NOM child
       'The child will certainly be beaten by your father.'

(117) a. ba-baked-ay   rakihan ki   aba
         PROG-beat-IRR.AF child   NOM Dad
         'The father will be beating a child.'

      b. ba-baked-ay   ki   punu wazu
         PROG-beat-IRR.LF NOM head dog
         'The head of a dog will be beaten.'

(118) ta-baked-aw   nita   ki   rakihan
       we-beat-IRR.PF we.GEN NOM child
       'Let's beat the child!'

(119) baked-i   nimu   ki   rakihan
       beat-IMP.PF you.GEN NOM child
       'You(PL) beat the child!'

When the only difference is between PF and LF of the same verb, as in (107–108), (111–112), and (115–116), a sentence with the LF verb is more definite than the one with the PF verb.

As mentioned above, the PF verb often refers to a past event in Pazeh, rather than to the future as in Atayal, as shown in the following examples from Tsuchida's texts.

(120) hau lia,   hutaxa'-en kuah   lia
       OK INCEP wait-PF   nothing ASP
       'Alright, it is useless to wait any longer.'

(121) pana'-en naki lia alaw a isia . . . iruma'-en, kita'-en, ini kita'-en
       shoot-PF I.GEN ASP fish   that   search-PF look-PF not see-PF
       'I shot that fish, . . . searched for it, and looked for it, but couldn't find it.'

7. Conclusion

Pazeh resembles most of the other Formosan languages in its aspect, focus, and case-marking system. It is similar to the Tsouic languages in making no distinction between common nouns and personal names, and in having the various NOM pronouns in the third person (see section 4 in this essay and Li 1997a). However, it does not have the complex system of auxiliary verb and main verb in each simple sentence, as in Tsou. Moreover, it does not have the infix -um- or its phonetic variants as com-

monly found in Formosan and Western Austronesian languages. Bunun does not have this infix either. Both Bunun and Pazeh have the prefixes *ma-* ~ *m-* to indicate Agent-focus. Pazeh is unique in using mostly the form *mu-*, which is considered more archaic than *-um-* for PAN.

To which Formosan language is Pazeh most closely related? Based on phonological evidence, Li (1985) suggests a close genetic relationship between Pazeh and Saisiyat. However, unlike Pazeh, Saisiyat does have the infix *-um-*, distinguishes between common nouns and personal names, and does not have short-form personal pronouns or locative pronouns as in Pazeh. Nevertheless, Pazeh and Saisiyat share the same sentence patterns in their existential and possessive sentences, not shared by other Formosan languages such as Rukai, Atayal, and Bunun. For example, the sentence 'I have a child' is expressed in these Formosan languages as follows:

Rukai (Tanan)		iakai	ka	lalak-li	
		exist		child-I.GEN	
Atayal (Mayrinax)		kiya	a'	'ulaqi'	mu
		exist	NOM	child	I.GEN
Bunun (Isbukun)		'aiza'	'inak	'uvaz	
		exist	I.GEN	child	
Pazeh	a.	yaku	ka	nahaza	rakihan
		I.TOP	TOP	exist	child
	b.	nahaza	rakihan	aku	
		exist	child	I.NOM	
Saisiyat		yako	hayza'	ka	korkoring
		I.NOM	exist	OBL	child

As illustrated here, the possessor does not serve as the subject of the sentence, but as the possessor in most Formosan languages such as Rukai, Atayal, and Bunun. In contrast, in both Pazeh and Saisiyat, the possessor functions as the subject (or topic) of the sentence, just as in English and Chinese. However, one piece of syntactic evidence is not adequate to establish a close genetic relationship between two languages. This requires further syntactic study and comparison.

Pazeh used to be a plains-tribe language spoken in the central western plains of Taiwan. However, both phonological and lexical evidence indicate that it differs greatly from all the other plains-tribe languages in the western plains: Taokas, Papora, Babuza, and Hoanya (Tsuchida 1982, Li 1985). Due to its close contact with Seediq and Bunun in the last century-and-a-half, Seediq and Bunun have exerted some lexical influence on Pazeh, as for example, Seediq *doriq*, Paz *dawrik* 'eye'; Bunun *lumaq*, Paz *xumak* 'house'.

REFERENCES

Ferrell, Raleigh. 1970. The Pazeh-Kahabu language. *Bulletin of the Department of Archaeology and Anthropology* 31/32:73–96.

Holmer, Arthur J. 1996. *A parametric grammar of Seediq.* Travaux de l'institut de Linguistique de Lund 30. Sweden: Lund University Press.

Huang, Lillian M. 1995. *A study of Mayrinax syntax.* Taipei: Crane Publishing.

Li, Paul. 1973. *Rukai structure.* Special Publications No.64. Taipei: Institute of History and Philology, Academia Sinica

———. 1976. Pazeh wordlist. Duplicated.

———. 1978. The case-marking systems of the four less-known Formosan languages. In *Second International Conference on Austronesian Linguistics: Proceedings,* ed. by S. A. Wurm and Lois Carrington, 569–615. Pacific Linguistics C-61. Canberra: Australian National University.

———. 1985. The position of Atayal in the Austronesian family. In *Austronesian Linguistics at the 15th Pacific Science Congress,* ed. by Andrew Pawley and Lois Carrington, 257–280. Pacific Linguistics C-88. Canberra: Australian National University.

———. 1996a. The pronominal systems in Rukai. In *Reconstruction, classification, description.* Festschrift in honor of Isidore Dyen, ed. by Bernd Nothofer, 209–230. Hamburg: Abera Verlag.

———. 1996b. *Formosan tribes and languages in I-Lan* (in Chinese). I-Lan: Government of I-Lan.

———. 1997a. A syntactic typology of Formosan languages: Case markers on nouns and pronouns. In *Chinese Languages and Linguistics, 4, Typological studies of languages in China,* ed. by Chiu-yu Tseng, 343–378. Taipei: Institute of History and Philology, Academia Sinica.

———. 1997b. *Formosan languages in Kaoshiung* (in Chinese). Kaoshiung: Government of Kaoshiung.

Li, Paul, and Ching-tsai Lin. 1990. The Pazeh ritual songs for ancestors and other types of songs (in Chinese). In *Field Materials: Institute of Ethnology,* 1–16. Academia Sinica Occasional Series 3. Taipei: Academia Sinica.

Lin, Ying-chin. 1988. Investigation of the Pazeh dialect in Puli (in Chinese). *The Taiwan Folkways* 39.1:176–200.

———. 1996. The Pazeh language (in Chinese). Unpublished MS, 42 pp.

Ross, Malcolm. 1995. Reconstructing Proto-Austronesian verbal morphology: Evidence from Taiwan. In *Austronesian studies relating to Taiwan,* ed. by Paul Jen-kuei Li, Cheng-hwa Tsang, Ying-kuei Huang, Dah-an Ho, and Chiu-yu Tseng, 727–791. Taipei: Institute of History and Philology, Academia Sinica.

Tsuchida, Shigeru. 1969a. Pazeh vocabulary. Duplicated.

———. 1969b. Pazeh texts. Tape-recorded.

———. 1982. *A comparative vocabulary of Austronesian languages of sinicized ethnic groups in Taiwan,* part 1, *West Taiwan.* Memoirs of the Faculty of Letters, University of Tokyo No.7.

Zeitoun, Elizabeth, Lillian Huang, Marie M. Yeh, Anna H. Chang, and Joy J. Wu. 1996. The temporal and modal systems of some Formosan languages: A typological perspective. *Oceanic Linguistics* 35:21–56.

8:   LEXICAL PREFIXES AND PREFIX HARMONY IN SIRAYA

Shigeru Tsuchida
Shung Ye Museum of Formosan Aborigines

1. Introduction[1]

When I was analyzing the Siraya data, I noticed an interesting fact, that a Siraya verb stem *-lpough* 'can, be able' occurs with quite a number of different prefixes, depending on the following verb. For example:[2]

(1)   assi   <u>mattæi-lpough</u>   ta   timamang   <u>mattæ'i-vlli</u>   tyni-æn
      assi   mattæ'i-lpuh   ta   ti-ma-mang   mattæ'i-vli   tini-an
      not   can:AF.NEUT   NOM   whoever   speak-back:AF.NEUT   3SGO

      ki   sasasat   ki   sou
      ki   sa-sa-sat   ki   su
      OBL   even-one   OBL   word
      'no man was able to answer him a word' (22:46)

(2)   Ka   <u>mou-lpough-ah</u>   mama   ki   mang   <u>mourbo</u>
      Ka   mu-lpuh-ah   mama   ki   mang   mu-rbu
      and   can:AF.fut   like   OBL   what   enter:AF.NEUT

      ta   timamang   ki   tallagh   ki   siouro   maligh
      ta   ti-ma-mang   ki   talah   ki   si-uru   ma-lih
      NOM   whosoever   OBL   house   OBL   most   strong
      'Or else how can one enter into a strong man's house?' (12:29)

See further the following (only the relevant words are extracted): *makou-lpough makou-Alilid* 'can pray' (26:53) (cf. *Alid* 'God'); *ni-smaki-lpough smaki-pæpænæh* 'could cast out' (cf. *mou-pænæh* 'come out' [4:9]); *ni-pæ'æ-lpoug pæ'ælikough* 'could restore' (F.14, 24b, p. 58) (cf. *æ'æ-lali-kough* 'again' [4:7]); *pou-lpough pou-sou* 'can speak' (12:34) (cf. *sou* 'word').

Besides, if we consider the following examples, we can say that the prefix *mættæ'i-* or *mættæ'-i* designates the meaning 'speak, say, talk' and the like: *mattæ'i-riang-â* 'bless!' (5:44) (cf. *ma-riang* 'good'); *mattæ'i-vli* 'answer' (11:4, etc.); *mattæi-pænæh-ah ki sou* 'profess'(7:23) (cf. *-pænæh* and *sou* above); *mattæ'i-doumma* 'say against' (26:62) (cf. *doumma* 'against' [5:23]); *mattæ'i-k'ma-hynna* 'say like this'(25:44) (cf. *k'ma hynna* 'like this'); *mattæ'i-la* 'speak again' (22:1; 26:44).

This type of prefix has already been noticed in the Tsouic languages by Tsuchida (1990), in Adelaar (1994; 1997a), and in Bunun by Nojima (1996) in Formosa, and in some Melanesian languages (e.g., Ezard 1978, Osumi 1995), and are often called "classificatory prefixes," but later renamed "lexical prefixes" by Nojima (1996).[3] I follow Nojima's naming hereafter. Nojima briefly touches upon the fact that the same prefix often appears both on the "lexical" verb and on the "adverbial" verb, as if the one contained in the following lexical verb is "copied" onto the adverbial main verb (Nojima 1996:17–18). Thus, for instance, in Bunun the prefix *mis-* 'burn' or *kis-* 'stab' occurs on both verbs in the following:[4]

> mis-utmag   mis-busuk
> burn-carelessly   burn-intoxicated
> '(He) carelessly became drunk.'

> kis-asu-a=s   mabananaz=tia kis-laupa
> stab-immediately=OBL   man=that   stab-stab
> 'Immediately after that, the man stabbed (the woman).'

Before dealing with Siraya lexical prefixes, a brief description of the Siraya verb structure will be given. Another source for the description of Siraya verb structure is Adelaar (1994), which was published in 1997.

---

2. Following the manner of representation used in Adelaar (1997a), in the following examples, the first line is the original orthography, the second line represents the morphological analyses in a probable phonemic form, the third line, the interlinear. The original orthography is not always clear about its phonetic/phonemic value, or what an apostrophe, a hyphen, or an accent mark meant, and I had to leave geminated consonants as they appear in many cases. For details, see Adelaar (1997b). Abbreviations in the interlinear are : A, Agentive pronoun; AF, Actor Focus; CAUS, causative prefix; E, emphatic pronoun; G, genitive pronoun; ITER, iterative; LF, Location Focus; LIN, linker; LOC, locative marker; N, Nominative pronoun; NOM, nominative marker; O, Oblique pronoun; OBL, oblique marker; OF, Object Focus; PERS, personal marker; Q, Question marker; EMPH, emphatic particle; EXCL, exclusive; FUT, future; IMP, imperative; INCL, inclusive; ITER, iterative; m-PRO, agentive pronoun beginning with *m-*; NEUT, neutral; PERF, perfective; PL, plural). For Siraya pronouns, see the appendix.

Examples are mostly from the Gospel of St. Matthew, but some, which are marked by (F), are taken from Gravius 1662, the figures after F representing the chapter number of Question and Answer, followed by the folio page number (a for face and b for back page), and the page number in Asai (1939). The first line is exactly as it appears in the original texts. Hyphenation in the original texts does not necessarily correspond to the morphological boundaries, which are, however, indicated in the second line. The mark = in the first line indicates that in the original text the parts preceding and following = are written separately, but morphologically they should be connected. Please notice that all the examples in the following, except for the second lines in interlinear, are the forms which appear in the original texts, and not the ones morphologically analyzed.

2. Verb Structure of Siraya

Siraya verbs inflect according to the focus system given in table 1.[5] The distinction between object focus (OF) and location focus (LF) is not clear, and actually there seems to be no major difference (see examples [3–7]). (Due to space limitations, not all inflected forms are exemplified.) {M}- represents the actor focus marker, which consists of allomorphs *m- ~ -m- ≈ ma- ≈ mi- ≈ Ø: m-* occurs with stems beginning with a vowel, *-m-* with stems beginning with a nonlabial consonant;[6] *ma-, mi-,* and *Ø* are morphologically conditioned. The AF Irrealis imperative aspect is marked by *-â/ -æ̂,* and AF Irrealis future is marked by *-ah.* The OF marker is *-en* for Realis, and *-au/ -æu* for Irrealis imperative and *-auh* for Irrealis future. LF marker is *-an* for Realis, *-ei* for Irrealis imperative, and *-eih* for Irrealis future. The suffixes *-ah, -auh,* and *-eih* have allomorphs as shown in the table: *-au/-ei* occur preceding an agentive pronoun beginning with *m-* such as *-mau* 'first singular agentive', *-moumi* 'second plural agentive', etc. (see example [26]). On the other hand, *-al, -aul,* and *-eil* occur preceding the particles *-appa* 'also' and *-ato* 'emphatic particle'. The agent in nonactor focuses is indicated by an agentive pronoun or by an oblique case marker *ki.*[7]

Table 1. Inflection of Siraya Verbs

		AF	OF	LF	CONDITIONS
REALIS	NEUTRAL	{M}-	-en	-an	
	PERFECTIVE	ni-{M}-	ni- (-en)	ni- (-an)	
IRREALIS	IMPERATIVE	{M}- -â	-au	-ei	
		{M}- ~ -æ	-æu		/æ__
	FUTURE	{M}- -ah	-auh	-eih	
		~ -au	~ -ei		/__m-PRONOUN
		{M}- ~ -al	~ -aul	~ -eil	/__-appa, /__ -ato

3. Nojima (1996:1) says, in footnotes: "They [= lexical prefixes] are very lexical in meaning in that they denote concrete events such as 'dying', 'dreaming', 'burning', or activities such as 'running', 'hitting', 'cutting', 'giving', which might be achieved by the use of independent verbs in other languages," and further: "The term 'lexical prefixes' might be reminiscent of the 'lexical suffixes' in the 'Mosan' languages in the Northwest Coast of North America, which are contrastively nominal in nature, however."

4. Grammatical abbreviations given by Nojima are omitted to avoid confusion.

5. I know that Stanley Starosta does not like to use the terms "focus" and "case markers," etc., and I feel guilty using them in this study dedicated to him. I hope he will overlook my obstinateness in consideration of my old age.

6. In the case of stems beginning with a labial consonant, it is not clear, but most likely a zero allomorph occurs: cf. *pa-pæh-piæh* 'scatter' (12:30), *papæhpiæh* 'divide' (25:32) / *papæh-piæh-eih* 'separate (LF)' (25:32); *pi'æ* 'give' (4:9, etc.) / *p'hæ-aneih* (LF) (14:9,etc.); *bæuut* 'bind' (12:29) / *væuut-au* (OF) (16:19, etc.); *mabys* 'pluck' (12:1); *mæiæ̈* (1:20), *mæ'jæ̈/ mæ'iæ̈* (19:9) 'marry', etc.

111

As in the Tsouic languages (Tsuchida 1976:43–110), Siraya verb stems can be classified into several classes depending on what forms appear in actor and nonactor focus inflections. Five classes are recognized. In the following examples, all the inflected forms are not necessarily available. Some forms are unattested, but can be inferred based on inflected forms of other verb stems belonging to the same class. Such forms are given in angle brackets < >.

2.1 Verb-stem Classes

Class I . Class I verb stems take {M}- for actor focus, but, unlike in other classes, there is nothing in nonactor focus except the OF marker *-en/-au/-auh* or the LF marker *-an/-ei/-eih* and their allomorphs. There is only one case where the word *-kytta* is inflected in object focus neutral form *(k'ita-en)*. All others are in location focus *(kytt'an, k'ytt'an, ni-k'yttan, ni-kytt'an, kytta-eih, kytta-ei, kytt'ey*, etc.).

(3)   ka   kaumang   ka   <u>k'ita-en</u>-hou   ta   houka ka   tou
      ka   kau-mang ka   kitta-en-hu   ta   huka ka   tu
      and   why   LIN   see:OF.NEUT-2SGA   NOM   mote   LIN   LOC

      matta   ki   tæ'i-apara   oho,   râ   ælogh [ka   pærænæh] ka
      matta   ki   tae'i-a-para   uhu,   râ   æluh [ka   pærænæh] ka
      eye   OBL   brother   2SGG,   but   beam [LIN   tree]   LIN

      tou   matta oho   assi ymhou   pou-liliht-in
      tu   matta uhu   assi imhu   pu-liliht-en?
      LOC   eye   2SGG   not   2SGN   consider:OF.NEUT
      'And why beholdest thou the mote that is in thy brother's eye, but considerest not the beam that is in thine own eye?' (7:3)

(4)   Irou   ka   k'ytt'an   tyn   ta   vare ka   maligh-da,
      iru   ka   kitta-an   tin   ta   vari ka   ma-lih-da,
      when LIN   see:LF.NEUT   3SGA   NOM   wind LIN   great-but,

---

7. Case markers in Siraya are as follows:

	NOMINATIVE	OBLIQUE	LOCATIVE
COMMON NOUNS	ta	ki	tou
PERSONAL NOUNS (=PN)	ta + ti	ti	ti + PN-an

An oblique case marker *ki* indicates any relations other than nominative and locative relations, so that it can indicate an object or instrumental/comitative 'with, together', etc. in a verb phrase, or a genitive 'of' in a noun phrase. A linker *(ka)* also indicates any relations between words, phrases, or sentences, which can often be translated in English as 'and, but, for, because', or as a relative pronoun.

Siraya pronouns are given in the appendix (for details, see Tsuchida 1996). The third person pronouns, both singular and plural, are independent forms, although all other personal pronouns are bound forms. The third person singular possessive *á-teni-ah* and the second person plural possessive *âmoumi-ah* are given in parentheses because they are AF-future forms, as indicated by the AF-future suffix *-ah*. (The forms *â-teni* and *â-moumi* are not attested as such.) Also, neither the inclusive nor the exclusive first-person plural possessive pronouns have been found in the extant data available.

mahtakout-ato
mah-takut-atu
afraid:AF.NEUT-EMPH
'But when he saw the wind boisterous, he was afraid' (14:30)

There are at least twelve cases of *k'ytt'an/kytt'an* (LF.NEUT), but all of them occur in a subordinate clause beginning with *irou* or *rou* 'when, if' as in (4). It is not clear yet whether or not this is significant.

Table 2. Inflection of Class I verb stem -*kiym/-ki'ym* 'search, seek'

		AF	OF
REALIS	NEUTRAL	kmi-'ym	kiym-en
	PERFECTIVE	ni-kmi'ym	ni-ki'ym-en
IRREALIS	IMPERATIVE	kmi-ym-â	<kiym-au>
	FUTURE	kmi'ym-ah	<kiym-auh>

Table 3. Inflection of Class I verb stem -*kytta* 'see, look'

		AF	OF	LF
REALIS	NEUTRAL	kmytta	k'ita-en	kytt'an
	PERFECTIVE	ni-kmytta		ni-k'yttan
IRREALIS	IMPERATIVE	kmiyttâ		kytt'ey
	FUTURE	kmytta-ah		kytta-eih

Class II. Class II verb stems take *m-* for actor focus, but *p-* for nonactor focus.'

Table 4. Inflection of Class II verb stem -*attæ'i-vli* 'answer'

		AF	OF
REALIS	NEUTRAL	mattæ'i-vli	pattæ'i-v'lli-en
	PERFECTIVE	ni-mattæ'i-vlli	ni-pattæ'i-vlli-en
IRREALIS	IMPERATIVE	<mattæ'i-vli-â>	<pattæ'i-vli-au>
	FUTURE	mattæ'i-v'lli-ah	<pattæ'i-vli-auh>

Class III. Class III verb stems take *m-* for actor focus, but *k-* for nonactor focus. Verb stems which belong to this class are mostly those that express some kind of emotions: *mavæangö* (AF.NEUT)/*kavæangö* (OF.NEUT) 'love', *mavahir* (AF.NEUT) (Vlis)/*kava-hir-au* (OF.IMP) (5:43) 'hate', *mamoei* (AF.NEUT) (6:5, etc.)/*kamoei-en* (OF.NEUT) (6:10, etc.) 'wish, will', and *marei* (AF.NEUT) (18:13)/*karei-en* (OF.NEUT) (25:21, etc.) 'rejoice, joy'.[8]

---

8. The following word does not seem to express emotion: *mavoë* (AF.NEUT) (24:19)/*kavoui-in* (OF.NEUT) (1:23) 'conceive, become pregnant', but this might show a part of the Siraya people's world view that considers conception as a very emotional event.

Table 5. Inflection of Class III verb stem *-a-væangö* 'love'

		AF	OF
REALIS	NEUTRAL	mavæangö	kavæango-en
	PERFECTIVE	<ni-ma-væangö>	<ni-ka-væango(-en)>
IRREALIS	IMPERATIVE	<ma-væangö-â>	kavæango-au
	FUTURE	mavæ'ango-ah	kavæango-au

Class IV. Class IV verb stems take *m-* for actor focus, but *h-* for nonactor focus. I have so far found only one example for this class, which is unfortunately very defective.

Table 6. Inflection of Class IV verb stem *-ali-da'æua* 'pass, depart'

		AF	OF
REALIS	NEUTRAL	mali-daæua	hali-dæu-en
	PERFECTIVE	ni-mali-daæua	
IRREALIS	IMPERATIVE		
	FUTURE		

Class V. Class V verb stems take Ø for both actor and nonactor focuses.

Table 7. Inflection of Class V verb stem *pæ'æ̈-tounnoun* 'deliver, give'

		AF	OF	LF
REALIS	NEUT	pæ'æ̈-tounnoun	<pæ'æ̈-tounnoun-en>	<pæ'æ̈-tounnoun-an>
	PERF	ni-pæ'æ̈-tounnoun	ni-pæ'æ̈-tounnoun-en	<ni-pæ'æ̈-tounnoun-an>
IRREAL	IMP	<pæ'æ̈-tounnoun-â>	<pæ'æ̈-tounnoun-au>	<pæ'æ̈-tounnoun-ei>
	FUT	pæ'æ̈tounnoun-ah	pæ'æ̈-tounnoun-auh	pæ'æ̈-tounnoun-eih
				pæ'æ̈-tounnoun-an-eih

Formally, three different forms appear: *pæ'æ̈-tounnoun-auh* for object focus future form (only one case), *pæ'æ̈-tounnoun-eih* (three cases) and *pæ'æ̈-tounnoun-an-eih* (two cases) for location focus future form. But all three appear in a very similar, or rather exactly the same environment, and thus no functional or semantic difference can be detected. See the following examples.

(5) ka <u>pae'æ̈-tounnoun-an-eih</u> ta Alak ki kaæuloung ki
ka pæ'æ̈-tunnun-an-aih ta Alak ki kaæulung ki
and deliver:LF.FUT NOM Son OBL man OBL

Sih-bavau ka Tama-p'higik
Sih-vavaw ka Tama-pi-hi-hik
chief LIN priest.PL
'and the Son of man shall be betrayed unto the chief priests' (20:18)

(6) K'ytt'ei   matouk-ato   ta   kidi ki tæ'lei, ka
    kitta-ay   ma-tuk-atu   ta   kidi ki tæ'lay, ka
    look:LF.IMP near:AF.NEUT-EMPH NOM time OBL hour,   LIN

    <u>pæ'æ̈-tounnoun-auh</u> ta   Alak ki ka'æuloung tou rima
    pae'"ae-tunnun-auh ta   Alak ki ka'aeulung tu   rima
    deliver:OF.FUT   NOM Son OBL man   LOC hand

    ki   tama-kavarau
    ki   tama-ka-varaw
    OBL sinner
    'Behold, the hour is at hand, and the Son of man is betrayed into the hands
    of sinners.' (26:45)

(7) ka   <u>pæ'æ̈-tounnoun-eih</u>   ta   Alak ki ka'æuloung tou
    ka   pæ'æ̈-tunnun-aih   ta   Alak ki ka'aæulung tu
    and deliver:LF.FUT   NOM Son OBL man   LOC

    pouna-væværæh-eih tyni-æn
    pu-na-væ-væræh-aih tini-an
    crucify:LF.FUT   3SGE
    'and the Son of man is betrayed to be crucified.' (26:2)

3. List of Siraya Lexical Prefixes

There are quite a number of lexical prefixes found in Siraya, but most of them are not clear in meaning. The following are the ones whose meaning could somehow be extracted. Some of these are already discussed in Adelaar (1997a). In some cases only one example is available and yet the meaning is clear. Such prefixes are listed with a note 'unique'. Stem class is indicated by a Roman numeral in parentheses immediately after the prefix. The initial *m-*, *p-*, or *k-*, which should be treated as allomorphs of the focus marker, are not separated from the stem in the following list. *R-* indicates a reduplication of the first two syllables (minus the last consonant of the second syllable) of the stem, whereas *r-* indicates a reduplication of the first consonant of the stem followed by the vowel *a*. N- indicates a noun stem.

   1. *æ'ia-* (V) 'be, exist, be with':
      æ'ia-apang 'be on the rooftop' (24:17)
      æ'ia-avang 'be in the ship' (14:33)
      æ'ia-laf 'be in the chamber' (24:26)
      æ'ia touk æ'ia-ngataf 'be near, be at the door' (24:33)
      æ'ia-hylli 'be near' (24:32)
      æ'iä-tounnoun ki vullu-vullum 'be in the heaven' (7:21)
      æiakoua 'be with' (26:11; 9:15)
      æia-ymmid æia-lam 'be all with' (13:56)

   2. *hali-* (IV or V) 'bind':
      hali-rima-ei 'Bind him hand (= *rima*)!' (22:13)
      hali-rahpal-ei 'Bind him foot (= *rahpal*)!' (22:13)

3. *h-m-ou/hou-/houh-* (I) 'drink' (cf. *myt* 'to drink'):
   hou-baæu-au-mau 'I drink it new' (26:29)
   hmou-lam 'drink with' (26:29) (cf. *lam* 'together, with')
   hou-ouro-en 'before drink it' (26:42)
   houh-tmouk-au 'Drink it all!' (26:27)

4. *itou-* (I) 'exist, be' (cf. *tou* 'locative marker'):
   itou-dung 'be in secret' (5:14)
   itou-qua 'exist, dwell' (2:13, 24:2, etc.)
   pa-itou-halap '(make be upon =) put upon' (24:2)
   itou-lam 'be together' (15:32, 8:11)
   itou-mala 'be outside' (12:46)

5. *k-* (V) 'eat' (cf. *k-m-an* 'to eat'):
   ni-k-ymmid 'ate all' (14:20)(cf. *ymmid* 'all')
   ni-k-da 'remained in eating' (14:20, 15:37)
   k-koua-ah 'Come to eat!' (26:18)(cf. *mou-koua* 'come, go')
   ni-k-lam 'ate together' (26:20)(cf. *lam* 'together')

6. *kæih-/kæi-/kmæih-/kmæi-/km-æi-/kmeih-* (I) 'say (with a bad connotation ??)':
   kæi-dmir-ou (< *kæi-dmir-au*) '(say-haphazardly >) prophesy (OF.IMP)'
      (26:68) (cf. *mattæ'i-rmir-ah* 'say haphazardly (AF.fut)' [24:48])
   ni-kmæi-beiaveiah '(they) all said (AF.pf)' (27:22) (cf. *veiagh* 'every one'
      [25:29])
   kmæih-tan '(say beforehand >) prophecy' (26:68) (cf. *si-tan* 'prepare'
      [22:4])
   kmeih-kahier 'condemn' (12:7), *kmæih-kahier* 'judge' (7:1)
   ni-kmæ-i-naunamou 'began (to rebuke) (AF.pf)' (16:22)

7. *kæuh-* (V) 'become':
   kæuh-boulas 'become sorry' (14:9) (*ma-voulas* 'sorry')
   ni-kæuh-darikah 'became healed' (8:13; 15:28) (*ma-darikagh* 'be whole')
   kæuh-kunnum-ato 'become waxed' (13:15)
   ni-kæuh-likough kæuh-riang (12:13) 'was restored (< became good again)'
   kæuh-pærænænæh-ato 'become a tree' (13:32)
   ni-kæuh-'mha-ato 'become grown' (13:32)
   kæuh-pæ-pæræh 'becoming a man = a young man' (19:20,22)
   ni-kæuh-vahto 'became headstone' (21:42)

8. *mahko-/mako-/pahko-* (II) 'say something bad':
   mahko-av-avagh 'curse' (27:39)
   ni-pahko-avavagh-en (25:41) 'curse'
   mahkou-doumma 'say-adversary' (5:11)
   barigh mahkou-talum '(blasphemy say-plant >) blasphemy' (9:3)

9. *ma(h)kou-/pa(h)kou-* (II) 'pray' (??):
   makou-Alilid/pakou-Alilid 'pray-god' (4:9)
   makou-tohkouko 'pray-stand' (6:5)
   makou-haradouhat 'pray-repeat' (6:7)
   Makou-sasoulat = Mahkou-sasoulat 'scribes'(23:15)(20:18)

ni-makou-v'lli 'return thanks = bless' (26:26)
pakou-riang-auh 'pray-good = bless (OF.fut)' (21:9)
pakou-tiktik-auh 'hallow (OF.fut)' (6:9)

10. *maki-/paki-* (II) 'find':
maki-valei, paki-valei-en 'find' (7:8)(12:44)
maki-ouro 'find first' (6:33)
maki-kaha 'finish to find' (27:7)

11. *makka-* 'for, during':
makka=sât ki wæ'i 'for one day' (20:6)
ni-makka-touro ki wæ'i 'was for three days' (12:40)

12. *mako-/makou-/pako-* (II) 'cry':
mako-ligh 'cry loud' (27:50) (cf. *ma-ligh* 'mighty')
mako-langæh 'cry' (21:15)
mako-dyllugh 'cry after' (15:23) (cf. *dmyllough* 'to lead' (6:13)')
mako-toun 'cry out' (27:23)
ni-makou-saoun 'cry more (AF.pf)' (20:31), *pakosaoun* 'id.' (OF.NEUT)
(27:23)

13. *mali-/hali-* (IV) 'away (??)':
mali-æurægh 'pass away' (26:42)
pahali-æurægh-â jau-an 'let pass from me!' (28:39)
mali-kaha'to mali-daæua ta kidi 'the time is now past' (14:15)
halida'æu-an (26:2), *Hali-da'æu-en* (26:17) 'passover'
pa-hali-dæuræ 'offended' (< 'made stumble and put away from belief' ??)
(11:6)

14. *matta-* 'accept' (??):
matta-rÿh 'accept in the heart' (= *rÿh*) (13:20)
matta-dys 'accept straight away' (= *madys*) (13:20)
matta-rei 'accept with joy' (13:20) (cf. *marei* 'joyful')

15. *matta'æu-* 'sleep' (??) (unique):
matta'æu-dyk merip 'only sleep' (9:24) (cf. *dyk* 'only', *merip* 'sleep')

16. *mattæ'i-/pattæ'i-* (II) 'speak':
mattæ'i-k'ma 'say' (27:63, etc.)
mattæ'i-limoulimou 'finish saying' (19:1; 26:1)
ni-mattæ'i-naunamou 'began to speak' (26:22)
mattæ'i-riang-â 'bless (AF.IMP)' (5:44)
ni-mattæ'i-sâl mattæ'i-tavætavæh 'said among themselves' (21:38)
mattæ'i-vli 'answer' (11:4, etc.)
mattæi-mam'-ah 'say-how' (7:4) (cf. *mama* 'like, as')
mattæi-pænæh-ah ki sou '(speak-come-out >) profess' (7:23)
ni-mattæi-mado tyni-æn 'said within herself' (9:21)

17. *mau-* 'laugh ?' (unique):
ni-mau-ak-ak-ak mattaua 'scorned' (9:24) (cf. *mattaua* 'laugh')

18. *mæææu-/mææu-/paææu-* (II) 'reciprocal':
    mæææu-amsing 'be-enough with each other' (20:2)
    mæææu-kla 'meet each other' (20:2)
    mæææu-ymd'-ah k'ytta maæuakla 'we shall all meet each other (fut.)'(F.30, 53a, p.115)

19. *mæææu-/mæ'æu-/æææu-/'ëaæu-/ë'æu-/eu-/a-æu-* (I)'take away, release, un-':
    mæææu-tætæææuf 'uncover, reveal'(11:27) (cf. *ta'æuf-en* 'to cover [10:26]', *tmagof* 'to cover [Vlis]')
    'ææu-lb-eih 'will be opened'(7:7,8) (cf. *ni-ilb-an* 'was shut' [25:10])
    ë'eu-kakirihg-ei '(untie =) loosen'(21:2) (cf. *kakirihg* 'tied'[21:2])
    æ'æu-vavæuut '(unbind =) loosen'(18:18)(cf. *bæuut* 'bind [12:29]')
    mæ'æumia 'release'(27:15)(cf. *akoumea* 'have [13:12, etc.]')
    eu-lpoug-en 'can be relieved' (F.11, 20a, p.49)
    a-æu=ymd-en 'all are set free' (F.33, 63a, p.135)

20. *mæ'æu-* 'save' (perhaps related to the above 'take away, release'):
    mæ'æu-lpough 'can save' (27:42)
    mæ'æu-mado 'save oneself' (27:42)

21. *mei-/pei-* (II) 'having' (??):
    mei-alak '(have a child=) give birth to' (1:25)
    mei-kakan '(have a meal=) eat' (26:26)
    mei-paringid '(have a flute=) play on flute' (9:23)
    mei-ra-rarei '(have fury=) upbraid' (11:20)
    mei-ringei '(have work=) work, make, build' (27:40, etc.)

22. *mou-/ou-* (I) 'locomotion':
    mou-koua, mou-qua, ou-koua 'go, come' (8:9; 25:22; 21:30, etc.)
    mou-naunang 'walk' (15:31)
    mou-rbo 'enter' (8:32, etc.) (cf. *mou-lpough-ah mou-rbo* 'will be able to enter' in example [2])
    mou-vavau 'go up' (20:18) (cf. *vavau* 'above')
    mou-pænæh 'come/go out' (8:28, etc.) (cf. *pænæh* 'open place')
    ni-mou-tallagh 'entered the house (= *tallagh*)'
    mou-voukkyn 'go into the mountain (= *voukkyn*)' (14:23)
    mou-Jerusalem 'go to Jerusalem' (20:18)

23. *pæ'ë-/pæ'æ-* 'give' (V) (cf. *pihæ* 'to give'):
    pæë-saoun-ah 'give-more (AF.fut)' (7:11)
    pæ'ë-ouryf 'give-gifts (AF.NEUT)' (7:11)
    pæ'ë-likough-ah 'give-back, reward (AF.fut)' (6:6)
    pæ'æ-kaha-kow pæ'æ-tarimukah ki si-ouro . . . 'thou hast paid the uttermost farthing.' (5:26)
    pæ'ë-tounnoun-ah 'will-give-deliver-up' (AF.fut)(10:21)

24. *pæ'i-(r)-R-* 'be divided into several and fight against each other' (??):
    pæ'i-pa=pouropouroh 'be divided into several lands and fight each other' (cf. *pourough* 'land') (12:25)

pæ'i-æumaæuma 'be divided into several cities and fight each other'
  (cf. æuma 'city') (12:25)
pæ'i-tallatallagh 'be divided into several houses and fight each other'
  (cf. tallagh 'house') (12:25)

25. pæh-/pag- (V) 'think':
pæg-dimdim 'think' (3:9; 5:17, etc.)
pæh-'kbou 'take care' (6:25, etc.)
pæh-alalei pæh-dimdim 'reason' (21:25) (alei 'for')
pæh-balei-en pæh-dimdim 'remember' (16:8) (maki-valei 'find')
pæh-dimdim pæh-madou 'think oneself' (16:7)

26. paha-r-N-, paha-R- (V) 'transform into':
paha-papa'ul-auh 'will be transformed into bread (pa'ul)'(4:3)
pahaviri-viri-in 'translate' (1:23)
pahava-vouël 'transform into body (vouël)' (F.20, 34a, p.77)
paha-vavatti 'transform into soul (vatti)' (F.20, 34a, p.77)
paha-ä-mamang 'transform into things (mamang)' (F.21, 35a, p.79)

27. pis- (V) 'die, kill':
pis-lalam mapatei 'die with' (26:35) (cf. lam 'with', ma-patei 'die')
pis-ma-mourang-en 'pestilence' (24:7)
pis-bouravourel pis-koupang 'lunatic' (17:15)
ni-pis-koua 'go to die' (AF.pf) (22:27)
pis-ilpough paka-patei 'can kill' (10:28) (cf. paka-patei 'kill')
ni-pis-tarimoukah 'perish' (??) (8:32)

28. pou- (V) 'put' (??):
pou-pænæh 'put forth, bring forth' (12:35, etc.) (cf. mou-pænæh 'come out')
pou-alak 'make child, beget' (23:15) (cf. mei-alak 'give birth to')
pou-vavau 'lift up' (12:11) (cf. mou-vavau 'come up [17:27]')

29. pou- 'say, speak' (perhaps related to 28, thus 'put into word' means 'speak'):
pou-sou 'speak' (12:30) (cf. sou 'word')
pou-ræmæh 'speak in light' (10:27) (cf. ræmæh 'light')

30. sa-/sah- (V) 'go through with some difficulty':
sa-likough 'return' (2:12) (cf. ni-ma-likough ma-voulas 'repent' [21:32])
sa-ma-mala 'go out' (9:31,32: 24:1, etc.) (cf. s-m-aki-mala 'cast out' [5:13])
sa-ra-rbo 'enter through' (7:13, etc.) (cf. mou-rbo 'go in, come in' (8:32, etc.)

31. saki-/s-m-aki- (I) 'throw away, cast away':
smaki-qua 'cast down' (7:6) (cf. mou-koua 'come, go')
smaki-pæpænæh 'cast out' (17:19)
smaki-mala 'cast out' (5:13)(cf. mou-mala 'go out')
smaki-tauagh 'off' (21:39,25:30) (cf. matau-tauagh 'afar off'[27:55])
smaki-vaoung 'cast into sea (vaoung)' (4:18)

32. saou-/sau-/sauh- (V) 'swear':
saulkut 'oath' (14:7)
ina saou-bongo saou-lkut ki vongo 'don't swear by head (= vongo)' (5:36)

119

ina saou-bullum 'don't swear by heaven (= *vullum*)' (5:34)
yna saou-jerusalem 'don't swear by Jerusalem' (5:35)
yna saou-næi 'don't swear by earth (= *næi*)' (5:35)
sauh-pyttou-appa kytti-æn sauh-pyttoul '(swear) seventy-seven times' (18:22)

33. *si-/si-a-* (V) 'do beforehand' (??):
   si-a-tæutæuugh '(teach beforehand >) appoint' (26:19)(cf. *matæutæuuh*
      'teach' [14:8])
   si-kakidi '(command beforehand >) precepts, instruct, command' (21:6)
   si-tan 'prepare' (22:4) (cf. *Mattæi-tan* 'prophet' [2:5])
   sitatanang-en 'prepare' (27:62) (cf. *tmananang* 'ready' [24:44])

34. *sou-/s-m-ou-* (I) 'touch':
   smoukla 'touch' (cf. *mæua-kla* 'gather, meet')
   smou-dyk-ah smoukla 'they might only touch (the hem of his garment)'(14:36)
   dou sou-dyk-en-au smoukla ki koulamough tyn 'if I could touch his gar-
      ment' (9:21)

35. *tæu-* 'from' (??):
   tæu-æuma 'from city (= *æuma*)' (15:22)
   tæu-æmag-an 'from blood (= *æmægh*) (LF.NEUT)' (9:20)

36. *tau-/'tau-* (V) 'go toward, go on, go through':
   tau-avang 'go on a ship' (14:13,34) (cf. *avang* 'ship')
   tau-ryttoung 'go around, travel' (4:23,23:15)(cf. *ryttoung* 'around')
   tau-qua, 'tauqua 'go toward, go on' (2:9,12:1)
   ni-tau-pourough 'went through the country' (14:13)

37. *tauki-/t-m-auki-* (I) 'ask':
   tmauki-lala 'ask again'(22:46) (cf. *ni-lala* 'did again'[26:72])
   tmauki-lymmugh 'finish asking' (2:16) (cf. *limou* 'end')
   ni-tauki-ymd-in 'all asked' (10:30) (cf. *ymmid* 'all')
   ni-tauki-ouro-en 'was asked first' (F.26, 43b, p.96) (cf. *ouro* 'first')
   tmauki-tounnoun 'ask to deliver' (F.26, 43b, p.96)

38. *tna-/tnæ-* (V) 'psychological action':
   tna-msing 'believe' (17:17) (*ma-msing* 'enough')
   tnæ-ghpang-ato 'were displeased' (21:15)
   tna-harum-meih 'will be sympathized' (cf. *ka-harum-an* 'mercy'[6:7])
   ni-tna-væingbing 'were moved to become angry' (20:24)

4. Serial Verb Construction and Prefix Harmony

As seen in the examples in the previous sections, a serial verb construction is found in
Siraya that consists of two (or sometimes three, but rarely four) verbs. Some verbs
that appear in the first position have no prefixes, as in *subbu-an-eih* 'will be casted'
(22); some always occur with a particular prefix, as in *ni-mah-takout* 'feared' (21);
and some occur with various prefixes, depending on the verb, as in *mattæ'i-lpough*
(1), *mou-lpough* (2), etc. 'can'. The last type of verb may be called an "auxiliary

verb," although an auxiliary verb does not differ from all other verbs as far as the inflection types are concerned.

Some auxiliary verbs are given in table 8, and more may be added (such as -*dung* 'do in secret' in section 4.2). Only one example is given for each here; for -*lpough*, see (1) and (2).

(8) Tou kidi k'anna <u>ni-matta-naunamou</u>   ta   ti   Jesus
     Tu kidi ka-ana ni-matta-nawnamu   ta   ti   Jesus
     LOC time that   speak-begin:AF.PERF   NOM   ti   Jesus

     <u>matta-sasou,</u>   mattæ'i-k'ma-hynna,
     matta-sasu,   mattæ'i-kema-hinna,
     preach:AF.NEUT,   say-like-this:AF.NEUT,
     'from that time Jesus began to preach, and to say,' (4:17)

(9) ka   <u>mali-kaha</u>'to   <u>mali-daæua</u>   ta   kidi:
     ka   mali-kaha-atu   mali-daæwa   ta   kidi
     and   away-end:AF.NEUT-EMPH   away-past:AF.NEUT   NOM   time
     'and the time is now past' (14:15)

(10) mou-vaoung-â,   pa-heiriryng-an-ei   ta   heiryng, ka   arau
      mu-vaung-â,   pa-hi-ri-ring-an-ay   ta   hiring, ka   ara-au
      go-sea:AF.IMP,   cast-hook:LF.IMP   NOM hook,   and   take:OF.IMP

      ta   t'hing ka   <u>mou-ouro</u>   <u>mou-vavau,</u>
      ta   thing ka   mu-uru   mu-vavaw
      NOM fish   LIN   come-first:AF.NEUT   come-up:AF.NEUT
      'go thou to the sea, and cast an hook, and take up the fish that first cometh up' (17:27)

(11) Siksik-au   ta   rima-oho: ka   ni-sik-ken   tyn,   ka
      sik-sik-au   ta   rima-uhu: ka   ni-sik-en   tin,   ka
      stretch.ITER:OF.IMP NOM   hand-2SGG: and   stretch:OF.PERF 3SGG, and

      <u>ni-kæuh-likough</u>   <u>kæuh-riang</u>   mama ki   idi   [tyn].
      ni-kæuh-likuh   kæuh-riang   mama ki   idi   [tin].
      become-back:AF.PERF become-good:AF.NEUT like   OBL other [3SGG]
      'Stretch forth thine hand. And he stretched it forth; and it was restored whole, like as the other.' (12:13)

Table 8. List of Auxiliary Verb Stems

1. -*lpough* 'can'	6. -*mado* 'do oneself'
2. -*naunamou* 'begin'	7. -*sâl* 'do together'
3. -*kaha* 'finish '	8. -*saoun* 'do more'
4. -*ouro* 'do first'	9. -*ymmid* 'do all'
5. -*likough* 'do again'	

(12) Timamang   ta   <u>pou-mado-ah</u>   <u>pou-rarim</u>   tyni-æn
Ti-ma-mang ta   pu-madu-ah   pu-rarem   tini-an
whosoever   NOM   make-self:AF.FUT make-down:AF.NEUT 3SGO

ki   kidi mama ki   rauwei k'atta,
ki   kidi mama ki   raway ka-atta
OBL   ?   like   OBL   child   LIN-this
'Whosoever therefore shall humble himself as this little child?' (18:4)

(13) wæ'ii   <u>ni-mi-sâl-koh</u>   <u>mierrong</u>   ymoumi-æn
wæ'i-i ni-mi-sâl-kuh   m-irung   imumi-an
day.PL   do-together:AF.PERF-1SGN sit:AF.NEUT 2PLO

mahtatæutæuugh   tou   Kouva ki   tatamd-den   ki   Alid
mah-ta-tæu-tæuuh tu   Kuva   ki   ta-tamd-en   ki   Alid
teach:AF.NEUT   LOC   hut   OBL   worshipping   OBL   God
'I sat daily with you teaching in the temple.' (26:56)

(14) Ka   rou   itouqua   tou   pasyt-en ki   rych   ka   malouaf,
Ka   ru   itu-kua   tu   pa-sit-en ki   riih   ka   ma-luaf,
and   when   be:AF.NEUT   LOC   struggle   OBL   heart   LIN   grievous,

<u>ni-makousaoun</u>   <u>makou-kilkil</u>   <u>makou-Alilid</u>   (ta   teni.)
ni-maku-saun   maku-kilkil   maku-Ali-lid   (ta   teni.)
pray-more:AF.PERF pray-earnest:AF.NEUT pray:AF.NEUT (NOM   3SGN)
'And in anguish of spirit he prayed the more earnestly.' (F.23, 38b, p.86;
Luk22:44)

(15) Ka   <u>ni-mi-ymmid</u>   <u>mavok,</u>   ka   mibangtau-ato
Ka   ni-mi-immed   m-avuk,   ka   mi-vangtaw-atu
and   do-all:AF.PERF   eat:AF.NEUT, and   full:AF.NEUT-EMPH
'And they did all eat, and were filled.' (15:37)

This phenomenon may be called "prefix harmony" in a serial verb construction. Another term for it is "anticipating sequence" (cf. Adelaar 1997a). The prefix harmony is observed even with the causative prefix *pa-* and *paka-*, the latter of which occurs with the stative verb stem:

(16) Ka   ni-ad-au   ta   teni   pa-ou-qua tou
Ka   ni-ad-au   ta   taini pa-u-kua   tu
and   bring:OF.PERF-1SGA   NOM   he   toward   LOC

Pahtatæutæu-en-hou,   ka   assi nein   <u>pa-lpoughin</u>   tyni-æn
Pah-ta-tæutæu-en-hu,   ka   assi nain   pa-lpuh-en   tini-an
disciple.PL-2SGG,   and   not   3PLA   CAUS-can:OF.NEUT 3SGE

<u>pa-darikægh</u>
pa-darikæh
CAUS-whole:AF.NEUT
'And I brought him to thy disciples, and they could not cure him.' (17:16)

(17) Ka kytt'ei   ta   Leproos ka ou-tatäum-min,   ni-iroua
     Ka kitta-ay   ta   Leproos ka u-ta-täum-en,   ni-irua
     and look:LF.IMP NOM   leper   LIN suffer?:OF.NEUT, come:AF.PERF

     deip   tyni-æn,koun,   Meirang, dou
     daip   tini-an, kua-en,   Ma-irang, du
     worship:AF.NEUT 3SGO, say:OF.NEUT, lord,   if

     mamoei-kow,   paka-lpough   kow   paka-kouptigh
     mamuy-kow,   paka-lpuh   kow   paka-kuptih
     wish:AF.NEUT-thou, CAUS-can:AF.NEUT 2SGN CAUS-clean:AF.NEUT

     iau-an-da
     yau-an-da
     1SGO-EMPH
     'And, behold, there came a leper and worshipped him, saying, Lord, if
     thou wilt, thou canst make me clean.' (8:2)

4.1 Determiner of Focus and Determiner of Prefixes

With the serial verb construction the following facts are observed: (a) in whatever
focus and aspect the auxiliary verb is inflected, all the following verbs are always in
AF neutral form, and (b) the lexical prefix that appears with an auxiliary verb is deter-
mined by the lexical prefix of the following (mostly the last) verb. In other words, it is
the auxiliary verb that is properly inflected according to the contexts, not the main
verb; on the other hand, it is the main verb that follows that determines the prefix, not
the auxiliary verb. There are numerous examples that indicate the above observation,
such as (8), (11), (12), (14), and (16). See further the following:

(18) tou euvan k'atta dou assi-appa mouni   ta   taoukka,
     tu   euvan ka-atta du   assi-appa m-uni   ta   taukka,
     LOC night LIN-this when not-also cry:AF.NEUT NOM cock,

     ka   t' touroul-ei-mhou   rmau   jau-an
     ka   t' turu-l-ai-mhu   r-ma-u   yau-an
     and ?   do-thrice:LF:FUT-thy deny:AF   1SGO
     'this night, before the cock crow, thou shalt deny me thrice' (26.34)

(19) Ka ni-mattæi-mado   tyni-æn, rou   sou-dyk-en-au
     Ka ni-mattæi-madu   tini-an, ru   su-dik-en-au
     and speak-self:AF.PERF 3SGO,   if   touch-only:OF.NEUT-1SGA

     smoukla   ki   koulamough tyn,   dou   miaka
     s-m-u-kla   ki   kulamuh   tin,   du   miaka
     touch:AF.NEUT OBL garment   3SGG, when then

     'ddærikæh-ah-koh.
     dærikæh-ah-koh.
     become-whole:AF.FUT-1SGN
     'for she said within herself, If I may but touch his garment, I shall be
     whole.' (9:21)

A main verb, however, does not necessarily appear in the last position in the serial verb construction. Thus:

(20) Neni-râ   ni-pæh-alalei   pæh-dimdim
        naini-râ   ni-pæh-al-alay   pæh-dimdim
        3PLN-but   think-reason:AF.PERF think-think:AF.NEUT

        pæh-madou   neini-æn
        pæh-madu   naini-an
        think-self:AF.NEUT 3PLE
        'And they reasoned among themselves,' (16:7)

Here the main verb *pæh-dimdim* 'think', which determines the appearing of the lexical prefix *pæh-* in the other two verbs, occurs in the second position.

If the first verb is not an auxiliary verb, prefix harmony does not occur, even if principle (a) above holds good that only the first verb is properly inflected in the serial verb construction and the second (main verb) stays always in actor focus neutral form:

(21) ni-mah-takout sah-ka-qua hynna ta   teni:
        ni-mah-takut   sah-ka-kua hinna ta   taini:
        fear:AF.PERF   go:AF.NEUT there NOM 3SGN
        'he was afraid to go thither:' (2:22)

(22) Assi-ato mariang amamang,   dyk   subbu-an-eih smaki-mala
        Assi-atu ma-riang a-ma-mang, dik   sebe-an-aih   s-m-aki-mala
        not-EMPH good   anything,   only cast:LF.FUT   cast-out:AF.NEUT

        ka g'itgitauh ki   ka'æuloung.
        ka hithit-auh   ki   ka'æulung.
        and tread:OF.FUT OBL men
        'It is thenceforth good for nothing, but to be cast out, and to be trodden under foot of men.' (5:13)

In (21) and (22), the verbs *ni-mah-takout* 'feared' and *subbu-an-eih* 'cast', which appear in the first position, do not take the same prefix as that of the main verb *sah-* 'go through with some difficulty' or *s-m-aki-* 'throw away' because they are not auxiliary verbs.

4.2 Possible Origins of Lexical Prefixes

The origin of lexical prefixes in Siraya is not at all clear. In a couple of cases, there is a vague formal resemblance between prefixes and semantically corresponding verbs. Thus, for instance, the lexical prefix *k-* 'eating action' appears to be similar to the first consonant of the verb *kman* 'eat' (< *k-em-dʔen*):[9]

(23) Ka ni-k-ymmid   kman   ka mi-bangtau ta   neni, ka
        Ka ni-k-immed   k-m-an   ka mi-vangtaw ta   naini, ka
        and eat-all:AF.PERF eat:AF.NEUT and full-stomach NOM 3PLN, and

        ni-ara   nein ki   na pyhpi ka ni-k-da   kman,
        ni-ara   nain ki   na pihpi ka ni-k-da   k-m-an,
        take:OF.PERF 3PLA OBL PL piece LIN eat-remain:OF.PERF eat:AF.NEUT,

    sat kytti-æn 'æb' ki rouha ka la'i-æn ta   matt'moej.
    sat kitti-an   'æb ki ruha   ka la'i-an   ta   ma-tmuy
    one ten   and of two   LIN basket NOM full
    'And they did all eat, and were filled: and they took up of the fragments
    that remained twelve baskets full.' (14:20)

(24) tou touk ta   kidi-au,   <u>k-koua-ah-koh</u>   <u>kman</u>
     tu   tuk   ta   kidi-au,   k-kua-ah-kuh   k-m-an
     LOC near NOM time-1SGG, eat-come:AF.FUT-1SGN eat:AF.NEUT

    ymhou-an [ki   joukko] ki Hali-daæu-en lam ki
    imhu-an [ki   yukku] ki Hali-daæu-en lam ki
    2SGO   [OBL sheep] of passover   with OBL

    Pahtatæutæuugh-en-au
    Pah-ta-tæu-tæuugh-en-au.
    disciple.PL-1SGG
    'My time is at hand; I will keep the passover at thy house with my disciples.' (26:18)

(25) Jrou ka mad-oung-ato <u>ni-k-lam</u>   meirrong ki sat
     Iru   ka madung-atu ni-k-lam   m-irung   ki sat
     when LIN evening-EMPH eat-with:AF.PERF sit:AF.NEUT OBL one

    kytti-æn 'æb ki   rouha
    kitti-an 'æb ki   ruha
    ten   and OBL two
    'Now when the even was come, he sat down with the twelve.' (26:20)

Pæ'æ̈- 'giving action' somewhat resembles its corresponding full verb *pihæ-,*
phæ-, p'æ-, p'hei-, and so forth, which may be cognate to Kanakanavu *mu-'a-vua?,*
Saaroa *mu-a-vura* 'give', and may ultimately, perhaps, come from **beRey*:

(26) Mang ta   <u>p'hei-anei</u>-moumi jau-an, ka <u>pæ'æ̈-tounnoun-eimau</u>
     Mang ta   p'hæy-an-ei-mumi yau-an, ka pæ'æ̈-tunnun-ei-mau
     what NOM give:LF.FUT-2PLA 1sgE, and deliver:LF.FUT-1SGA

    tyni-æn ymoumi-æn?
    tini-an imumi-an
    3SGO   2PLO
    'What will ye give me, and I will deliver him unto you?' (26:15)

    In some cases, it happens that not a prefix but the first syllable of the main verb is
carried over to the auxiliary verb.

---

9. There are two full words for 'eating': *kman* (< *kan-*) and *mavok* (< *avok-*). Semantic differences which are expected to exist are not clear. *Kman* requires the lexical prefix *k-*, whereas *mavok* seems to take a prefix *mi-* as in *mi-sâl mavok* 'eat together' (9:11; 14:9), *mi-sal-ah mavok* (AF.FUT) (24:49), and *ni-mi-ymmid mavok* '(they) did all eat' (15:37). Although there is only one full word for 'drinking', *myt* (< *yt-*), there are two prefixes: *hou-* and *mou-*, as in *hou-baæu-au myt hmou-lam* 'drink it new with (you)' (26:29), and *moulpough myt* 'can drink' (20:22), *mou-sâl-ah myt* 'will drink together' (24:49). Here again possible semantic differences between the two are not known.

(27) Alei ka ka'æuloung, ka mamado   Tama-kavarau, assi
Aley ka ka'æulung, ka ma-madu   Tama-ka-varaw, assi
for LIN man,   LIN oneself:AF.NEUT sinner,   not

<u>hmalpough</u> <u>hmavoung</u> ki na varau ki rarouma-râ
h-m-a-lpouh h-m-avung ki na varaw ki ra-ruma-râ
can:AF.NEUT pay:AF.NEUT OBL pl sin of other-person-EMPH
'Because a man who is himself a sinner cannot pay for someone else.'
(F.13, 22a, p.53)

(28) ka Raraman-oho ka <u>kmi-dung</u>   <u>kmytta.</u>
ka Ra-rama-n-uhu ka k-m-i-deng   k-m-itta
and father-2SGG LIN see-secret:AF.NEUT see:AF.NEUT

pæ'æ̈-pænæh-æh pæ'æ̈-vli   ymhou-an-da
pæ'æ̈-pænæh-æh pæ'æ̈-vli   imhu-an-da
give-openly:AF.FUT give-reward:AF.NEUT 2SGO-EMPH.
'and thy Father which seeth in secret himself shall reward thee openly.' (6:4)

In (28), we are not sure if -*dung* is an auxiliary verb because of the paucity of examples.[10] But judging from the words equivalent to it in the Tsouic languages, it is most likely to be an auxiliary verb.[11] At any rate, it is extremely interesting to see the emergence, as it were, of pseudo-prefixes. Unfortunately, however, there are no other inflected forms found so far in our Siraya data currently available, such as **ha-lpough-en, *ki-dung-auh*, and so forth, but once it happened that such inflected forms were produced by analogy with other forms, we would have seen the birth of new prefixes.[12]

There is another type of prefix that presents an interesting case:

(29) ti-mamang ta <u>mas-aou-moei</u> mararau   ki Ina,
ti-ma-mang ta ma-sau-mui   m-araraw   ki Ina,
anybody   NOM more-wish:AF.NEUT look:AF.NEUT OBL woman,

dæh-kahâ-to   rbou-rbo tyni-æn tou tyntin tyn.
dæh-kaha-atu   rebu-rebu tini-an tu   tintin tin.
commit:AF.NEUT-EMPH adultery 3SGO LOC heart 3SGG.
'whosoever looketh on a woman to lust after her hath committed adultery with her already in his heart.' (5:28)

Here *mas-aou-moei* can be interpreted as consisting of three elements, *ma-saou-moei*. The first element *ma-* is a copy of the first syllable of *m-ararau* 'look at (AF neutral form)', the result of prefix harmony. The third element *moei-* is the stem-root mean-

---

10. I have found only one more case: *itou-dung* 'be hidden, be in secret' (5:14).

11. For instance, in Saaroa, *-sakave* meaning 'do in secret' takes various lexical prefixes as follows: *ma-a-saka-sakave* 'drink in secret', *ku-a-saka-sakave* 'eat in secret', *mu-a-saka-sakave* 'go/come in secret', etc.

12. The following two examples might be related to a possible prefix *ki-* 'seeing': *ni-ky-heil-an* 'pleased' (14:6) (cf. *gmail* 'witness' [23:31], *hgeil-in* 'testimony' [10:18]); *ky-hau-ei-en* 'vision' (17:9) (cf. *mäi-hga-hauwei-ah* 'appear' [6:18]). But there are too few examples, and it is hard to say anything certain.

ing 'want, wish, will' (see 2.1.3). The second element *saou-* is problematic. There are many examples of the lexical prefix *saou-* meaning 'swear' (see item 32 in the list given in section 3), but semantically the prefix we are concerned with here cannot be this particular one. It seems more likely that it is the stem of the auxiliary verb *-saoun* meaning 'do more' (see (14) above), but without its final consonant *-n*. If it is not a case of a simple misspelling or misprint, there are two possible ways in which this prefix developed: (i) there was a rule to attach the whole stem excluding its final consonant; (ii) there was a rule to attach the whole stem, but later its final consonant was lost in the weak position of a prefix. At any rate the three cases above might be remnants of an embryo or a germ of possible lexical prefixes which would have developed into full prefixes had the language continued to exist.

APPENDIX: SIRAYA PRONOUNS

	NOMINATIVE		GENITIVE	AGENTIVE
	INDEPENDENT	BOUND		
1SG	jau	(-)koh/kog	-au	-au/-nau/-mau
2SG	ymhou	(-)kow	-oho(u)/-'uho/-noho	-oho(u)/-hou/-mhou
3SG	teni		tyn (=)	tyn
1PL.INCL		-kytta	-eta	-mytta
1PL.EXCL	ymi-æn/ymi-an	-kame	-ian/-jan	-niæn/-miæn
2PL	ymoumi/eimoumi	-kamou	-oumi/-noumi	noumi/-moumi
3PL	neini/neni		nein (=)	nein

	OBLIQUE/EMPHATIC	POSSESSIVE
1SG	jau-an/jauan/jauwan	â-jau
2SG	ymhou-an/ymhouan	â'mouhou/â'mhou
3SG	tyni-æn/tyni-an	(á-teni-ah)
1PL.INCL	ymit'æn	
1PL.EXCL	ymiæn-æn/ymi-an-an/ymianæn	
2PL	ymoumi-æn/ymoumi-an	(âmoumi-ah)
3PL	neini-æn/neiniæn/neini-an	â-neini/â-neni

REFERENCES

→ Adelaar, K. Alexander. 1997a. Grammar notes on Siraya, an extinct Formosan language. *Oceanic Linguistics* 36: 362–397.
———. 1997b. Retrieving Siraya phonology: A new spelling for a dead language. Paper presented at the 8th International Conference on Austronesian Linguistics, Taipei.
———. 1994. Grammar notes on Siraya, an extinct Formosan language. Paper presented at the 7th International Conference on Austronesian Linguistics, Leiden.
Asai, Erin. 1939. *Gravius's Formulary of Christianity in the Siraya language of Formosa.* Memoirs of the Faculty of Literature and Politics, Taihoku Imperial University, vol. 4, no.1. Taihoku (= Taipei): Taihoku Imperial University.

Ezard, Bryan. 1978. Classificatory prefixes of the Massim Cluster. In *Second International Conference on Austronesian Linguistics: Proceedings*, ed. by S. A. Wurm and L. Carrington, 1159–1180. Pacific Linguistics C-61. Canberra: Australian National University.

Gravius, Daniel. 1661. *Het Heylige Euangelium Matthei en Johannis. Ofte Hagnau ka D'llig Matiktik, Ka na sasoulat ti Mattheus, ti Johannes appa. Overgeset inde Formosaansche tale, voor de Inwoonders van* Soulang, Mattau, Sinckan, Bacloan, Tavokan, *en* Tevorang. t'Amsterdam, 1661. Reprinted in *The Gospel of St. Matthew* in Formosan (Sinkang Dialect) by Rev. Wm. Campbell. London: Trübner & Co., 1888. Reproduced in 1996 from SMC Publishing Inc., Taipei.

———. 1662. *Patar ki Tna-'msing-an ki Christang, ka Tauki-papatar-en-ato tmaeu'ug tou Sou ka Makka Si-Deia. Ofte / 't Formulier des Christendoms, Met de Verklaringen van dien, Inde Sideis-Formosaansche Tale.* t' Amsterdam, by Michiel Hartogh. Photo-reprint by Asai (1939).

Murakami, Naojiro. 1933. *Sinkan manuscripts.* Memoirs of the Faculty of Literature and Politics, Taihoku Imperial University vol.2, no.1. Taihoku (= Taipei): Taihoku Imperial University.

Nojima, Motoyasu. 1996. Lexical prefixes of Bunun verbs. *Journal of the Linguistic Society of Japan* 110:1–27.

Osumi, Midori. 1995. *Tinrin grammar.* Oceanic Linguistics Special Publication No. 25. Honolulu: University of Hawai'i Press.

Tsuchida, Shigeru. 1976. *Reconstruction of Proto-Tsouic phonology.* Study of Languages & Cultures of Asia & Africa, Monograph Series No.5. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.

———. 1990. Classificatory prefixes of Tsou verbs. *Tokyo University Linguistic Papers '89*, 17–52.

———. 1996. Personal pronouns in Siraya (Formosa). In *Reconstruction, classification, description: Festschrift in honor of Isidore Dyen* , ed. by Bernd Nothofer, 231–247. Hamburg: Abera-Verlag Meyer & Co.

van der Vlis, C. J. 1842. Formosaansche woorden-lijst, volgens een Utrechtsch Handschrift. *Verhandelingen van het Bataviaasch Genootschap* 18:437–488. Reprinted with an English translation in Murakami 1933: 154–203.

9: WHAT PART OF SPEECH IS *NÍI* 'THIS' IN THAI?

Amara Prasithrathsint
Chulalongkorn University

1. Introduction[1]

The word *níi* 'this' in Thai is apparently a simple word, but its classification is rather problematic. First, there has not been a consensus as to what part of speech or word class *níi* belongs to. It has been classified as a determinative adjective/adverb by Upakit-silapasarn (1937:92);[2] a demonstrative adjective/adverb by the Royal Institute Dictionary (1982:445); a demonstrative adjective by Haas (1964:274); just a demonstrative (subclass of substantives) by Noss (1964:102);[3] a deitic word, which is a subclass of noun substitutes (or pronouns) by Bandhumedha (1984:14, 34);[4] a pronoun and a determiner by Panupong (1989:66, 72); and just a determiner by Savetamalya (1989:37, 53, 55, 84, 103), Indrambarya (1994:179, 201, 213; 1996:514), Starosta (1994), and Warotamasikkhadit (1996:24).[5] Second, the word *níi* has been used by some syntacticians who consider it to be a determiner as a criterion to decide the word class of such ambivalent words as *nay* 'in, inside', *bon* 'on, top', and so on. They argue that such words should be classified as nouns, not as prepositions, as most linguists would classify them, because these words can have *níi* as their dependent (e.g., *nay níi* 'this inside'), the same as other common nouns (e.g., *bâan níi* 'this house'). It seems to me that using *níi* as the

---

1. This study is part of my research project on "Parts of speech in Thai: A syntactic analysis based on a two-million word corpus of current Thai," supported by the Thailand Research Fund (TRF). I am grateful to Videa De Guzman and an anonymous referee for their constructive comments, which have helped me develop the final version of this essay. I would also like to thank Stanley Starosta and Pranee Kullavanijaya for sharing with me their knowledge of Thai syntax, and thoughts on how to approach syntactic problems.

2. Upakitsilapasarn classifies *níi* as *níyamáwísèet*, meaning 'determinative adjective/adverb'. The morpheme *níyamá-* means 'pointing, determining', and *wísèet* is a form that qualifies a noun or modifies a verb, or an adjective or adverb.

3. According to Noss (1964:102), a demonstrative is "any substantive [that] is invariably the last free lexeme in any nonpredicate construction in which it occurs (e.g., enumerations and endocentric substantive expressions). Demonstratives by themselves do occur in most of the typical substantive positions—topic, subject, object, complement—but of course are never heads. Their chief function, however, is that of modifier, hence they are a special case of complementatives." What Noss seems to point out here is that *níi* is a demonstrative that functions as a modifier of a head noun, or occurs as sister of the head of a noun phrase, which is in the topic, subject, object, or complement position. Noss regards *níi* as a base lexeme from which *nîi* 'here, this' (with a falling tone) derives. He calls *nîi* a derivative of *níi*.

4. Bandhumedha names the class to which *níi* belongs *khambòŋchíi*, meaning 'deitic word', which is a subclass of the major class *khamtheen* 'substitutes' (= pronouns). However, according to her, *níi* normally functions as a noun modifier.

5. Warotamasikkhadit names the class to which *níi* belongs *níyamálák*, meaning 'feature that determines' or 'determiner'.

criterion for determining nouns is problematic, because it can be observed that *níi* has a number of different distributions. In addition to occurring after nouns, as in (1–4), *níi* is also found to occur after nonnouns, as in (5–8).

(1)	naǎŋsǐi níi	'this book'		(5)	kɔ̀ɔn níi	'before this'
(2)	rooŋrian níi	'this school'		(6)	càak níi	'from this'
(3)	sàpdaa níi	'this week'		(7)	tâŋtèɛ níi	'since this'
(4)	prayòok níi	'this sentence'		(8)	taam níi	'according to this'

Superficially, the constructions in (1–4) and that in (5–8) are identical: WORD + *níi*. However, the words *kɔ̀ɔn* 'before', *càak* 'from', *tâŋtèɛ* 'since', and *taam* 'according to,' as in (5–8), never behave like nouns. That is, they cannot be head of a noun phrase that functions as the subject or object of a sentence. Thus, it is clear that *níi* in (5–8) is not in the same word class as *níi* in (1–4).

Here, we are faced with a dilemma in word classification. Should we regard *níi* as one word or two separate words of the same form? Taking the latter position, I argue in this essay that there are two *níi* words in Thai. One is a determiner, and the other is a demonstrative pronoun. I will also suggest that one should not use *níi* as a criterion for determining the word-class of its sister, because its own status is ambiguous.

2. Reasons for Analyzing *níi* as Two Words

If I choose to analyze *níi* as one word, then I will have to explain why *níi*, which normally occurs with nouns, can also occur with prepositions. One explanation would be that *níi* follows a preposition only in the surface structure, and that in the deep structure it follows a noun, which is its sister constituent. In this scenario, both *níi* and the noun form a noun phrase that is sister constituent of a prepositional-phrase construction. Then the noun is deleted by a transformational rule. What is left is a preposition + *níi* on the surface. This model of grammar is too powerful, and we will be faced with the problem of recoverability, one of the greatest weaknesses of Transformational Grammar. Another way out is to regard a preposition + *níi* construction (e.g., *kɔ̀ɔn níi* 'before this'), which looks like an exception, as a compound word instead of as a prepositional phrase, which poses a problem of irregular distribution for *níi*. In this case, we can maintain the principle of one syntactic distribution for one word; that is, that *níi* always follows a noun, with which it forms a noun-phrase construction. However, forcing a phrase to become a word poses certain problems relating to the intuition, and to productivity. Indeed, it is hard for a native speaker of Thai to accept that all strings of preposition + *níi* are compound words. Moreover, we will have to count such constructions with other deitic words similar to *níi* as compound words, too. That would go against the principle of grammatical generalization. The other solution is to allow several different syntactic environments for one word class or to adopt the notional or semantic approach to word classification. Thus, in whatever environment, *níi* would be considered to be just one word because of a common deitic meaning. This approach has many drawbacks and causes many inconsistencies in word classification, as evidenced in Traditional Grammar.

In this study, I adopt the principle of word classification proposed by Starosta (1994), which uses syntactic criteria (that is, distribution and cooccurrence) to determine membership in syntactic word classes. This approach has several advantages, as pointed out by Starosta. The most important one, in my opinion, is that it is not as elusive and subjective as a notional approach, and can also be supported by data of real language use. Thus, since *níi* is found to have two different distributions, it should be classified as two homophonous members of two different classes.

3. The Determiner *níi*

Following Starosta (1994), who emphasizes the importance of taking into consideration cross-linguistic generality, I classify the *níi* that is a constituent of a noun phrase construction as a determiner, rather than as an adjective. According to him, a class of adnominal modifiers that always occur at the periphery of an NP and that express deitic meanings are better analyzed as determiners than adjectives. The determiner *níi* in Thai always occurs as the right-most constituent of a noun phrase construction.[6] The following are examples of noun phrases containing the determiner *níi* 'this'. Note that *níi* is always at the end of NP constructions, no matter how long they are.

   (9) mɨaŋ **níi**
         city   this
         'this city'

  (10) mɨaŋ yày **níi**
         city   big   this
         'this big city'

  (11) mɨaŋ yày thîi   mii   khon   nǎa-nèn **níi**
         city   big   which have people crowded this
         'this crowded big city'

  (12) mɨaŋ yày thîi   mii   khon   nǎa-nèn lɛ́   aakàat rɔ́ɔn càt **níi**
         city   big   which have people crowded and weather hot very this
         'this big city which is crowded and very hot'

As a determiner, *níi* marks specificness on the head noun. It has a deitic meaning of 'being close to the speaker'. There are three other words in Thai that have the same high tone and the same syntactic behavior as *níi*, and very similar deitic meanings. They are *nán* 'that', *nóon* 'over there', and *núun* 'way over there'. These are also classified here as determiners.

---

6. Noss (1964:102) also takes note of this. He labels *níi* as a demonstrative, which he defines as "the last lexeme in any nonpredicate construction."

4. The Demonstrative Pronoun *níi*

Evidence shows that *níi* can also occur with words other than nouns, forming a construction that is superficially identical with a noun phrase, but that is actually not a noun phrase, as in the following examples taken from Thai newspapers and journals. The constructions with *níi* are in italics.[7]

(13) khamphûut *chên*   *níi* mây troŋ kàp khwaam pen ciŋ
     word   such as (like) this not straight with NOM   be real
     'Words like this (such words) do not correspond to reality.'

(14) *nɔ̀ɔkcàak níi* rát khuan sâaŋ sathiàanraphâap khɔ̌ɔŋ khâa ŋən   bàat
     besides this state should create stability   of   value money baht
     'Besides this, the state should create the stability of the *baht* value.'

(15) khwaam mây thâwthiam kan   khɔ̌ɔŋ kaan kracaay raaydây
     NOM   not equal   each other of   NOM distribute income

     cà   tham hây kə̀ət   wíkrìttakaan thîi runrɛɛŋ *kwàa níi*
     will make give occur crisis   that strong   than this
     'The inequality in the distribution of income will cause a crisis that is more serious than this.'

The words *chên, nɔ̀ɔkcàak,* and *kwàa* in (13), (14), and (15), respectively, are prepositions. Following Upakitsilapasarn (1937:97) and Starosta (1994), I define a preposition as a word that forms an exocentric construction with a phrase the head of which is a noun, pronoun, or verb. An exocentric construction is defined as one in which all the constituents are obligatory. Also, particularly in Thai, prepositions cannot be negated; that is, they are never preceded by *mây* 'not' (Indrambarya 1994:41, Intratat 1998:54–56). Examples (16–18) show that the words *chên, nɔ̀ɔk-càak,* or *kwàa* are prepositions; they precede a noun phrase (NP), with which they form a prepositional phrase (PP)—an exocentric construction—and they cannot be preceded by *mây* 'not'.

(16) PP[P[chên]P NP[nók]NP]PP   (*mây chên . . .)
     such as   bird   not such as . . .
     'such as birds'

(17) PP[P[nɔ̀ɔkcàak]P NP[mahǎawítthayaalay nay prathêet]NP]PP
     besides   university   in   country
     'besides universities in the country'
                                     (*may nɔ̀ɔkcàak . . .)
                                       not besides. . .

(18) PP[P[kwàa]P NP[sǎam méet]NP]PP   (*mây kwàa. . .)
     than   three meter   not than . . .
     'than three meters'

---

7. Abbreviations used include CL, classifier; NOM, nominalizer; and PRT, particle.

The noun phrases in (16–18) and the word *níi* in (13–15) have the same status. With the prepositions preceding them, they form PP constructions. I classify *níi* in this syntactic environment as a demonstrative pronoun.[8]

The following minimal pairs of sentences containing P + *níi* (prepositions + *níi*) and P + NP (prepositions + noun phrases) will serve as more evidence supporting the analysis of *níi* as a pronoun. The prepositions are *troŋ* 'straight at', *càak* 'from', *taam* 'following; according to', *kwàa* 'than', *kɔɔn* 'before', and *tâŋtɛɛ* 'from; since'. The contrasting prepositional phrases (P + *níi* vs. P + NP) are italicized.

(19) P + *níi*
karunaa khĭan khamtɔ̀p *troŋ   níi*
please   write answer straight at this
'Please write the answer right here.'

P + NP
karunaa khĭan khamtɔ̀p *troŋ   chɔ̂ŋwâaŋ thîi hây wáy*
please   write answer straight at blank   that give keep
'Please write the answer in the blank provided.'

(20) P + *níi*
càak níi pay raw tôŋ   yùu yàaŋprayàt
from this go we must live economically
'From this (time) on we must live economically.'

P + NP
càak dian nâa pay raw tôŋ   yùu yàaŋprayàt
from month next go we must live economically
'From next month on we must live economically.'

(21) P + *níi*
thâa khun tham *taam   níi* khun cà   pràsɔ̀p phŏnsămrèt
if   you do   following this you will meet   success
'If you follow this, you will succeed.'

P + NP
thâa khun tham *taam   khamnɛ́nam khɔ̌ɔŋ chăn* khun cà
if   you do   following advice   of   me   you will

pràsɔ̀p phŏnsămrèt
meet   success
'If you follow my advice, you will succeed.'

(22) P + *níi*
chăn hây khun mâak *kwàa níi* mây dây
I   give you much than this not can
'I cannot give you more than this.'

_____

8. Demonstrative pronouns are a subclass of pronouns, which are regarded here as a subclass of nouns.

P + NP

chăn hây khun mâak *kwàa nɨŋ láan   bàat* mây dây
I   give you   much than one million baht not can
'I cannot give you more than one million baht.'

(23) P + *nɨ́ɨ*
kɔ̀ɔn   nɨ́ɨ raw yùu yàaŋ-sabaay
before this we live comfortably
'Before this (time), we lived comfortably.'

P + NP
kɔ̀ɔn   sǒŋkhraam raw yùu yàaŋ-sabaay
before war   we live comfortably
'Before the war, we lived comfortably.'

(24) P + *nɨ́ɨ*
tâŋtɛ̀ɛ   nɨ́ɨ thúk yàaŋ khoŋ cà dii khɨ̂n
from   this every thing likely will good more
'From this time on everything will probably get better.'

P + NP
tâŋtɛ̀ɛ   piimày   thúk yàaŋ khoŋ cà dii khɨ̂n
from   New Year every thing likely will good more
'From New Year on everything will probably get better.'

5. The Limited Distribution of the Demonstrative Pronoun *nɨ́ɨ*

It should be noted that the occurrence of the demonstrative pronoun *nɨ́ɨ* is consider-
ably limited, both quantitatively and qualitatively, when compared to that of the deter-
miner *nɨ́ɨ*. Quantitatively, it seems to occur less frequently than the determiner. I tried
counting all the tokens of the occurrences of *nɨ́ɨ* in a sample text of ten pages and
found that approximately 20 percent of all the tokens were pronouns and approxi-
mately 80 percent were determiners.

Qualitatively, whereas the determiner *nɨ́ɨ* is found to occur with all kinds of nouns,
the demonstrative pronoun *nɨ́ɨ* has a limited distribution in two respects. In the first
place, it seems to be compatible only with certain prepositions, with which it forms a
phrase that usually signifies time and some other abstract concepts. Examples of
these prepositions are *càak* 'from', *tâŋtɛ̀ɛ* 'since', *kɔ̀ɔn* 'before', *taam* 'according to',
nɔ̂ɔkcàak 'besides', and *kwàa* 'than'. In addition, Panupong (1989:67) points out that
nɨ́ɨ tends to occur with some prepositions indicating space, such as *nay* 'in', *bon* 'on',
tâay 'under', *khâaŋ* 'beside', *troŋ* 'at', and *thɛ̌ɛw* 'around'. In the second place, the
distribution of the demonstrative pronoun *nɨ́ɨ* is also limited in terms of its extremely
rare appearance as subject and object of a sentence. Indeed, I have not found any real
use of *nɨ́ɨ* as such. However, Upakitsilapasarn (1937:81) gave an example of a sen-
tence in which *nɨ́ɨ* is the subject, shown in (25). He used this example to make the
point that even though he classified it as a demonstrative adjective/adverb, he

accepted the possibility that it could sometimes function as a pronoun. Noss (1964:102) also gave an example in which *níi* is the subject, shown in (26).

(25) **níi** khɔ̌ɔŋ khray
    this of   who
    'Whose is this?'

(26) **níi** bâan yày[9]
    this house big
    'This is a big house.'

It seems that the reason for the limited occurrence of *níi* as subject and object of a sentence is not grammatical but pragmatic. In Thai, subjectless and particularly object-less sentences are frequently used, especially in consultative and informal styles or in situations in which the speaker and the hearer understand each other so well that they can leave many words unsaid, particularly those representing given information. As pronouns represent given information, they are rarely used in most environments, except in exocentric constructions, in which their occurrence is obligatory, or for special emphasis. In short, *níi* could occur in the subject and object positions, but it normally does not because Thai ways of speaking require that the subject and the object not be mentioned when the context is clear. Examples (27a) and (28a), in which there is no subject or object, are commonly used sentences, whereas (27b) and (28b), which contain *níi* 'this' or *man* 'it' in subject or object position, are rarely used.

	COMMON		RARE[10]	
(27) a.	sǎmkhan mâak	b.	níi/man sǎmkhan mâak	
	important very		this/it important very	
	'This/it is very important.'			
(28) a.	yàa hây khray hěn ná	b.	yàa hây khray hěn níi/man ná	
	don't give anyone see PRT		don't give anyone see this/it PRT	
	'Don't let anyone see this/it.'			

Regarding the occurrence of *níi* with only certain prepositions, the reason for this restriction seems to be semantic. The demonstrative pronoun *níi* denotes something close to the speaker, and also seems to connote abstractness or intangibility. It is mostly used to substitute for nouns that refer to abstract things, such as space, time, ideas, arguments, reasons, and so forth. Probably, this is an explanation for its nonexistence with some prepositions that normally occur with concrete nouns, such as *dûay* 'with' (signifying 'instrument' *(*dûay níi)*, *hây* 'for', signifying 'beneficiary' *(*hây níi)*, *kɛ̀ɛ* 'to', signifying 'recipient' *(*kɛ̀ɛ níi)*, and so forth.

Another explanation for the limited distribution of the demonstrative pronoun *níi* may have to do with the fact that Thai has another deitic that is phonologically and semantically similar to *níi*, but has a falling tone—*nîi*, meaning 'here' or 'this'. This

---

9. This sentence and the one in (25) are verbless. They are called "verbless sentences" by Warotamasikkhadit (1969:74), and "equation constructions" by Noss (1964:75). Indrambarya (1994:105) regards the sister constituent of *níi* in such constructions as "nonverbal predicates."

10. I asked a few native speakers of Thai how (27b) and (28b) sounded to them. All found the two sentences odd. However, when asked whether they could say them, they said they could, but that they would normally use just the verb without *níi* or *man* in either of the sentences.

deitic *nîi* has a syntactic distribution that is almost complementary to that of *níi*. Indeed, it is striking that in whatever environment in which *níi* cannot be used, the other—*nîi*—seems to fit very well. The difference between *níi* and *nîi* is expanded on below.

6. The Words *níi* and *nîi* in Thai

Just like *níi*, *nîi* is also ambiguous. That is, the same form represents members of two different word classes: determiners and pronouns. Semantically, *nîi* is also deitic, and, like *níi*, it has three counterparts: *nân, nôon,* and *nûun,* meaning 'that/ there', 'over there', and 'way over there', respectively. However, there is a contrast in the distribution of the *níi* words and the *nîi* words. In the case of *níi*, the demonstrative pronoun *níi* has a limited distribution, whereas the determiner *níi* does not, as already shown. By contrast, in the case of *nîi*, the determiner *nîi* has a limited distribution, whereas the demonstrative pronoun *nîi* does not. The following examples show the contrasts in the distribution of *níi* and *nîi*. Sentences (29) and (30), in which *nîi* is subject and object, respectively, are more acceptable than (31) and (32), in which *níi* has parallel roles.

(29) *nîi*   lék   pay sămràp chăn
     this small too for   me
     'This is too small for me.'

(30) waaŋ *nîi* wáy bon tó   nòy
     place this keep on table PRT
     'Please put this on the table.'

(31) ?*níi* lék pay sămràp chăn
     this small too for   me
     'This is too small for me.'

(32) ?waaŋ *níi* wáy bon tó   nòy
     place this keep on table PRT
     'Please put this on the table.'

As a constituent of a PP, unlike *níi*, *nîi* seems to occur with prepositions that normally precede nouns signifying concrete things, such as *dûay* 'with', *kàp* 'together with', *phîa* 'for', and so forth. However, there is overlap, because both *nîi* and *níi* can occur with the same prepositions but convey different meanings. The former makes the prepositional phrase convey a concrete meaning, especially about locality, while the latter conveys the concept of time or abstract things. The following are examples of prepositional phrases with *nîi* (33a–35a) and *níi* (33b–35b) occurring with the same prepositions. Notice the difference in the meanings of the phrases.

(33) a. càak *nîi*
     from here
     'from here'

    b. càak *níi*
       from this
       'from this (time)'

(34) a. tâŋtὲε *nîi*
     from   here
     'starting from here'

    b. tâŋtὲε   *níi*
       from/since this
       'from/since this time'

(35) a. nɔɔkcàak *nîi*
     besides   this
     'besides this particular
     thing right here'

    b. nɔɔkcàak   *níi*
       besides   this
       'besides this (abstract thing)'

As a determiner, *nîi* seems to be used less frequently than *níi*. Moreover, its distribution is more restricted. Whereas *níi* occurs with all kinds of nouns, *nîi* does not. The co-occurrence of *nîi* and a noun classifier[11] is not permitted, but that of *níi* and such a word is always grammatical. The following examples show contrast between the two determiners *nîi* and *níi*. Note that the NP constructions (in italics) with *nîi* in (36a) and (37a) are ungrammatical, but those (also in italics) with *níi* in (36b) and (37b) are grammatical.

(36) a. **naŋsɨ̆ɨ lêm nîi   mây nâasŏncay*
     book   CL   this/here not   interesting
     'This book is not interesting.'

    b. *naŋsɨ̆ɨ lêm níi   mây nâasŏncay*
     book   CL   this   not   interesting
     'This book is not interesting.'

(37) a. **sɨ̂a tua nîi   lék   pay sămràp chăn*
     shirt   CL   this/here small   too for   me
     'This shirt is too small for me.'

    b. *sɨ̂a tua níi lék pay sămràp chăn*
     shirt   CL   this small too for me
     'This shirt is too small for me.'

Apart from noun classifers, *nîi* is also found to be incompatible with some other nouns. These include, for example, ones that designate abstract concepts, including time, that normally occur with *níi*, such as *khwaamkít níi* 'this idea', *khwaamrák níi* 'this love', *kaanplìanpleeŋ níi* 'this change', *weelaa níi* 'this time', and *wan níi* 'today'.

Despite the discrepancy between *nîi* and *níi* shown above, the distributions of *nîi* and that of *níi* sometimes overlap. That is, they can be used with the same nouns, but convey different meanings. The following examples of noun phrases with *nîi* (38a–39a) and *níi* (38b–39b) show that the ones with *nîi* signify strong deixis and that those with *níi* convey the meaning of 'specific' as opposed to 'generic'.

(38) a. sɨ̂a nîi   kàw lɛ́ɛw
     shirt here old   already
     'This shirt (the one right here) is already old.'

    b. sɨ̂a níi   kàw lɛ́ɛw
     shirt this old   already
     'This shirt (some specific shirt, not any shirt) is already old.'

(39) a. mahăawítthayaalay nîi   mii-chɨ̂ɨsĭaŋ
     university   this/here have fame
     'This university (the one in which we are now) is famous.'

    b. mahăawítthayaalay níi mii-chɨ̂ɨsĭaŋ
     university   this have fame
     'This university (the one we are talking about) is famous.'

---

11. Noun classifiers are regarded here as a subclass of nouns.

The analysis of the difference between *nîi* and *níi* in this section shows that both are in complementary distribution in some environments but contrast in others. The fact that the use of *nîi* as a demonstraive pronoun is much more frequent than that of the demonstrative pronoun *níi* may serve as an answer to the question as to why the latter has a restricted distribution: most of the function is performed by the former instead.

7. Problems Arising from the Ambiguity of *nîi*

The form *níi* is ambiguous because it can be either a determiner or a demonstrative pronoun. This can cause a problem in syntactic analysis. Because *níi* is mostly found to be a dependent of a noun, some linguists identify any word occurring with it as a noun. This is not valid, because *níi* itself can be a pronoun following a preposition, as shown earlier. Therefore, to regard *nay* 'in' and *bon* 'on' as nouns because they are found to occur with *níi*—*nay níi* and *bon níi*—is not well-motivated, due to the ambiguous status of *níi*.

In addition, it is important to note that the words *nay* and *bon* never occur alone as subject or object of a sentence, as in (40) and (41).

(40) **nay* plɔ̀ɔdphay (41) **chan* chɔ̂ɔp *bon*
    in   safe   I   like   on
    *'In is safe.'   *'I like on.'

When *nay* 'in' and *bon* 'on' are constituents with *níi,* they can occur in the position that is apparently the subject or object position, as in (42) and (43). Of course, one can interpret *nay níi* and *bon níi* as noun phrases meaning 'this inside', and 'this top', respectively. However, one cannot be sure whether they are the subject or object of the sentence, because, in Thai, subjectless and objectless sentences are quite common, as has been mentioned earlier.

(42) *nay*   *níi*   plɔ̀ɔdphay (43) chan chɔ̂ɔp *bon* *níi*
    in   this   safe   I   like   on   this
    'In this (it) is safe.'   'I like (it) up here.'

What is more, a phrase containing *nay* 'in' or *bon* 'on' and other noun phrases cannot function as subject or object of a sentence, as in (44) and (45). This evidence shows that to classify *nay* 'in' and *bon* 'on' as nouns would be poorly motivated.

(44) **nay* *kaan* *tham* *wícay* pɔ̀ɔdphay
    in   NOM do   research safe
    *'In doing research is safe.'

(45) **chăn*   chɔ̂ɔp *bon* r   âakthăan *thîi* *khɛ̌ŋrɛɛŋ*
    I   like   on   foundation that strong
    *'I like on the foundation that is strong.'

As can be seen, the last two sentences containing phrases with *nay* 'in' and *bon* 'on' do not make any sense, because the phrases are in the subject or object position.

If they were in the position of prepositional phrases, they would make the sentence grammatical, as in (46) and (47).

(46) khǎw   ramátrawaŋ   *nay* *kaan*   *tham* *wícay*
      he/she   careful   in   NOM   do   research
      'He/she is careful in doing research.'

(47) khǎw   sâaŋ bâan   *bon* *râakthǎan* *thîi* *khěŋrɛɛŋ*
      he/she   build house   on   foundation   that   strong
      'He/she built a house on a strong foundation.'

Although *nay* 'in' and *bon* 'on' in (42) and (43) are ambiguous (that is, they could be either nouns or prepositions), in (46) and (47) it is clear that they are prepositions. They fit in the class of words that precede a noun phrase to form an exocentric construction and cannot be negated by *mây* 'not'.

8. Conclusion

In the foregoing, I have attempted to provide evidence that the simple form *níi* 'this' in Thai is ambiguous. It can be either a determiner or a demonstrative pronoun. I have also pointed out the differences between *níi* and *nîi,* and that although the two forms are extremely similar, they do not seem to be variants of the same form.

It may be interesting to note that my analysis of *níi* as a determiner and as a demonstrative pronoun agrees with Panupong's (1989:66, 72) and also accounts for the observations of the existence of *níi* as a pronoun by Upakitsilapasarn (1937:81) and Noss (1964:102). My study of *níi* in the Sukhothai inscriptions shows that it had a high frequency of occurrence, whereas *nîi* was found only once in the documents (Prasithrathsint 1999). This may lead to the conclusion that *nîi* may have developed from *níi.* At first, *nîi* may have had the restricted function of a pronominal, but later on it may have maintained its own identity while developing another homophonous form for a different function, that of determiner. If this proves true, it will agree with Noss's analysis of *nîi* as a derivative of *níi* (1964:55).

It is safe to say that both *níi* and *nîi* have been in the process of change. The former, which must have originated as a determiner (functioning as a noun modifier), generated a homophonous form whose limited grammatical function is that of a pronoun. The latter, perhaps originally a pronoun, also gave birth to a homophonous form that is used as a determiner. Thus, we see both *níi* and *nîi* having overlapping functions.

What can be inferred from this study is that new members of a part of speech keep coming into being, especially function words. The pronoun *níi* and the determiner *nîi* are good examples of the emergence of new members of these functional categories.

REFERENCES

Bandhumedha, Nawawan. 1984. *Thai grammar (wayyakɔɔn thay).* Bangkok: Rungreung-sarn Karnphim.

Haas, Mary R. 1964. *Thai–English student's dictionary.* Stanford: Stanford University Press.

Indrambarya, Kitima. 1994. Subcategorization of verbs in Thai: A lexicase dependency approach. Ph.D. dissertation, University of Hawai'i.

————. 1996. On impersonal verbs in Thai. In *Pan-Asiatic Linguistics: Proceedings of the Fourth International Symposium on Languages and Linguistics,* vol. 2, 505–521. Bangkok: Mahidol University.

Intratat, Charatdao. 1998. Grammaticalization of verbs into prepositions in Thai. Ph.D. dissertation, Chulalongkorn University.

Noss, Richard B. 1964. *Thai reference grammar.* Washington D.C.: Foreign Service Institute.

Panupong, Vichin. 1989. *The structure of Thai: Grammatical system (khrooŋ sâaŋ khɔ̌ɔŋ phaasǎa thay: rabòp wayyakɔɔn).* Bangkok: Ramkhamhaeng University Press.

Prasithrathsint, Amara. 1999. The syntactic and semantic features of *níi* 'this' and *nán* 'that' in Sukhothai Thai. *Journal of Language and Linguistics* 17.2:69–83.

Royal Institute. 1982. *The Royal Institute dictionary (of Standard Thai).* Bangkok: Aksorn Charoenthat Publishers.

Savetamalya, Saranya. 1989. Thai nouns and noun phrases: A lexicase analysis. Ph.D. dissertation, University of Hawai'i.

Starosta, Stanley. 1994. The identification of word classes in Thai. Revised version of a paper presented at the Fifth International Conference on Thai Studies at the Centre of Southeast Asian Studies, School of African and Oriental Studies in London, July 1993.

Upakitsilapasarn, Phraya. 1937. *Principles of Thai: Analysis of words (làk phaasǎa thay: wáci-wíphâak).* Bangkok: Thai Watanapanich.

Warotamasikkhadit, Udom. 1969. Verbless sentences in Thai. *Linguistics: An International Review* 47:74–79.

————. 1996. *Thai reference grammar (wayyakɔɔn thay nay phaasǎasàat).* Bangkok: Roongphim Chuanphim.

10: MULTIPLE LEXICAL ENTRIES OF *Kɔ̂ɔ* IN THAI

Saranya Savetamalya
Chiangmai University

1. Introduction

The lexical entry of *kɔ̂ɔ* in Thai has long been a mystery in the description of Thai syntax. The occurrence of *kɔ̂ɔ* is varied. It can appear as a single word in a clause such as *kɔ̂ɔ* 'then', or it can combine with other words to form compounds, such as *kɔ̂ɔ taam* 'of no consequence', *kɔ̂ɔ dây* '(it) is all right (to)', *kɔ̂ɔ dii* '(indicate) more than one', *kɔ̂ɔ lɛ́ɛw kan* 'that is it', or *kɔ̂ɔ lɛ́ɛw pay* '(then) let it go, leave it at that'. This study will discuss only the occurrence of *kɔ̂ɔ* as a word and as a component of compounds.

Most linguists working on Thai syntax have analyzed *kɔ̂ɔ* as belonging to the syntactic class of conjunction (Haas 1964, Warotamasikkadit 1963, Panupong 1970, Thonglor 1972, and Bandhumedha 1975). Under a Lexicase dependency grammar analysis, *kɔ̂ɔ* cannot be analyzed as a conjunction, because by definition a conjunction forms an exocentric construction by conjoining two or more phrases from the same major phrase type such as NP, or S, and so forth (Starosta 1988:52). But *kɔ̂ɔ* does not function to combine two or more phrases of the same type, so it cannot be considered a conjunction. Working within the constrained Lexicase dependency grammar framework, it is the purpose of this study to determine the syntactic category of *kɔ̂ɔ*.

2. The Analysis

There are four different lexical entries for *kɔ̂ɔ* in Thai: $kɔ̂ɔ_1$ means 'also', $kɔ̂ɔ_2$ means 'then', $kɔ̂ɔ_3$ means 'well', and $kɔ̂ɔ_4$ means 'surprisingly' or 'it is unexpected'. In Lexicase, each word consists of three basic components, namely, sound, distribution, and meaning. Distribution plays a significant role in distinguishing a single lexical item from other homophonous forms. The general principle in Lexicase in setting up lexical items states that forms that occur in two distinct syntactic environments must be represented as two distinct lexical entries (Starosta 1988:45). In other words, if a form X can occur in syntactic patterns A and B, but some other form Y can occur only in pattern A, but not B, or only in pattern B but not A, the forms X and Y are considered to belong to separate lexical items distinguished by their distinct syntactic distributions.

2.1 $Kɔ̂ɔ_1$ as an Adverb Meaning 'also'

The first subclass of *kɔ̂ɔ* is analyzed as an adverb. There are two syntactic distributions of this *kɔ̂ɔ*. Either it appears in a sentence without any preceding predicative clause or it appears with a preceding predicative clause. In view of this, *kɔ̂ɔ* will be treated as belonging to two different lexical entries, that is, $kɔ̂ɔ_1$ and $kɔ̂ɔ_2$. Both of them are treated here as adverbs according to the arguments that follow. $Kɔ̂ɔ_1$ is discussed first; the discussion of $kɔ̂ɔ_2$ is in 2.2.

Consider the following data:

(1) chăn kɔ̂ɔ, chôop kin
    I   also   like eat
    'It is also true that I like to eat.'

There are two possible interpretations for kɔ̂ɔ, in (1). The first meaning emphasizes
the action of the verb, which means 'Besides doing other things, I also like to eat'.
The second meaning emphasizes the subject, which means 'Besides other people, I
too like to eat'. The semantic ambiguity of sentence (1) would not affect the determi-
nation of the syntactic class of kɔ̂ɔ,. That is, the semantic ambiguity can be recovered
by the use of intonation. It is possible to make use of intonation to clarify pragmatic
ambiguity. There are two ways of assigning intonation; that is, intonation can either
fall on a subject noun or on a verb. If the intonation falls on the verb, we get the first
interpretation, because in a dependency grammar, the verb acts as a proxy for the
whole clause. If the intonation falls on the noun, we get the second interpretation, in
which the scope of kɔ̂ɔ, is only the subject (Starosta 1993).

The reason kɔ̂ɔ, is treated as an adverb is that it behaves syntactically like an
adverb. One of the characteristics of an adverb is that, unlike a verb, it cannot appear
as an answer to a question, as in the following:

(2) khun yàak pay rɨ̆ɨ   mây pay
    you want go   or not not go
    'Do you want to go or not?'

(3) yàak
    want
    'Yes, I want to.'

(4) pay
    go
    'Yes, I (want to) go.'

(5) khun kɔ̂ɔ, cà? pay rɨ̆ɨ
    you also will go or not
    'Are you also going?'

(6) pay
    go
    'Yes, I am going.'

(7) *kɔ̂ɔ,
    also

(8) *cà?
    will

We see that *yàak* and *pay*, being verbs that appear as the heads of a clause, can
answer the question in (2). On the contrary, the question in (5) cannot be answered by
either kɔ̂ɔ, or *cà?*. Only *pay* can answer this question. The reason is that kɔ̂ɔ, and *cà?*

are not heads of a clause, whereas *pay* is. Thus, they cannot be analyzed as verbs in the same way that *pay* can. Both *kɔ̂ɔ₁* and *cà?* are treated here as adverbs.

The second characteristic of an adverb is that it cannot be negated. The following data illustrate that *nâa* 'is likely to', *yàak* 'want', *khuan* 'should', and so forth, which are all verbs, can be negated with *mây* 'not'. On the other hand, *kɔ̂ɔ₁* 'also', *cà?* 'will', and *mây* 'not' cannot be negated with *mây* 'not'. These elements are something other than verbs. They are therefore treated as adverbs.

(9) fon   mây   nâa   cà?   tòk
    rain   not   likely   will   fall
    'It is unlikely to rain.'

(10) chăn   mây   yàak   tham
     I   not   want   do
     'I do not want to do (it).'

(11) khun   mây   khuan   wâaynáam
     you   not   should   swim
     'You should not swim.'

(12) *fón   mây   kɔ̂ɔ₁   cà?   tòk
     rain   not   also   will   fall

(13) *chăn   mây   cà?   tham
     I   not   will   do

(14) *chăn   mây   mây   wâaynáam
     I   not   not   swim

From the data (12–14), it can be seen that *kɔ̂ɔ₁* has the same distribution as *cà?* and *mây* in that none of the three can be negated. *Kɔ̂ɔ₁*, *cà?*, and *mây* fall into the same class, namely adverbs. The position of *kɔ̂ɔ₁* as an adverb is fixed. That is, it only appears after an overt subject, if there is one, and it always appears before the first commanding verb, as shown in the following:

(15) chăn   kɔ̂ɔ₁   chôop   pay   thîəw
     I   also   like   go   sightseeing
     'I also like to go sightseeing.'

(16) *chăn   chôop   kɔ̂ɔ₁   pay   thîəw
     I   like   also   go   sightseeing

(17) *chăn   chôop   pay   kɔ̂ɔ₁   thîəw
     I   like   go   also   sightseeing

(18) *chăn   chôop   pay   thîəw   kɔ̂ɔ₁
     I   like   go   sightseeing   also

Within a Lexicase analysis, the tree structure of a clause containing *kɔ̂ɔ₁* with an adverb function as in (15) may be illustrated as follows:

(15) a.

```
                            chôop
                            like
                            3ndex
   chan        kɔ̂ɔ₁        +V          pay
   I           also        +prdc       go
   index       2ndex       1[+Nom]     4ndex
   Nom         +Adv        1[+PAT]     +V
   PAT                     4([+V])     +prdc       thîəw
                           2([+Adv])   −fint       sightseeing
                           3<2([+Adv]) 5([+V])     5ndex
                           4[−fint]                +V
                           3<4[−fint]              +prdc
                                                   −fint
```
'I also like to go sightseeing.'
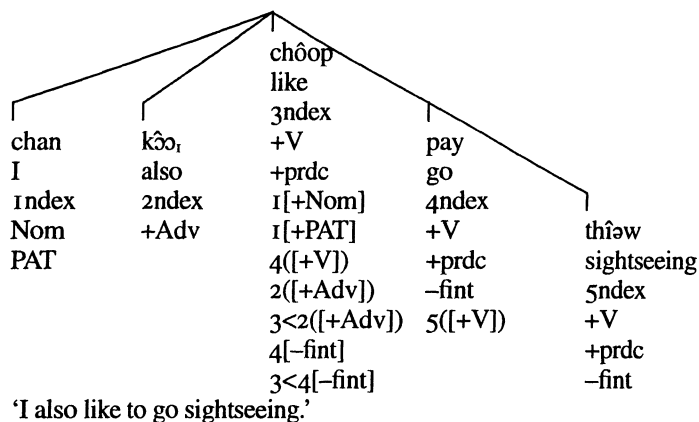
This section has discussed the distribution of *kɔ̂ɔ₁* meaning 'also', which appears without a preceding predicate. The next section discusses *kɔ̂ɔ₂* meaning 'then', which also functions as an adverb but appears with a preceding predicate.

2.2 *Kɔ̂ɔ₂* as an Adverb Meaning 'then'

The lexical entry of *kɔ̂ɔ₂* meaning 'then' also bears the category adverb. It differs from *kɔ̂ɔ₁* syntactically in that it requires a preceding predicate. When it appears, it refers to a previous discourse, and its lexical meaning is 'sequential'. *Kɔ̂ɔ₂* introduces the main clause in a conditional sentence structure. This characteristic function applies only when there is no overt subject (cf. [20] and [21] where its occurrence is identical to *kɔ̂ɔ₁*, that is, between the subject and the first verb of the main clause). Consider the following data:

(19) aayú   khâw   keen   lɛ́ɛw   kɔ̂ɔ₂   tông   pay   riən
     age   reach school-age already   then   must   go   study
     '(When) (the child) reaches school-age, (he) must then go to school.'

(20) thǐng   fǒn   cà?   tòk   chǎn   kɔ̂ɔ₂   cà?   pay
     though   rain   will fall I   then   will go
     'Even though it is going to rain, I will go.'

(21) thâa   khun   tham   phìt   tamrùət   kɔ̂ɔ₂   cà?   càp   khun
     if   you   do   wrong police   then   will arrest you
     'If you do something wrong, the police will then arrest you.'

Kɔ̂ɔ₂ in (19–21) is also treated as an adverb based on the criterion that it cannot be negated, as shown in (22–24).

(22) *thǐng   fǒn   cà?   tòk   chǎn   mây   kɔ̂ɔ₂   cà?   pay
      though   rain   will fall I   not   then   will go

(23) *thâa   khun   tham   phìt   tamrùət   mây   kɔ̂ɔ₂   cà?   càp   khun
      if   you   do   wrong police   not   then   will arrest you

(24) *aayú khâw keen lέεw mây kɔɔ₂ tông pay riən
     age reach study-age get not then must go study

Again, the linear position of *kɔɔ₂* as an adverb is fixed. That is, *kɔɔ₂* only appears after the subject, if any. This is evidence against treating *kɔɔ₁* and *kɔɔ₂* as conjunctions, because a conjunction does not require a subject. Only a predicate takes a subject. The subject, which is marked as [+Nom] (nominative), cannot appear as the dependent sister of a conjunction. If *kɔɔ₂* were analyzed as a conjunction, the result would be an ill-formed structure as in (19a) below:

(19) a. *

khâw			kɔɔ₂	tông		
reach			then	must	pay	
2ndex			5ndex	6ndex	go	riən
aayú	+V	keen	+cnjn	+V	7ndex	study
age	+prdc	school-age		−fint	+V	8ndex
1ndex	1[+PAT]	3ndex		7([+V])	−fint	+V
Nom	1[+actr]	Acc	lέεw	5([+cnjn])	8[+V]	−fint
PAT	3[+COR]	COR	already	7[−fint]	8[−fint]	
actr		4([+Adv])	4ndex	6<7[−fint]	7>8[−fint]	
			+Adv			
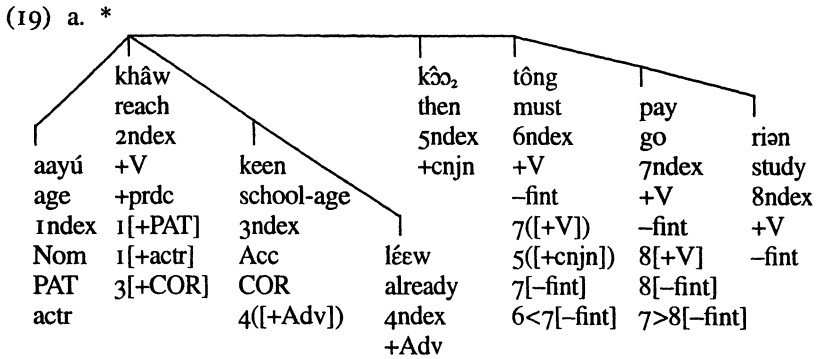
Even though *kɔɔ₁* and *kɔɔ₂* are both adverbs, they differ from each other in that each has a different meaning. *Kɔɔ₂* means 'then', whereas *kɔɔ₁* means 'also'. The structure of sentences with *kɔɔ₂* can be compared to conditional clauses in English. That is, in a conditional clause, the conditional forms such as *if, though, even though, however,* and so forth, function to introduce a subordinate clause and indicate that there must be a main clause after the subordinate clause. For example, in the sentence *if the student misses his class, the teacher will punish him, if* as a conditional form functions to introduce the subordinate clause *the student misses his class,* and implies that the clause *the teacher will punish him* is a main clause that comes after the subordinate clause.

However, *kɔɔ₂* in Thai exhibits some properties that are different from conditionals in English. *Kɔɔ₂* implies a preceding rather than a following subordinate clause. The preceding subordinate clause may or may not have a conditional form as in (19), repeated here as (25).

(25) aayú khâw keen lέεw kɔɔ₂ tông pay riən
     age reach school-age already then must go study
     '(When) (the child) reaches school-age, (he) must then go to school.'

The clause in which *kɔɔ₂* appears as a dependent sister is always a main clause whose meaning implies the continuing, resultative consequence of the previous clause. The internal structure of the sentence having *kɔɔ₂* in (25) is illustrated below as (25a):

(25) a.

```
            khâw                    kɔ́ɔ₂  tô ng
            reach                   then  must      pay
            2ndex                   5ndex 6ndex     go      riən
  aayú      +V          keen        +Adv  +V        7ndex   study
  age       +prdc       school-age        –fint     +V      8ndex
  Index     1[+PAT]     3ndex             7([+V])   –fint   +V
  Nom       1[+actr]    Acc         lέεw  7[–fint]   8[+V]  –fint
  PAT       3[+COR]     COR         already 5([+Adv]) 8[–fint]
  actr                  4([+Adv])   4ndex   6<7[–fint]
                                    +Adv    2([+fint])
                                            6>2([+fint])
                                            2([+fint])<5([+Adv])
```
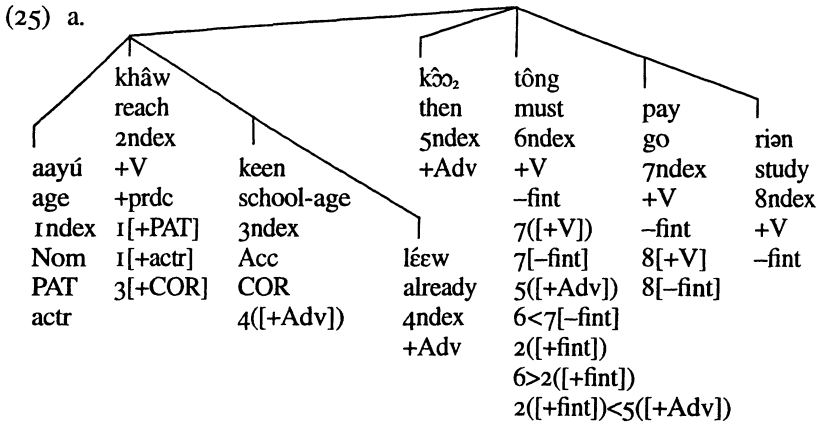
'(When) (the child) reaches school-age, (he) must then go to school.'

The clause *aayú khâw keen lέεw* '(the child) reaches school-age' is treated here as a subordinate clause, and the clause *kɔ́ɔ₂ tô ng pay riən* '(he) must then go to school' is treated as a main clause. There is no clear syntactic evidence to prove that the clause in which *kɔ́ɔ₂* 'then' appears is a main clause. However, the determination for a main clause can be explained by a pragmatic criterion. That is, *kɔ́ɔ₂* introduces additional information, the purpose of which is to complete the proposition previously introduced. The previous perspective is represented as a subordinate clause, whereas the new perspective is conveyed in a main clause (Starosta, pers. comm.).

Another bit of supporting evidence for treating the clause having *kɔ́ɔ₂* as a main clause is that reversing the order of the subordinate clause and the main clause is not permitted. If applied to (25), the result is the ungrammatical sentence (26).

(26) *kɔ́ɔ₂ tô ng pay riən   aayú khâw   keen   lέεw
     then must   go   study age   reach school-age already

Thus, the internal structure of a clause having *kɔ́ɔ₂* is considered to be that of a main clause, not a subordinate clause. The regent *tô ng* imposes the valency features on its dependent sisters. That is, it allows an adjunct *(aayú) khâw (keen lέεw)* and an adverb *kɔ́ɔ₂* to appear as its dependent sisters on the left, and the complement *pay* to appear as its dependent sister on the right (as shown in the tree stemma in [25a]).

2.3 *Kɔ́ɔ₃* as an Interjection

Kɔ́ɔ₃ meaning 'well' is treated as an interjection. In Lexicase dependency grammar, the syntactic category INTERJECTION has not been claimed to be a particular part of speech. So far, only eight parts of speech have been established, namely, noun, verb, adjective, adverb, preposition, determiner, conjunction, and sentence particle (Starosta 1988:27). By looking at the distribution of *kɔ́ɔ₃* in Thai, the category INTERJECTION may have to be introduced into the theory. Consider the following sentences:

(27) pay thîəw   sanùk măy
     go sightseeing fun   or not
     'Do you enjoy going sightseeing?'

(28) kɔ̂ɔ₃ sanùk sì
     well fun   indeed
     'Well, it is fun indeed.'

(29) càʔ tham măy
     will do or not
     'Will you do it?'

(30) kɔ̂ɔ₃ mây rúu   càʔ tham r̆ɨ plàw
     well not know will do or not
     'Well, (I) don't know whether I will do it or not.'

The characteristic of *kɔ̂ɔ₃* is that it behaves like *kɔ̂ɔ₁* and *kɔ̂ɔ₂* in that it cannot be negated as in (31–32). This means that *kɔ̂ɔ₃* cannot be analyzed as a verb.

(31) *mây kɔ̂ɔ₃ sanùk sì
      not well fun   indeed

(32) *mây kɔ̂ɔ₃ rúu   càʔ tham r̆ɨ plàw
      not well know will do or not

However, as demonstrated in (28) and (30), *kɔ̂ɔ₃* can answer a question, which disqualifies it from belonging to the adverb class. While it can answer a question, it does not take any dependent sisters, as in (34) and (36) below. It always appears by itself without any dependent sisters, as in (28) and (30).

(33) pay thîəw   sanùk măy
     go sightseeing fun   or not
     'Do you enjoy going sightseeing?'

(34) *chan kɔ̂ɔ₃ sanùk sì
      I   well fun   indeed

(35) càʔ tham măy
     will do or not
     'Will you do it?'

(36) *chan kɔ̂ɔ₃ mây rúu   càʔ tham r̆ɨ plàw
      I   well not know will do or not

The preceding analysis of the syntactic characteristics of *kɔ̂ɔ₃* leads to the conclusion that it does not belong to any of the syntactic classes established thus far in Lexicase theory. Thus, it will be claimed in this essay that *kɔ̂ɔ₃* belongs to a new syntactic class labeled INTERJECTION. It is claimed that an interjection typically appears by itself with no dependent sisters. Furthermore, it is inflected for the category [±root]. The feature [+root] means that an entity is the head of an independent occurrence. The feature [−root] means that it is not the head of an independent occurrence; it can only appear as a dependent of a predicate (Starosta 1998:283ff.).

147

In Thai, an interjection only appears in the initial position of a sentence. Some other words in Thai that can be treated as interjections are *âaw, ôo, um, ahâa,* all of which have the same meaning as the interjection "well" in English. These interjections can substitute for *kɔ̂ɔ₃*, as seen in the following:

    (37) âaw/ ôo/ um/ ahâa sanùk sì
          well   fun   indeed
          'Well, it is fun indeed.'

The question arises as to how to determine the internal structure of sentences containing *kɔɔ₃*. Considering sentence (28), repeated here as (38), we can propose four possible ways of assigning its internal structure.

    (38) kɔ̂ɔ₃ sanùk sì
          well fun   indeed
          'Well, it is fun indeed.'

The first hypothesis is that *kɔɔ₃* is treated as an independent clause, and it cooccurs with another independent clause. The structure would look like the following:

    (38) a.

kɔ̂ɔ₃	sanùk	
well	fun	sì
1ndex	2ndex	indeed
+Ntjn	+V	3ndex
–root	2([+Adv])	+Adv

However, this structure is not possible, because the structure implies that *kɔɔ₃* appears as an independent clause without any dependent sister, which is contrary to the fact that *kɔɔ₃* cannot appear independently in a simple clause, as seen in the following dialogue:

    (39) A: ngaan mïikhïin níi sanùk mǎy
            party   last night this fun   or not
            'Was the party fun last night?'

    (40) B: kɔ̂ɔ₃ sanùk sì
            well fun   indeed
            'Well, it was fun indeed.'

    (41) B: *kɔ̂ɔ₃
            well

The feature [–root] marked on the interjection *kɔ̂ɔ₃* indicates that *kɔ̂ɔ₃* cannot be a sentence. It is an incomplete sentence, therefore it cannot be the highest node in a sentence. If it were to occur by itself as in (41) above, it would be a fragment and not a grammatical sentence.

In the second hypothesis, *kɔ̂ɔ₃* appears in a coordinate clause, in which it forms a coordinate structure with another clause. The structure would look like that shown in (38b).

(38) b.

```
kɔɔ₃      sanùk
well      fun        sì
ındex     2ndex      indeed
+Ntjn     +V         3ndex
          3([+Adv])  +Adv
```
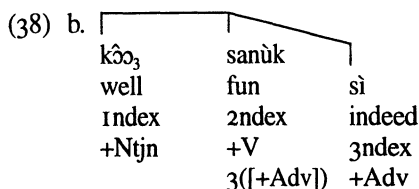
This structure is not possible either, because by Lexicase definition, a coordinate structure requires its phrasal heads to be of the same syntactic category (Starosta 1988:246). Here, *kɔɔ₃* is an interjection and *sanùk* is a verb. That is, they belong to different syntactic categories, thus contradicting the property of coordination. Based on the ungrammaticality of sentences (31) and (32), *kɔɔ₃* cannot be treated as a verb, and it cannot be coordinated with the verb *sanùk* either. Forcing these two different categories to be coordinated results in mis-assigning a proper internal structure for sentence (38).

Another reason to support the view that *kɔɔ₃* and *sanùk* cannot be coordinated is that in a coordinated construction, the conjoined items can switch order. However, the order of *kɔɔ₃* and *sanùk sì* in (38) cannot be interchanged to *sanùk sì kɔɔ₃. The whole conjuncts would change, not just their head words, as seen in the following stemma. This provides further confirmation that *kɔɔ₃* and *sanùk* cannot be coordinated.

(38) b'. *

```
sanùk
fun        sì         kɔɔ₃
ındex      indeed     well
+V         2ndex      3ndex
3([+Adv])  +Adv       +Ntjn
                      –root
```
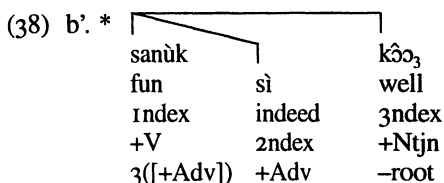
The third hypothesis treats *kɔɔ₃* as the superordinate of a clause. *Kɔɔ₃* is the regent predicate of the whole clause. The structure would look like the following:

(38) c.

```
kɔɔ₃
well      sanùk
ındex     fun        sì
+Ntjn     2ndex      indeed
–root     +V         3ndex
2[+V]     –fint      [+Adv]
          3[+Adv]
```
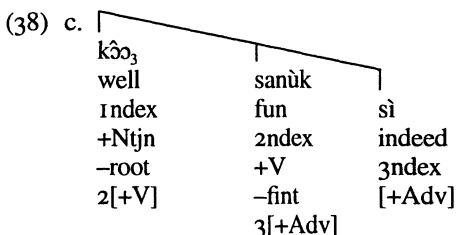
This structure is also not possible, because if *kɔɔ₃* is a superordinate, which requires a dependent sister *sanùk*, it implies that there cannot be any intonation break between *kɔɔ₃* and *sanùk*. This implication is contrary to the fact that an intonation break is permitted between *kɔɔ₃* and *sanùk*. Thus, the structure makes an erroneous prediction as far as the data are concerned.

The last hypothesis considers *kɔɔ₃* to be one of the dependent sisters of a main clause. The syntactic relationship of *kɔɔ₃* and its regent predicate *sanùk* is that *kɔɔ₃* is a subordinate (which means dependent in Lexicase dependency grammar [Stan Starosta, pers. comm.]) and *sanùk* is the main clause. The structure would look like (38d).

(38) d.

	sanùk	
kɔ̂ɔ₃	fun	sì
well	2ndex	indeed
1ndex	+V	3ndex
+Ntjn	−fint	+Adv
−root	1([+Ntjn])	
	3([+Adv])	
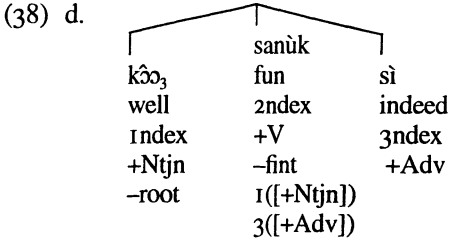
This is the only possible structure for *kɔ̂ɔ₃*, as it permits *kɔ̂ɔ₃* to appear as the subordinate dependent sister of the regent clause headed by *sanùk* 'be fun'. What appears to be the problem here is in what meaningful way is *kɔ̂ɔ₃* a "dependent" of *sanùk* (or any other V for that matter) that is different from the other dependent *sì*. The syntactic argument that supports this hypothesis is that the structure implies that an intonation break between *kɔ̂ɔ₃* and *sanùk* is possible. This analysis shows that *kɔ̂ɔ₃* is part of a higher clause. The whole sentence is well-formed and there is no violation of any features. The structure satisfies the well-formedness condition on the dependency tree, which shows that the highest node is a predicate verb that allows an interjection ([+Ntjn]) *kɔ̂ɔ₃* and an adverb *sì* to appear as its dependent sisters.

As stated earlier, there are other Thai words that can appear as interjections. However, their syntactic characteristics are quite different from those of *kɔ̂ɔ₃* in that they can appear by themselves without any dependent sisters. On the other hand, *kɔ̂ɔ₃* appears as the dependent sister of its regent predicate. This indicates that the class INTERJECTION can be divided into two different subclasses: an interjection with the feature [+root] and an interjection with the feature [−root]. The set of words that belongs to the first subclass includes *âaw, ôo, um,* and *ahâa.* The single lexical item *kɔ̂ɔ₃* belongs to the second subclass.

From the discussion above, it will be claimed in this study that the syntactic class INTERJECTION needs to be set up in the current Lexicase dependency grammar in order to account for the occurrence of *kɔ̂ɔ₃* in Thai.

2.4 *Kɔ̂ɔ₄* as an Interjection

The last lexical entry of *kɔ̂ɔ* is *kɔ̂ɔ₄* and it is treated as an interjection like *kɔ̂ɔ₃*. However, it is different from *kɔ̂ɔ₃* in that the meaning of *kɔ̂ɔ₃* is 'well', whereas, the meaning of *kɔ̂ɔ₄* is 'surprisingly', or 'it is unexpected'. *Kɔ̂ɔ₄* is treated as belonging to the syntactic category INTERJECTION because it has the same syntactic characteristics as an interjection. First, it does not take dependent sisters. Second, an intonation break between *kɔ̂ɔ₄* and its regent clause is also allowed. *Kɔ̂ɔ₄* is treated as being the subordinate dependent sister of a regent clause.

Consider sentence (42) and its corresponding internal structure.

(42) kɔ̂ɔ₄   chǎn mây rúu   wâa hètkaan níi cà? kə̀ət   khîn
unexpected I   not know that incident this will happen go up
'It is unexpected (for me) that this incident would happen.'

(42) a.

			rúu know		cà?̀	
kɔ̂ɔ₄	chǎn	mây	4ndex	wâa	will	
unexp.	I	not	+V	that	8ndex	
1ndex	2ndex	3ndex	+prdc	5ndex	+V	kə̀ət
+Ntjn	Nom	+Adv	1([+Ntjn])	+P	9[+V]	happen khîn
−root	PAT		2[+Nom]	7[+V]		9ndex go up
			2[+PAT]			+V 10ndex
			3([+Adv])			+prdc +Adv
			5([+P])	hètkaan		−fint
			4>1([+Ntjn])	incident	níi	10([+Adv])
			4>3([+Adv])	6ndex	this	
				Nom	7ndex	
				PAT	+Det	
				7([+Det])		

'It is unexpected (for me) that this incident would happen.'
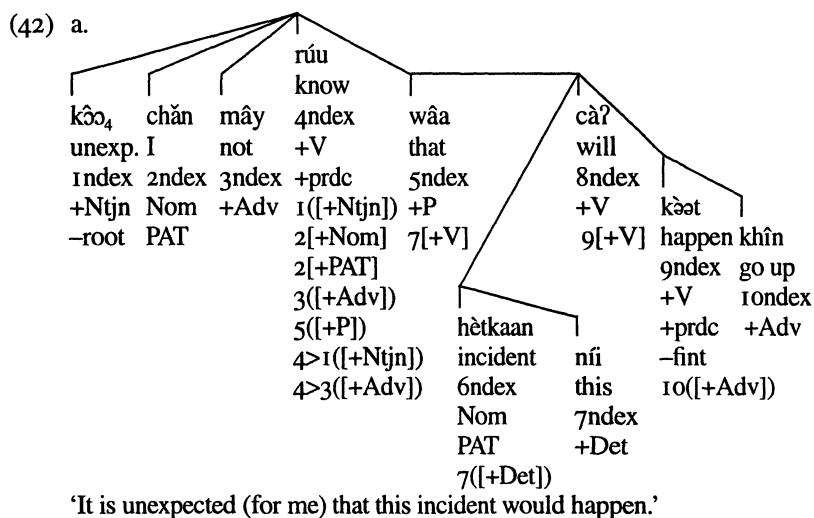
The structure shows that *kɔ̂ɔ₄* appears as the subordinate dependent sister of the regent verb *rúu* 'know'. The verb *rúu* imposes a selectional restriction on *kɔ̂ɔ₄* to appear as one of its dependent sisters on the left. *Rúu* also requires other dependent sisters, *chan* 'I' as its nominative subject, *mây* 'not' as its adverbial sister, and *wâa (hètkaan nìi cà? kə̀ət khîn)* 'that (this incident would happen)' as its dependent sister complement.

3. Conclusion

This study has shown the various kinds of syntactic distribution manifested by *kɔ̂ɔ* in Thai using the Lexicase dependency grammar framework. Based on the criterion that forms that have different syntactic distributions belong to different lexical entries, there are two syntactic subclasses of *kɔ̂ɔ*, namely, adverb and interjection. Each subclass is further divided into two different subclasses based on distributional evidence and differences in meaning.

One of the major discoveries that has been made while determining the syntactic category of *kɔ̂ɔ* is that a new syntactic class of INTERJECTION needs to be established in Lexicase theory. The syntactic distribution of an interjection is quite different from all other syntactic categories in that an interjection only appears by itself without any dependent sisters.

REFERENCES

Bandhumedha, Banchop. 1975. *Láksan̆?à phaasǎa thai* (Characteristics of the Thai language). Bangkok: Kurusaphaa.

Haas, Mary R. 1964. *Thai–English's students dictionary.* Stanford: Stanford University Press.

Panupong, Vichin. 1970. *Inter-sentence relations in modern conversational Thai.* Bangkok: Siam Society.

————. 1977. *Kroongsâang phaasǎa thai: rabòp waiyakoon* (The structure of the Thai language: Syntax) Bangkok: Ramkhamhaeng University Press.

————. 1978. Some basic problems of semantics concerning certain types of homophonic-graphic words in Thai. In *Proceedings of the 12th International Congress of Linguists*, ed. by Wofgang U. Dresser and Wolfgang Merid, 217–221. Innsbrucker Beitraege zur Sprachwissenschaft. Innsbruck: Institut für Sprachwissenschaft der Universität Innsbruck.

Savetamalya, Saranya. 1987. A reanalysis of auxiliaries in Thai. University of Hawai'i *Working Papers in Linguistics*, 19.1:1–44.

Starosta, Stanley. 1988. *The case for lexicase: An outline of lexicase grammatical theory*. London: Pinter Publishers.

————. 1993. Word order and focus in constrained dependency grammar. In *Proceedings of the conference: Functional description of language*, ed. by Eva Hajičová, 237–252. Prague: Faculty of Mathematics and Physics, Charles University.

————. 1998. Ergativity, transitivity, and clitic coreference in four Western Austronesian languages. In *Case, typlogy, and grammar: In honour of Barry J. Blake*, ed. by Anna Siewierska and Jae Jung Song, 277–307. Amsterdam: John Benjamins.

Thonglor, Kamchay. 1972. *Làk phaasǎa thai* (The structure of the Thai language). 4th ed. Bangkok: Bamrungsarn.

Warotamasikkhadit, Udom. 1963. Thai syntax: An outline. Ph.D. dissertation, University of Texas.

11:   HUNGER ACTS ON ME: THE GRAMMAR AND SEMANTICS OF BODILY AND MENTAL PROCESS EXPRESSIONS IN KALAM

Andrew Pawley
Simon Peter Gi
Ian Saem Majnep
John Kias
Department of Linguistics, Research School of Pacific and Asian Studies,
Australian National University; Simbai, Papua New Guinea

1. Introduction

This essay examines the way one broad domain of human experience, namely, bodily and mental processes (BMP), is represented in the Kalam language of Papua New Guinea.[1] Kalam is spoken by about 15,000 people living around the junction of the Bismarck and Schrader Ranges on the northern fringes of the Central Highlands, in the southwest corner of Madang Province.[2]

The term "bodily and mental processes" covers a diverse range of phenomena grounded in human physiology, including (i) transient observable processes like sweating, bleeding, vomiting, sleeping, snoring, sneezing, and giving birth, (ii) stable visible conditions like having boils, dandruff, or warts, or being pregnant, (iii) sensations like feeling sick or hungry, or tasting bitter, sweet, or salty, (iv) emotions such as

---

2. Kalam belongs to the Madang branch of the large Trans New Guinea Phylum, to which most so-called "Papuan" languages belong. There are two main dialects that in morphology are about as different as Spanish and Italian and so might well be regarded as different languages. *Etp Mnm* is spoken in the Upper and Middle Simbai Valley and Kaiment Valleys and much of the Upper Kaironk. *Ti Mnm* is spoken in the Asai Valley and parts of the Upper Kaironk. Examples used here are in the Etp Mnm dialect as spoken in the Upper Kaironk Valley.

    Linguistic analyses of Kalam include a general grammar by Pawley (1966), a detailed study of serial verb constructions by Lane (1991), and treatments of particular topics by Givón (1990), Pawley (1987, 1993), and Pawley and Lane (1998). There is a dictionary soon to be published (Pawley et al., in press).

feeling angry, sad, or jealous, and (v) cognitive processes like being awake, thinking, remembering, and dreaming.

Most kinds of BMP are usually expressed in Kalam by a single verbal clause, although some are represented by a multi-clausal construction. As is the case in many Papuan languages, two main classes of single-clause BMP constructions can be distinguished, which we will call *Experiencer Subject* and *Experiencer Object* constructions. The two types are illustrated by the sentences in (1) and (2), respectively.[3]

Table 1. Consonant Allophones

	WORD-INITIAL	MEDIAL	FINAL
ORAL OBSTRUENTS			
p	[ɸ]	[β]	[p] or [b]
t	[t]	[ř]	[ř]
c	[č]	[č]	[č]
s	[s]	[s]	[s]
k	[k]	[ɣ]	[k]
PRENASALIZED OBSTRUENTS			
b	[mb]	[mb]	[mp]
d	[nd]	[nd]	[nt]
j	[ňǰ]	[ňǰ]	[ňč]
g	[ŋg]	[ŋg]	[ŋk]
RESONANTS			
m	[m]	[m]	[m]
n	[n]	[n]	[n]
ñ	[ñ]	[ñ]	[ñ]
ŋ	[ŋ]	[ŋ]	[ŋ]
l	[ł]	[ł]	[ł]
w	[w]	[u] /C_C	[w]
y	[y]	[i] /C_C	[y]

---

3. The consonants and their allophones (a rough indication of the pronunciation of each, but excluding the consonant-release vocoid) are given in table 1. There are three vowel phonemes, *a, e, o,* realized as as [aˑ], [eˑ] and [oˑ], and two semivowels, spelled *w* and *y* when nonsyllabic and *u* and *i* when syllabic.

Many words contain no phonemic vowels. Consonant phonemes standing alone or before another consonant in a word are released with a predictable epenthetic vowel, which may be called a "consonant release" vowel, as in /wsn/ [wusín] 'sleep', /ytk/ [yirík] 'forest', /mlp/ [miłíp] 'dry', /gpnp/ [ŋíβiníp] 'I might have done'.

In words consisting of a single consonant other than a palatal, the release vowel is a very short stressed [ə́], e.g., [mbə́] 'man'. In the context C_CVC it may be a very short, near copy of V or a short central or high central vowel, e.g., /lkañ/ 'blood' is [laɣá.ñ] or [lɨɣá.ñ]. Elsewhere before and after /y/, the epenthetic vowel is usually [i]. Elsewhere before and after /w/, it is usually [u].

(1) a. Nad   suk   a-sp-an.
         you (SUBJ)   laughter   say-PROG-2S[4]
         'You are laughing.'

     b. Nuk   kñik   g-s-a-p.
         he   saliva   do-PROG-3S
         'He is spitting.'

     c. Yad   ss   ki-sp-in.
         I   urine   excrete-PROG-1S
         'I am urinating.'

     d. Nuk   añŋ   ay-s-a-p.
         he   breath/pulse put-PROG-3S
         'He is breathing'/ 'He is catching his breath.'

     e. Yad   ñb   nŋb-in.
         I   eat   perceive-PF-1S
         'I tasted (it) (deliberately).'

(2) a. Np   suk   ow-p.
         you (OBJ)   laughter   come-PF.3S
         'You felt like laughing/You couldn't help laughing.'

     b. Nup   kñik   ow-p.
         him   saliva   come-PF.3S
         (i) 'He is salivating.' (also [ii] 'He feels a craving.' [iii] 'He feels revulsion.')

     c. Yp   ss   yow-p.
         me   urine   fall-PF.3S
         'I feel like urinating/I need to urinate.'

     d. Np   añŋ   yow-p.
         you (OBJ)   breath/pulse fall-PF.3S
         'You are out of breath.'

     e. Yp   ydk   g-p.
         me   tasty   do-PF.3S
         'It tasted good to me.'

In grammatical terms, the sentences in (1) have the following in common: (a) the Experiencer is represented by a pronoun from the Subject set of pronouns and (b) the person-number suffix on the verb, which marks Subject, is coreferential with the Experiencer. The sentences in (2) differ in that (a) the Experiencer is represented by a

---

4. Key to abbreviations used in glosses: D, dual; DS, different Subject (from following verb); DUR, durative; FUT, future ; HORT, hortative; I, Instrument; IMM, Immediate past; L, Locative; lit., literally; NDR, noun derivative suffix; OBJ, Object (case); OPT, optative; P, plural; PF, perfect (denotes today's past, present perfect, or present-iterative); PAST, remote past (yesterday or earlier); PAST.HAB, past habitual; PRIOR, prior or preceding (the event denoted by following verb); REC, recent past; PROG, present progressive; S, singular; SIM, simultaneous (with the event denoted by following verb); SS, same Subject (as following verb); SUBJ, Subject.

pronoun from the Object set and (b) the person-marker suffix on the verb always marks 3rd person singular. It is clear that in (2) the suffix does not refer to the Experiencer, but it is not always clear whether or not it refers to the other nominal in the clause. Foley (1986:121–127, 190–194) discusses comparable examples from a selection of Papuan languages under the rubric of events that the animate participant controls or does not control.

BMP expressions are a rich field for the study of human perceptions of the body and mind, a domain of considerable interest to philosophers, psychologists, and anthropologists. For linguists, such expressions are of special interest in connection with the debate about how iconic the design of languages is with human perceptions of the world, because BMP expressions refer to a domain of experience that is presumably the same for all people.

In the transformation of experience into conventional linguistic representations, there are certain interfaces where there is potential either for iconicity or for arbitrariness. One such interface resides in the speaker's initial interpretation of sensations and perceptions as meaningful entities. "Making sense" of experience is a complex process that involves paying attention to particular bits of sensory input, while ignoring others, and placing a construction on those bits that are attended to. For example, someone tasting a bottle of wine at dinner may experience many different sensations at once—the feel of the glass on the lips, the temperature of the room, the flickering of lights, the hum of conversation, etc.—and pay little or no attention to any of the sensations except for the effect of the wine on the palate and the nose. Categorizing that sensation will be a matter of judgment, and translating that judgment into words will require still other judgments.[5]

A second interface resides in the matching of syntax and semantics. If choice of grammatical category and construction were fully determined by semantic categories and relations, there would, by definition, be a perfect correspondence between syntax and meaning. However, in all languages there are more semantic distinctions than there are grammatical ones, and this mismatch creates a design problem—how to squash a wide range of semantic structures into a smaller range of formal structures—which guarantees some degree of arbitrariness in the coding of meaning. Even though all formal constructions may be in large part semantically motivated, in that each represents a canonical semantic type, most construction types also accommodate more than one kind of conceptual construction. Thus, transitive constructions in English represent both canonical transitive relations such as "The man ate his lunch" and "Mary kicked the cat," where an agent does something to a patient, and quite different sorts of relations, such as "The man heard the noise" and "Mary feared the worst."

In the present essay, our primary concerns will be with certain basic questions in the domain on grammar and semantics. What kinds of syntactic constructions are used to represent particular bodily and mental processes? To what extent is the choice of construction for particular BMP expressions determined by semantic consider-

---

5. In his book *The Linguistic Construction of Reality*, George Grace (1987) discusses these issues at length.

ations? Investigation of these matters will lead us to consider some questions about other issues in Kalam grammar and semantics—for example, what properties define "Subject" and "Direct Object?" How do notions like 'cause', 'volition' and 'control' figure in the grammar? What factors trigger switch reference morphology on verbs?

Comparison of the examples in (1) and (2) suggests that Kalam BMP constructions show a neat fit between syntax and semantics. The sentences in (1) denote processes that are controlled by the Experiencer (at least to some extent) and the sentences in (2) represent processes that are *involuntary* or *uncontrolled*. However, examination of a wider range of processes shows that things are not so straightforward.

Before we go deeper into BMP expressions, a few remarks are needed on relevant points of Kalam grammar. (For phonology and orthography see n. 3 and table 1.)

2. Notes on Kalam Grammar

2.1 Verbal Clauses

A verbal clause consists minimally of an inflected verb. It may also contain a number of core and adjunct noun phrases and adverbial elements.

2.2 Verbs

2.2.1 Verb roots and verb morphology. Verb roots or bases are a closed class with about 130 recorded members. Their morphological and syntactical properties are highly distinctive. They are the only part of speech to carry inflectional suffixes marking tense, aspect, or mood, Subject person-and-number, and switch reference.

Inflected verbs are either independent or dependent. Independent verbs carry suffixes marking absolute tense, aspect, or mood, and suffixes marking Subject-reference independently of any other verb. Dependent verbs carry suffixes marking relative tense (prior, simultaneous with, or subsequent to) and relative Subject reference (same or different), the comparison being with the next verb in the construction. These markers of relative tense and Subject-reference are discussed in 2.6 under the heading "switch reference and relative tense marking." Some dependent verbs are also marked for absolute tense, aspect, or mood, and person-number of Subject.

2.2.2 Generic verb roots. In forming its verbal lexicon and in certain other respects, Kalam relies heavily on a small number of "generic" verb roots. These are verbs that have a very broad meaning (e.g., *g-* 'happen, act, do, make, work', *ag-* 'make a sound, say', *ay-* 'put, form, stabilize, become', *d-* 'touch, have, control, finish', *ŋ-* 'perceive, know, see, hear, smell, feel, be conscious, be aware, observe') and that, as well as occurring alone, occur as a component in many conventional expressions, in sequence with verbal adjuncts (2.2.3) and/or with other verb roots (2.2.4). Fifteen generic roots alone account for nearly 90 percent of verb-root tokens in text. Thirty-five verb roots account for some 98 percent of all such tokens.

2.2.3 Complex verbs: Verbal adjuncts + verb root. A verbal adjunct is a word or phrase that occurs only in partnership with a verb root, helping to define its meaning more specifically and that, together with the verb root, forms a complex verb. They differ from clausal adverbs in that verbal adjuncts always form a constituent with the verb, although not necessarily a continuous constituent. True verbal adjuncts cannot stand in a case relation with the verb, that is, cannot be Subject, Direct Object, Locative, and so forth, cannot be possessed or modified by an adjective, and cannot modify a noun. However, there is a cline between true adjuncts and quasi-adjuncts that have adverb-, adjective- and noun-like properties. In the following examples, *si* 'illegally' and *dad* 'carrying, being in possession of, controlling' are true adjuncts and *wsn* 'sleep, sleeping' is a quasi-adjunct. Verbal adjuncts differ from adverbs in that they are always constituents of a complex verb and cannot modify larger constituents.

(3) a. Kik   wog-yad   si   amn-ya-k.
        they   garden-my illegally go-3p-PAST
        'They trespassed on my garden.'

    b. Cn np   mdak ayn-bogs dad   n-ng   g-p-un.
       we you(OBJ) later   iron-box carrying join-SS.FUT do-PF-IP
       'We will join you later bringing the metal trunk.'

    c. Cn wsn   kn-nu-k.
       we sleep(ing) lie-IP-PAST
       'We slept.'

A large part of the verbal lexicon of Kalam consists of adjunct + verb-root sequences. For example, the adjunct *si* 'unlawful, illegal' occurs paired with a variety of verbs, as in *si am-* (illegally go) 'trespass', *si bsg-* (illegally sit) 'steal s.o.'s seat', *si d-* (illegally take) 'steal', *si g-* (illegally do) 'transgress', *si ñag-* (illegally shoot) 'poach (game)', *si tb-* (illegally cut) 'poach timber', *mgn si daŋ-* (vagina illegally copulate) '(of a man) commit adultery'.

2.2.4 Complex verbs: Serial verb constructions. Kalam makes extensive use of serial verb constructions, in which one or more bare verb stems precede an inflected final verb, usually under a single intonation contour. With certain exceptions, the nonfinal verbs in such a series share the same Subject and the same tense, aspect, or mood as the final verb, as in (4).

(4) a. Am kmn   pk d   ap   ad   ñb-igp-ay.
        go game.mammal kill get   come cook eat-PAST.HAB-3P
       'They used to hunt game mammals for food.'(lit. 'They used to go and
       kill game mammals and bring back and cook and eat.')

    b. Yp wik d   ap   tan   d   ap   yap   g-a-k.
       me rub touch come ascend touch come descend do-3S-PAST
       'He massaged me.' / 'He gave me a rubdown.'

Serial verb constructions are an important means of augmenting the verbal lexicon. The number of conventional expressions consisting of two or more verb stems runs

into the thousands. A few examples: *d am-* (get go) 'take', *d ap-* (get come) 'bring', *am d ap-* (go get come) 'fetch', *ptk am-* (fear go) 'flee', *kn am-* (sleep go) 'drift off to sleep', *ag ñ-* (say transfer) 'tell', *ag ask-* (say avoid) 'refuse', *ag ay-* (say stabilize) 'make an appointment, ask to stay', *ag ŋ-* (say perceive) 'ask', *ag slok-* (say slip.off) 'reneg', *ag g-* (say do) 'insist, command', *pk cg-* (strike adhere) 'attach, stick on', *tk ay pk ay-* (give.birth complete strike complete) 'commit infanticide', *kat d ap puŋi tb tk-* (car hold come press cut interrupt) 'put the brakes on a car', *d ap tan d ap yap-* (hold come ascend hold come descend) 'move back and forth', *wik d ap tan d ap yap g-* (rub hold come ascend hold come descend do) 'massage something, give someone a rubdown'.

2.3 Other Parts of Speech

Several parts of speech other than verbs and verbal adjuncts can be distinguished, chiefly on syntactic grounds. The following discussion treats only those word classes relevant to this essay.

2.3.1 Personal pronouns. There are two main sets of personal pronouns: the Subject (or nominative) and Object (or accusative) sets. They are most often used with humans, less often with higher animals, and rarely with low-ranking animals such as insects. They very rarely occur with inanimate referents, which are usually represented by a full NP or, anaphorically, by zero or by a demonstrative such as *mey* 'that (previously mentioned)'. The most basic forms in each pronoun set are given in table 2.

Table 2. Subject and Object Pronouns (Animate)

	1S	2S	3S	1D	2D	3D	1P	2P	3P
SUBJECT	yad	nad	nuk	ct	nt	kikmay	cn	nb	kik
OBJECT	yp	np	nup	ctp	ntp	kuypmay	cnp	nbp	kuyp

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.

2.3.2 Common nouns. Common nouns can serve as Subject or Object of a verb and can take articles and demonstratives, can be modified by a possessor or adjective, and can serve as modifiers before another noun, as in *smi kotp yob ak* (dance.festival house large that) 'that large dance festival house'.

2.3.3 Proper nouns. Proper nouns can serve as Subject or Object of a verb but cannot be possessed and cannot take articles or demonstratives.

2.3.4 Question-words. The question-words *an* 'who?', *an-nup* 'whom?', *etp* 'what?', *etp-nen* 'how, what for?, *akay?* 'where?', *won akay?* 'when?' are nominals.

2.3.5 Adjectives. Adjectives, for example, *kls* 'strong', *sketk* 'wrong', *tep* 'good', *yob* 'big', *ydk* 'tasty', can occur after a noun or verb root as a modifier. All adjectives can also occur in partnership with a verb in predicate position, e.g., *kls g-* (strong do) 'be strong, do strongly'. Unlike common nouns, adjectives cannot be possessed and cannot be modified except by reduplication and by intensifiers such as *tmey* 'bad, extremely' and *yb* 'true, truly, extremely'. However, certain groups of adjectives show some noun-like properties. In meteorological and certain bodily process expressions, an adjective can occur as the only element other than the verb (as in 5a), or the only element other than the Experiencer (5b).

(5) a. Takl g-p.
     cold do-PF.3S
     'It's cold.'

   b. Yp km g-p.
     me bitter do-PF.3S
     'It tastes bitter to me.'

2.3.6 Negation. The negative clitic *ma-* (*m-* before bases beginning with *o*) can occur only once in a clause. It always precedes the verb. Although it most often attaches directly to the inflected verb, it can also occur earlier in the clause, attached to the first element in a complex verbal group (to a verbal adjunct or the first bare verb root in a serial verb sequence) or, less commonly, to a Direct Object, Instrument, or Locative phrase. It cannot precede a Subject or Object pronoun but does occasionally occur before Subject NPs when they are in second (nontopic) position, as in (6).

(6) a. Yp ma- mablep ay-p.
     me not wart form-PF.3S
     'I don't have warts.' ('Warts do not form on me.')

   b. *Ma- yp mablep ay-p.
     not me wart form-PF.3S.
     'I don't have warts.' ('Warts do not form on me.')

2.4 Nominal Clauses

A nominal clause consists of a topic nominal and a comment or question nominal, as in (7).

(7) B yob yb-nuk Biski.
    man big name-his Biski
    'The big-man's name is Biski.'

2.5 Verbal Clauses and Case Relations

A verbal clause or simple sentence consists minimally of an inflected verb. In canonical transitive clauses, the distinction between NPs representing the core grammatical relations, Subject and Object, is usually well-marked.

2.5.1 Subject. The Subject NP can be identified using the following criteria: (a) The Subject is the NP that is coreferential with the person-number suffix on the main verb.

(b) When the Subject is a free pronoun, it selects a pronoun from the Subject class (which marks subject, independent or topic referent, or non-Object possessor).

(c) Subjects control switch reference (within certain bounds discussed in 2.6).

(d) Subjects prefer first position in the sentence.

(e) Subject denotes the participant that plays the most active role in transitive verbs.

(f) The verbal negative clitic *ma-* seldom precedes Subjects, but more readily (under some conditions) precedes Objects (when these are a full NP) and Instruments.

2.5.2 Agreement between subject and verb. Among the Subject person-and-number suffixes, 1st, 2nd, and 3rd singular and plural are distinguished; in the dual the distinction is between 1st person and 2nd/3rd. 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. With lower animal Subjects, the suffix seldom distinguishes number but may do so. With inanimate Subjects, the verbal suffix is always 3rd person singular.

2.5.3 Direct and indirect objects. Direct and Indirect Object NPs differ from Subjects as follows: (a) When the Object NP is a pronoun, the pronoun is drawn from the Object class. (However, only animate referents are represented by independent pronouns.)

(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.

(d) The verbal negative clitic *ma-* can, under certain conditions, readily precede an Object NP other than pronouns.

(e) The Direct Object NP is the Patient of two-place action verbs and the Location of existential and posture verbs. There are a few three-place verbs that take both a Direct Object and an Indirect Object in a Beneficiary or Associative role.

The order of core constituents has some degree of flexibility. In transitive clauses with an animate Subject and a single Object (Direct or Indirect), the canonical word order is SOV, as in (8).

(8) a. Bin   kaj-nup   piow-ya-k.
        woman pig-it   search-3P-PAST
        'The women searched for the pig.'

   b. An   np   ñ-a-k?
       who you give-3S-PAST
       'Who gave (it) to you?'

However, a highly topical Object (animate or inanimate) can precede an animate Subject. Direct Objects occasionally follow the verb. When the Object is animate and the Subject is inanimate, the Object usually comes first, as in (2a–e) above and (9).

161

(9) Np   et   g-p?
     you what do-PRES:PF.3S
     'What happened to you?'

Direct Objects usually follow Indirect Objects (10a), but a Direct Object that is topic (10b) or head of a relative clause (10c) precedes the Indirect Object.

(10) a. Np   moni   ñ-n-k.
        you money give-1S-PAST
        'I gave money to you.'

     b. Yad np   b   ñ-ng-ayn!
        I   you man give-FUT-1S
        'I will give you to a man!'

     c. Moni   np   ñ-n-k,   mey   akay?
        money you give-1S-PAST,   that where
        'The money I gave you, where is it?

2.5.4 Peripheral or adjunct NPs. Peripheral or Adjunct NPs include Location, Time, and Instrumental phrases. Locative phrases are sometimes marked by postpositions. Peripheral phrases other than Time usually follow Objects. For Locatives, the preferred order is SOLV or SOVL. Seldom does a simple verbal clause contain a peripheral instrumental NP. Usually the instrument role is coded as a Direct Object in a separate "instrumental" clause.

2.6 Switch Reference and Relative Tense Marking

Switch reference and relative tense are grammatical relationships that hold between successive clauses. Relative tense has to do with the sequential relationship between the events marked by the verbs. Both categories are signaled by a portmanteau suffix on the first verb in the pair. In Kalam, switch reference refers to marking of identity or change of Subject in the second of two inflected verbs. In this respect, it differs from some Papuan languages (Farr 1997, Foley 1986:190–195, Roberts 1997) in which switch reference tracks clause topics rather than Subjects.

2.6.1 Suffixes marking same subject. Four suffixes indicate Same Subject reference combined with relative tense: *-i* 'action prior to next verb' and *-tag* 'action prior to next verb' (the latter is used mainly in folk tales) (both are abbreviated as SS.PRIOR),
   -ig 'action simultaneous with next verb' (SS.SIM),
   -ng 'action subsequent to next verb' (SS.FUT). *-ng* is not inherently a Same Subject marker but it has this interpretation unless it is preceded by *-e-*, which marks Different Subject (see 2.6.2).
   When a verb carries a Same Subject suffix, as in (11) below, it usually has no suffix marking absolute person-and-number of its Subject. That information is normally carried by the next independent verb, usually the final verb in the sentence.

(11) a. Am-i,   ap-i,   wog g-ng   g-p-in   ag-p.
        go-SS.PRIOR come-SS.PRIOR work do-SS.PRIOR do-PF-IS say-PF-3S.SS.FUT
        'He said he will do the work when he gets back.' (lit. '"Having gone,
        having come, intending to do work I do," he said.')

   b. Nad cnp gos   nŋ-ig   md-enm-n.
      you us thought perceive-SS.SIM stay-OPT-2S
      'You should keep thinking of us.' (lit. 'You should stay thinking of us.')

2.6.2 Suffixes marking different subject. Two suffixes indicate Different Subject reference combined with relative tense: *-e-* (or its conditioned variant *-o-*) denotes action prior to next verb (DS.PRIOR), and *-nŋ* marks action simultaneous to next verb (DS.SIM). The sequence *-e-ng-* (DS-FUT) marks future action by Different Subject.

Any verb marked for Different Subject in the next clause is also marked for the person and number of the Subject of its own clause. Examples are given in 2.6.3.

2.6.3 Rules for same and different subject marking. The following is a rough approximation of the rules for marking switch reference. Successive animate Subjects take Same Subject marking on the verb if at least one referent of the first Subject is shared with the second. For example, IS and IP count as Same Subject. If the Subjects of both the first and second verbs are animate and if they do not share at least one participant in common, the first verb is marked for Different Subject, as in (12).

(12) a. An   ag-e-k   g-na-k?
        who say-DS.PRIOR-3S.PAST do-2S-PAST
        'Who told you to do it?'

   b. Nad ag-e-na-k   g-n-k.
      you say-DS.PRIOR-2S-PAST do-IS-PAST
      'I did it because you told me.'

   c. Kik bsg md-ya-knŋ   ow-ya-k.
      they sit stay-3P-DS.SIM   come-3P-PAST
      'While they$_i$ stayed sitting they$_j$ arrived.'

   d. Kaj yp   suw-e-k   yuwt g-p.
      pig me   bite-DS.PRIOR.3S-PAST pain do-PF.3S
      'The pig bit me and it's painful.'

If the Subject of the first verb is inanimate, and does not share at least one participant in common with the following independent verb, the first verb may be marked either for Same or Different Subject. The choice seems to depend, in at least some uses, on whether the speaker wants to present the two clauses as representing closely integrated or separate events. Compare (13a) and (13b).

(13) a. Ygen ap   cec-kotp d-e-k,
        Wind come tent   take-DS.PRIOR-3S.PAST

        kuptut   a-s-a-p.
        vibrating.noise sound-PROG-3S
        The wind caught it (the tent) and it's making a vibrating noise.'

163

b. Ygen ap   cec-kotp d-i   ap   yow-p.
   Wind come tent   take-ss.PRIOR come fall-PF.3S
   'The wind blew the tent over.' (lit. 'The wind caught the tent and it fell
   down.')

3. Bodily and Mental Processes: Experiencer Object Expressions

Let us now return to BMP constructions, beginning with the Experiencer Object type. In this type there is, it seems, a perfect fit between syntax and semantics in at least one respect. All such constructions express processes that are involuntary. To be more precise, the Experiencer does not initiate or control these processes; they just happen to him or her. The encoding of the Experiencer as Object is consistent with this conception. There are three reasons why the Experiencer NP must be regarded as Object rather than Subject of the verb. First, it is represented by a pronoun from the Object set. Second, the person-number suffix on the verb does not agree with the Experiencer NP; it is always marked for 3rd person singular. Third, switch reference marking, a syntactic process sensitive to subject identity or change across clauses, cannot be triggered by the Experiencer NP in Experiencer Object constructions.

For the grammarian, the vexing problem is to understand the grammatical roles played by the other nominal or noun-like elements in Experiencer Object constructions.

3.1 Single Clause Constructions

Such constructions typically contain, in addition to the VERB and EXPERIENCER, one or two NPs (or NP-like constituents) labeled here as CONDITION and BODY-PART.

CONDITION is a loose cover term for an NP or constituent whose head denotes a bodily or mental condition, product, process, or sensation. Condition nouns (we use "noun" to include noun-like adjuncts) include the following rough semantic subgroups: (i) Visible conditions or products, e.g., *mablep* 'wart', *sbek* 'pimple', *sŋl* 'boil', *soy* 'sore, ulcer', *slañ* 'scab', *si-ñg* 'tears', ss 'urine', *sb* 'dung', *wsb* 'sweat'.

(ii) sounds (i.e., vocal products) such as *mukbel* 'belch', *suk* 'laughter', *mnm* 'speech', *si* 'crying', *jup* 'squeak', *wal* 'scream, shriek, strident call'.

(iii) invisible persisting conditions or products such as *tap* 'sickness', *suŋ* 'good health, ritual purity', *asŋ* 'state of ritual contamination', *nŋep* 'comprehending, knowing', *saki* 'uncomprehending, forgetting, deaf, insane'.

(iv) sensations or feelings such as *slk* 'itching', *yuwt* 'pain', *yuan* 'hunger', *wokep-tek* 'feeling like vomiting', *ydk* 'tasty, tastiness, *tep* 'good, pleasure'.

The nouns in semantic groups (i) and (ii) can take a determiner and are readily modified by a possessive pronoun, quantifier or other modifier. Those in groups (iii) and (iv) seldom take any modifier other than an intensifier.

BODY-PART refers to an NP whose head names the locus of a condition or sensation in the body, such as *alkjon* 'armpit', *jun* 'head', *kogi* 'belly', *mapn* 'liver', *ss-kogi* 'bladder', *sb-wt* 'guts, innards', *wak* 'skin, body'. The Body-Part noun usually stands alone but may be accompanied by a possessive pronoun and/or determiner, quantifier or another kind of modifier.

Note that in most constructions containing a Body-Part noun, an Object pronoun can occur either as Possessor of a body-part, immediately after the possessed noun, or as a Direct Object immediately before the possessed noun, but not in both positions. The following arrangements of constituents occur:

EXPERIENCER	CONDITION	VERB	
EXPERIENCER	BODY-PART	VERB	
EXPERIENCER	BODY-PART	CONDITION	VERB
EXPERIENCER	CONDITION	BODY-PART	VERB

These can be reduced to the formula:

EXPERIENCER   (BODY-PART) (CONDITION) VERB
: CONDITION may precede BODY-PART

In most Experiencer Object constructions, the selection of verb stem(s) is narrowly constrained by the choice of Condition noun. Some Condition nouns are compatible with only one verb stem. Others allow some variation, the choice of verb stem either being free or correlating with "aspectual" meaning. Thus the Condition noun *mablep* 'wart' takes either *ap-* 'come' (which marks onset or incipience) or *ay-* 'form, put, stabilize' (which marks an established, persisting condition).

What grammatical relations hold between the verb and the nominals in this class of constructions? Do all Experiencer Object constructions have a Subject and if so, which NP has this function? The alternatives that must be considered are that: (i) all such constructions have a Subject, which, in some instances, is the Condition and in others is the Body-Part NP; (ii) some types have a Subject, others do not; (iii) none have a Subject.

The following sentence exemplifies some of the issues.

(14)   Sb-wt   yp   gullag   a-s-a-p.
         Innards/stomach my (OBJ) rumbling(s) sound-PROG-3S
         '(a) My stomach is rumbling.' (b) 'Rumblings sound in my stomach.'

It is tempting here to interpret *sb-wt yp* 'my stomach' as the Subject, parallel to the English translation (a). However, the occurrence of an Object pronoun as the possessor in this NP rules out this analysis: *sb-wt yp* is a Locative Object here. A possessive pronoun in the Objective form can never appear in a Subject NP. A second analysis would treat *gullag* 'rumbling(s)' as the Subject (in accord with the translation 'Rumblings sound in my stomach'), by analogy with, say, (15a–b).

(15) a. Tob-yp   soy   ay-p.
         leg-my (OBJ.) ulcer form-PF.3S
         'Ulcers have formed on my leg'.

     b. Tumuk   a-s-a-p   Sbay   ebneŋ.
         thunder sound-PROG-3S Simbai upriver
         'Thunder is sounding at Simbai in the up-river direction.'

165

The chief difficulty with this last analysis is that *gullag* does not behave like a canonical noun in that it cannot be possessed. It resembles a true verbal adjunct (see 2.2.3) in occurring only as a satellite of a verb *ag-* 'make a sound, say, emit'. It is possible to interpret *gullag ag-* 'make rumbling sounds' as a complex predicate in which *gullag* is regarded either (i) as a verbal adjunct of *ag-*, which does not stand in any sort of case relation to the verb, or (ii) as a defective noun acting as generic object of *ag-*. Finally, a third analysis might be proposed, in which sentence (14) is said to have no Subject, only a complex verb *gullag ag-* 'to rumble' with Locative Object *sb-wt yp* '(in) my innards'. For rough analogies we could appeal to "meteorological" and "impersonal" verbs, which in many languages occur with no Subject, or with a dummy Subject, as in *It is raining*.

With these alternatives in mind, let us now look at a range of different Experiencer Object expressions. It is convenient to subclassify such expressions, in the first place, by the interaction of (i) the verb root and (ii) the Condition nominal.

3.2 Cases Where the Condition Nominal Is Clearly the Subject

Let us consider constructions where the Condition NP can be uncontroversially interpreted as Subject because it shows the same features as canonical inanimate Subjects. These are clauses where the head of the Condition NP is clearly a noun that can be possessed or modified by adjectives of, for example, quantity or size, where it stands in the same semantic relation to the verb as other nominals that are clearly the Subject, and where it can control the marking of switch reference (see 3.3.4).

3.2.1 Stable conditions marked by *ay-* **'form, put, stabilize'.** The verb *ay-* 'form, stabilize, become' characteristically occurs with Condition nouns denoting a stable, visible external condition, such as *kñowŋ* 'birthmark', *magi-wt* 'scar', *sbek* 'pimple', *sŋl* 'boil, abscess', *soy* 'sore, ulcer', *slañ* 'scab', *tmd sb* 'ear-wax', *wdn-sgalb* 'sleep (dry secretion in eyes)', *ypl-nep* 'baldness', to indicate that the condition is already established. In such cases the Condition constituent behaves like a canonical noun and a canonical Subject of the verb *ay-* (16a–c). However, a few other Condition nominals denoting stable Conditions that are not visible occur with *ay-*, as for example *saki* 'uncomprehending, disoriented, deaf, crazy' (16d). See 3.3.2 for another use of *ay-*.

(16) a. Nup sŋl   yob alkjon ay-a-k.
             him boil big   armpit form-3S-PAST
             'A large boil has formed in his armpit.'

      b. Sŋl   alkjon-nup yob ay-a-k.
             boil   armpit-his large form-3S-PAST
             'He had a large boil in the armpit.'

      c. Kogm-np   mablep koŋay ay-p.
             knee-your warts   many   form-PF.3S
             'You have many warts on your knee.'

    d. Yp   tmud   saki   ay-p.
       me   ear   deaf   form-PF.3S
       'I am deaf.' / 'I have become deaf.'

3.2.2 Stable conditions marked by *pag-* 'broken, creased'. Certain Body-Part nominals occur with the intransitive verb *pag-* 'change shape: break, broken, creased, bent, dented, shattered, etc.' and the Condition nominal *gu* 'depression, furrow, crease, hollow' to refer to the presence of furrows, creases or dimples in the skin.

  (17) a. Ñn   yp gu   pag-p.
          hand my depression break-PF.3S
          'My hand has creases.'

       b. Mkem nup gu   pag-p.
          cheek her depression break-PF.3S
          'Her cheeks are dimpled/furrowed.'

3.2.3 Onset of a sensation or visible condition marked by *ap-* 'come'. When the Condition noun denotes a stable, visible skin condition (see 3.2.1), the use of *ap-* marks incipience of the condition, whereas *ay-* marks stability or persistence. When the Condition noun denotes a transient process or an unstable product such as *kuñk* 'saliva', *slg* 'cramp', *si-ñg* 'tears', *suk* 'laughter', *wsn* 'sleep', the verb *ap-* 'come' marks the onset of the condition. These nouns never select the verb *ay-*.

  (18) a. Yp   wsn   (wsn) ow-p.
          me   sleep sleep come-PF.3S
          'I felt (quite) sleepy.'

       b. Yp   si-ñg ow-p.
          me   tears   come-PF.3S
          'I feel like crying/I'm ready to cry.'

       c. Yp ypl   slg   ow-p,   sayn g-p;
          me muscle cramp come-PF.3S soft do-PF.3S

          ypl   tp   ok   kawboŋ g-p.
          muscle again that relaxed do-PF.3S
          'The cramp that came in (my) muscle has eased; the muscle has relaxed again.'

3.2.4 Processes marked by *yap-* 'fall', *jak-* 'rise' , *tan-* 'grow, climb'. The verb *yap-* 'fall, descend' combines with certain terms for body parts or products to indicate a process in which the part or product falls or is displaced.

  (19) a. Yp   lkañ   yow-p.
          me   blood fall-PF.3S
          'I am bleeding.'

    b. Np   wdn   pug   ju   yow-p.
       you   eye   stretched displace fall-PF.3S
       'Your eyes have become glazed.'

    c. Yp   komlŋ yow-p.
       me   groin   fall-PF.3S
       '(The gland in) my groin is swollen.'

When the body part or product has to do with body wastes (*ss-kogi* 'bladder', *ss* 'urine', *sb* 'bowel, excrement'), the conventional implicature is that the speaker has become aware of the need for elimination, as in (2c) and (20).

    (20)   Yp   sb   yow-p.
            me   excrement fall-PF.3S
            'I need to defecate/I feel like defecating.'

The verbs *jak-* 'rise' and *tan-* 'grow, climb' combine with certain terms for body parts or products (e.g., *kas* 'hair', *meg* 'teeth', *kogi* 'belly', *wsb* 'sweat') to indicate a process in which the part or product grows, swells, or rises.

    (21) a. Np   wsb   jak-p.
          you (OBJ) sweat rise-PF.3S
          'You are starting to sweat.'

        b. Yalk   nup   dsn   jak-p.
          Yalk   him   beard rise-PF.3S
          'Yalk has grown a beard.'

    (22) a. Nup   dsn   skoy   tan-b.
          him   beard small grow-PF.3S
          'He has grown a small beard.'

        b. Cnp   tob-pñg tan-b.
          us   toe-nail grow-PF.3S
          'Our toe-nails have grown.'

        c. Np   meg   omŋal tan-b.
          you (OBJ) tooth two   grow-PF.3S
          'You have grown two teeth.'

        d. Yp   lkañ -kogi   tan-b.
          me   blood-swelling grow-PF.3S
          'I have a bloodblister.'

3.2.5 Internal discomfort, marked by *ñŋ-* 'consume, eat'. The verb *ñŋ-*, loosely glossed in examples as 'eat', most commonly means 'consume (eat, drink, smoke) something'. However, it also has the senses '(of a sharp object) nick someone, make a cut in the skin or body', and 'cause internal pain or pressure'. Sometimes it is preceded (in a serial verb construction) by a verb of penetration *puɲi-* 'pierce', *su-* 'bite' or *ak-* 'scrape'.

(23) a. Yp   ña-pan-ŋaŋ   ñb-s-a-p.
     me   baby   eat-PROG-3S
     'I feel labor pains.' (lit. 'the baby is eating me.')

    b. Yp sb-wt   su   ñŋ-b.
      me guts-cluster bite eat-PF.3S
      'My stomach hurts. / I have a pain in my gut.'

    c. Tglm wagn yp   puŋi ñŋ-b.
      ribs   base   me   press eat-PF.3S
      'I have a pain in my rib cage.'

    d. Tap ñb-ab-in   yp   mdmagi ak   ñŋ-b.
      food consume-REC-IS me heart   scrape eat-PF.3S
      'After I ate, I felt a sharp pain in the heart.'

3.2.6 Involuntary processes marked by other verbs. A few other verb stems and series of verb stems occur as the main verb in an Experiencer Object construction. For example, *am-* 'go' indicates movement away, as when one's life force (*noman*) or spirit (*kawnan*) leaves the body temporarily during dreams and permanently on death; *kum-* 'die, cease to function' indicates numbness or paralysis of a body-part, *pag-* 'break, distort, disturb, fold' indicates breaking or distortion of a body-part', *tk-* 'separate, sever' indicates cracking, and any of *pag yk-* 'burst open', *pk-* 'strike' and *plg-* 'block' can indicate pulsation or throbbing.

(24) a. Tob yp   kum-b.
     leg   my die-PF.3S
     'My leg is numb/paralyzed.'

    b. Tŋl   np   pag-p.
      bone your (OBJ) break-PF.3S
      'Your bone is broken.'

    c. Spsep yp   pag   yk-ng   g-s-a-p.
      temple me   break open-SS.FUT do-PROG-3S
      'My temple is pulsating/throbbing.' (lit. 'my temple veins are pulsating as though about to burst.')

    d. Yad spsep np   d-i   nŋ-b-in,
      I   temple your touch-SS.PRIOR perceive-SS.PRIOR

      añŋ- magi plg-p.
      pulse.unit throb-PF.3S
      'When I touch your temple I can feel the pulse throb.'

    e. Tob-yp pub nŋ-i   gug   tk-p.
      foot-my sun perceive-SS.PRIOR crack cut.across-PF.3S
      'My feet have developed cracks from the sun.'

f. Yp lkañ-mñ   pag   ow-a-k.
    me blood-cord break come-3S-PAST
    'I had a miscarriage.' (lit. 'umbilical cord break it came')

g. Yad wsn nŋ-i,   kceki   nŋ-ab-in,
    I   sleep perceive-SS.PRIOR forest-goblin perceive-REC-IS

    ju   d-ab-in,   yp kawnan am-b.
    withdraw seize-REC-IS, me spirit   go-PF.3S
    'I dreamed that when I saw a forest demon, I was startled and my spirit
    left me.'

3.3 Cases Where the Role of the Condition Nominal Is Less Certain

Let us turn now to cases where the grammatical role of the Condition nominal is
somewhat less certain. These are chiefly expressions where the main verb is *g-* 'act,
do, make' or *ay-* 'become' and where the Condition nominal denotes a sensation or
feeling. Sections 3.3.1 and 3.3.2 introduce the kinds of Condition nominals that occur
with these verb stems and a later section (3.3.3) considers evidence concerning the
grammatical role of the Condition nominals.

3.3.1 Sensations and feelings with the verb *g-* 'happen, act, do, etc.'. The verb *g-*
can be used either intransitively ('happen, occur, act, function') or transitively ('do,
make, act or work on, affect something'). In BMP clauses, *g-* appears always to be
transitive, because the Experiencer is invariably marked by an Objective pronoun. *g-*
combines with nominals and adjectives representing several types of Condition,
including the following: (i) Sensations or processes that involve sensations such as
yuan 'hunger', *yuwt-bt* 'exhaustion', *pboŋ* 'heat, hot, warm', *ygen* 'cold (from wind)',
takl 'cold' (general)', *km* 'bitter', *ydk* 'good taste, tasty' (it also means 'salt'), *slk* 'hot-
tasting, pungent' (it also means 'itching'), *ñekñek* 'hiccups', *jlken* 'cough, head-cold',
tap 'sick, sickness', *kajknm* 'wince'. Names of sensations tend not to be so nouny as
most other Condition nouns. The sensation terms listed above cannot be possessed or
quantified. They can, however, be modified for intensity, as in (25).

(25) a. Yp   ydk   tmey g-p.
      me   good-tasting bad   do-PF.3S
      'It tastes awfully good to me.'

    b. Yp   takl   tmey g-p.
      me   cold   awful act-PF.3S
      'I am terribly cold.'

    c. Yp   ytuk   naban   g-p.
      me   lethargy extremely act-PF.3S
      'I feel extremely lethargic/lazy.'

However, a few sensation nouns, such as *yuwt* 'pain', can be possessed:

(26)   Yuwt yp   sayn g-p.
       pain me   weak do-PF.3S
       'My pain has eased.' / 'The pain has gone.'

(ii) Nouns denoting emotions and feelings, as for example *nabŋ* 'shame, shyness, embarrassment', or body parts that are the seat of emotions, such as *sb* 'bowels, sympathy', *mapn* 'liver, sympathy', *mluk* 'nose, bad humor'. Certain emotion words are not so nouny. They can be modified but they cannot be possessed or quantified.

(27) a. Yp nabŋ   yob g-a-k.
        me shame   big do-3S-PAST
        'I was very ashamed/shy/embarassed.'

     b. Yp sb   g-a-k.
        me bowels   act-3S-PAST
        'I was emotionally moved (upset, happy, angry).'

     c. Yp tep   yb   g-a-k.
        me good true do-3S-PAST
        'I felt truly happy/pleased.'

     d. Sb-wt   yp   tmey g-p.
        guts-cluster my   bad do-PF.3S
        (i) 'I feel angry/upset.' (ii) 'My guts feel bad.'

Note that in (27c–d) the adjectives *tep* 'good' and *tmey* 'bad' probably do not function as modifiers marking opposing extremes on a good–bad scale. Instead they refer to kinds of feelings: 'pleasure, feeling good' and 'unpleasant, feeling bad', analogous to 'feeling sick', 'being hungry', or 'being tasty'.

3.3.2 Feelings marked by derived nominal and *ay-* 'become, form, stabilize'.

An adjective or noun can be derived from any verb root by adding the suffix *-eb* or *-ep* 'V-ing, characterized by V-ing, associated with V-ing', for example, *ñŋ-* 'consume, eat' yields *ñŋ-eb* 'for consumption'. A complex noun can be formed by adding *-eb, -ep* to an Object + verb sequence. For example, *tap ñŋ-* 'eat food' yields *tap ñŋ-eb* 'thing for eating, food'; *kotp am-* 'go home' yields *kotp ameb* 'going home'. Any such derived word (call it X) can be used as a noun together with the clitic *-tek* 'like, resembling, as if' and the verb *ay-* 'become, form, stabilize' to form a construction (OBJ) *X-tek ay-*, which can be translated fairly literally as (a) 'X is about to/is likely to happen (to OBJ)' or (b) 'A need/desire to do X has come over OBJ'. A more colloquial translation of (b) is '(Experiencer) feels like doing X', as, for example:

(28) a. Yp kotp   am-eb   -tek ay-p.
        me house go-NDR   like form-PF.3S
        'I feel like going home.'

     b. Np   wok-ep   -tek ay-p?
        you vomit-NDR   like form-PF.3S
        'Do you feel like vomiting?'

c. Tap   wok-ep   -tek   ay-p!
thing   vomiting   -like   form-PF.3S
'It's a thing that makes one want to vomit!'

3.3.3 Do all experiencer object clauses have a subject? In Experiencer Object clauses, we know that the Experiencer NP is not the subject. The difficulty is to know which, if any, of the other NPs is Subject. Unfortunately, the candidates for Subjecthood are almost always inanimate NPs, and the two most reliable tests for identifying Subjects, namely (i) selection of a nominative pronoun and (ii) Subject-verb agreement, do not work for inanimate NPs. The first test does not work because inanimate nouns cannot be pronominalized (see 2.3.1). The second test does not work because in Experiencer Object clauses the verb "agreement" suffix is invariably 3rd singular. Pluralizing an inanimate NP makes no difference, for reasons noted in 2.5.2.

It might be argued that, if there is only one inanimate NP present in an Experiencer Object construction, that NP must, by default, be the Subject, and it must be that NP that governs verb agreement. However, in the light of comparative evidence, such an argument is not persuasive. It is a commonplace that some languages have "dummy subjects" manifested as a 3rd singular Subject pronoun (compare English 'It is a pity it is raining') or as a 3rd singular ending on the verb. As the next most useful diagnostic procedures for identifying Subjects, we are left with (i) seeking parallelism with canonical transitive clauses in the semantic relations between NP and verb, (ii) capacity to trigger switch reference, and (iii) tests designed to eliminate all but one candidate.

As we have seen, the first procedure works fairly well for those Experiencer Object clauses discussed in 3.2, but not for those exemplified in 3.3.1 and 3.3.2, where the semantic relation between the NP denoting a sensation or feeling is not transparently similar to that in canonical Subject-verb relations. Let us now turn to other tests.

3.3.3.1 Switch reference as a test of subjecthood. It was noted earlier (2.5, 2.6) that verbs that belong to separate clauses within the same sentence typically carry suffixes marking identity or change of Subject. (There are certain exceptions, chiefly having to do with embedded clauses, that we need not discuss here.) This system of "switch reference" marking provides a potential test of Subjecthood for the nominals in BMP constructions.

Each of the sentences in (29) consists of two or more clauses denoting a sequence of bodily processes. In each case the Subject of the first clause is the Experiencer NP and this NP is coreferential with the (overt or understood) Direct Object of the last clause.

(29) a. Ñb-e-n   (yp)   tep   g-p.
eat-DS.PRIOR.IS   me   good   act-PF.3S
'It tastes good.'(lit. 'When I eat, good/pleasure acts [on me].')

b. Añŋ   ay-e-y,   tep   md-p.
pulse   form-DS.PRIOR.2S.HORT   good   stay.PF.3S
'Your pulse is good.' / 'Your breathing is normal.'

172

c. Ñŋ-in!   ag-e-n,   (yp) yuwt g-s-a-p.
eat-IS-HORT   say-DS.PRIOR.IS me   pain do-PROG-3S
'It hurts when I try to eat.' (lit. 'When I say 'I'll eat!' pain is acting [on me].')

d. Ñb-e-n   amn-aŋ,   (yp) yuwt g-s-a-p.
eat-DS.PRIOR-IS go-HORT.IS me pain act-PROG-3S
'It hurts when I swallow.' (lit. 'When I eat it goes down, it is painful [to me].')

e. Sb   ki-ng   g-e-k,   (nup) kls   g-i,
faeces, guts excrete-SS.FUT do-DS.PRIOR.3S him   tight do-SS-PRIOR

(nup) plg-p.
(him) block-PF.3S
'He was constipated.' (lit. 'He tried to defecate, it was tight, it was blocked.')

The examples in (30) differ from those in (29) in that in each sentence *both* clauses have Experiencer Objects. In (30a) the first clause has an inanimate Subject: *sŋl* 'boil', with the Body-Part nominal as Locative Object; the second clause refers to the pain (*yuwt*)) caused to the (optionally specified) Experiencer. In (30b) the first clause also has a Body-Part as Locative Object ('our innards'), but the clause has an animate Object.

(30) a. Sŋl nñ-yp   ay-e-k,   (yp)   yuwt g-p.
boil arm-me form-DS.PRIOR.3S (me) pain do-PF.3S
'A boil has formed on my arm and it's painful (to me).'

b. Sb-wt-cnp gley boley ag-a-knŋ,   np yuan   g-p.
innards-our rumbling say-3S-DS.SIM us hunger act-PF.3S
'When our stomachs rumble we are hungry' ('When rumblings are sounding in our innards, hunger has worked on us.')

In (31) the first clause has an Experiencer Object, the second an Experiencer Subject. The first verb is marked for change of Subject.

(31)   Toytk   nup mñak g-e-k,   kum   md-e-k.
yesterday him sick do-DS.PRIOR-PAST.3S indisposed stay-DUR.3S-PAST
'Yesterday sickness affected him and he wasn't able to do anything.'

The fact that in (29)–(31) the verb in the penultimate clause is always marked for change of Subject is consistent with the pronominal marking, which shows the Experiencer is the Object of this clause. As the only remaining overt NP is the Condition nominal, we are left with this as Subject—unless we want to posit verbal sentences with no Subject.

The weakness in the switch reference test is that the marking of different Subject is only automatic when the first verb has an animate subject. The sentences in (32) each consist of two clauses with different Subjects. In (32a–c) the Subjects are two different Condition nominals. In (32d) they are a Condition and a Body-Part nominal. (See also [13b] in 2.6.2.)

(32) a. Sŋl 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).'

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

c. Yp   tap   g-i,   tap-ñŋeb ma-tek   ay-a-k.
   me   sickness   act-SS.PRIOR food   not-like form-3S-PAST
   'I got sick and lost my appetite.' (lit. 'Sickness having acted on me, a
   feeling for food did not form [in me].')

d. Nup   sud-cp   g-i,   (nup) bay   ay-a-k.
   him   malaria   do-SS.PRIOR   him   spleen   form-3S-PAST
   'After he got malaria, it affected his spleen.' (lit. 'Malaria having acted
   on him, his spleen formed, i.e., became enlarged.')

3.3.3.2 Clauses with both body-part and condition NP. Many BMP clauses with
Experiencer Object also contain both a Body-Part NP and a Condition NP, as in
(30a–b) above and (33). Which of these NPs is the Subject?

(33) a. Wak   ytuk   g-p.
        body lethargy do-PF.3S
        'The body feels lethargic.'/ 'Lethargy is affecting the body.'

     b. Wdn   lkañ   ya-s-a-w.
        eye   blood fall-PROG-3S
        'The eye is bleeding.' / 'Blood is coming from the eye.'

     c. Alkjon   magi   yk   yow-p.
        armpit   lump   open fall-PF.3S
        'The armpit has a lump' / 'A lump has swelled up in the armpit.'

A crucial observation is that if a pronominal possessor of the Body-Part nominal is
specified, that pronoun must be drawn from the Object set, as in (34):

(34) a. Wak- yp   ytuk   g-p.
        body-my (OBJ)   lethargy do-PF.3S
        'My body feels lethargic.'/ 'Lethargy is affecting my body.'

     b. Wdn-np   lkañ   ya-sa-w.
        eye-your (OBJ) blood fall-PROG-3S
        'Your eye is bleeding.' / 'Blood is coming from your eye.'

     c. Alkjon-nup   magi   yk   yow-p.
        armpit-his (OBJ) lump   open fall-PF.3S
        'His armpit has a lump.' / 'A lump has swelled up in his armpit.'

Thus, if there is an overt Subject in the sentences in (33) it can only be the Condition
NP. It is, however, possible to have an Object pronoun and a Body-Part NP as sepa-
rate phrases, as in (35). In such cases the Body-Part NP can be interpreted as a Loca-
tive Object (see also [26] and [27]).

174

(35) a. Nup wdn lkañ a-s-a-w.
  him eye blood fall-PROG-3S
  'His eye is bleeding.' (lit. 'Blood is falling from his eye.')

  b. Yp sb-wt tmey g-p.
   me guts bad do-PF.3S
   'I am upset.' (lit. 'A bad feeling is affecting me in the guts'.)

3.3.3.3 Clauses where there are two condition nominals. Some BMP clauses contain two condition nominals, neither of which is possessed. In such cases, three alternative interpretations of their grammatical relations suggest themselves. One is that one nominal modifies the other, forming a complex NP that is the Subject. In (36a) the first Condition nominal, *ñg-saki-nen* 'for liquor' arguably modifies *yuan* 'craving, hunger, thirst'. In (36b) *bok* 'pus' may modify *soy* 'sore'. In (36d) *mñak* 'non-serious, short-lived sickness' may modify *jlken* 'cough' or vice versa.

(36) a. Yp ñg-saki-nen yuan g-p.
  me liquor-for hunger make-PF.3S
  'I have a craving for liquor.' / 'A craving for liquor is affecting me.'

  b. Yp soy bok ay-p.
   me sore pus form-PF.3S
   'I have an infected sore.'/ 'Pus has formed in my sore.'

  c. Yp sŋl yuwt g-s-a-p.
   me boil pain act-PROG-3S
   'A boil is causing me pain.' / 'A painful boil is affecting me.'

  d. Yp jlken mñak g-p.
   me cough sickness do-PF.3S
   'I am a bit sick with a cough.' ('A cough sickness is affecting me.')

A second possibility is that one of the Condition NPs is the Subject and the other is an Adjunct NP (36a–c). A third is that one Condition NP is Subject and the other a verbal adjunct in a complex verb. Both the first and second analyses make sense when one of the nominals carries the clitic *-nen* 'purpose: for, after', as in (36a) and (37a–b), where the *-nen* phrase nominal could be either a reason or cause Adjunct of the verb or a modifier of the following nominal. The third interpretation makes better sense in the case of (36d).

(37) 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.'

3.3.3.4 Constructions where the sole condition nominal is not the subject. There are certain Experiencer Object constructions containing a single Condition nominal

that is apparently not the Subject. This class is illustrated by the three paraphrases in
(38a–c), all grammatical and all translatable as 'I don't like the taste of pork' or 'Pork
is not tasty to me'.

(38) a. Kaj   yp   ydk   ma-g-p.
       pork me   tasty not-do-PF.3S

   b. Yp   kaj   ma-ydk   g-p.
      me   pork not-tasty do-PF.3S

   c. Yp   kaj   ydk   ma-g-p.
      me   pork tasty not-do-PF.3S

In (38) the sensation noun *ydk* is best interpreted as a verbal adjunct that is part of a
complex verb: *ydk g-* 'be tasty'. The Subject of *gp* 'it does' is presumably *kaj* 'pig,
pork', naming the source or cause of the sensation, yielding the literal translation 'Pork
does not cause tastiness in me'. It is possible, however, to argue that *kaj* is a peripheral
Adjunct ('Tastiness does not act on me with respect to pork.'). Compare (36c) where
one might propose competing grammatical analyses corresponding to either of two lit-
eral glosses: 'A boil is causing me pain' or 'Pain is affecting me in a boil.'

3.4 Conclusion

Experiencer Object clauses are a mixed bag. In many cases (3.2) the Condition nominal
is clearly the Subject. In others (3.3) it is hard to find compelling grounds for choosing
between an analysis that treats the Condition nominal as a "funny" Subject and one that
treats it as a verbal adjunct within a complex verb. The clear cases are chiefly those
where the Condition nominal denotes a visible and/or stable entity. The problematic
cases are those where the Condition nominal denotes a sensation, emotion, or feeling.
When we consider the sharp semantic difference between these two classes of Condi-
tion nominal, it is not surprising that they behave somewhat differently.

4. Experiencer Subject Constructions

Section 2 began with a binary division of BMP clauses, on grammatical grounds, into
Experiencer Subject and Experiencer Object types. We noted that this grammatical
division is apparently motivated by semantics. In one class of expressions, the Expe-
riencer typically initiates the process and so has some measure of control over it. In
the other class, the Experiencer is the involuntary undergoer.

On closer examination, the correlation between syntactic and semantic types turns
out to be less than perfect. Experiencer Subject constructions are not a unified class
with respect to the semantic role of the Subject. Although in most cases the Subject is
clearly the initiator and controller of the process, there are a fair number of exceptions
to this equation. This section will first briefly review a small sample of Experiencer
Subject constructions and then seek to make sense of the exceptions.

4.1 Types of Experiencer Subject Constructions

4.1.1 Perception and cognition, marked by *nŋ-* 'perceive'. The verb *nŋ-* 'perceive, sense, be aware, cognize' is neutral between intent and lack of intent. It can be translated, in context, as any of the following: 'see, hear, listen, smell, taste, feel, know, think, understand, remember, try, notice, recognise, consider, observe'. When there is intent *nŋ-* always takes an Experiencer Subject. Deliberate feeling and tasting are treated as two successive actions by the Subject, as in (39).

(39) a. Yad   kaj   d   nŋ-b-in.
   I   pig   hold   perceive-PF-IS
   'I (deliberately) felt the pig.'

  b. Yad   kaj   ñb   nŋ-b-in.
    I   pig   consume   perceive-PF-IS
    'I tasted the pig.'

Smelling something accidentally and on purpose can be expressed by either an Experiencer Object or an Experiencer Subject construction:

(40) a. (Yp)   kaj   kuy   ow-p.
   (me)   pig   odor   come-PF.3S
   'I can smell pork/pig.'

  b. (Yad)   kaj   kuy   nŋ-b-in.
    (I)   pig   odor   perceive-PF.IS
    'I smell pork/pig.'

A distinction between different kinds of perception can be shown either by naming a specific organ as an instrument (41a–b), or by naming the thing produced or perceived (41c–d). Note that although seeing and hearing may be involuntary experiences, they are expressed with an Experiencer Subject.

(41) a. Yad   wdn   nŋ-b-in.
   I   eye   perceive-PF-IS
   'I saw (it)'.

  b. Tumuk ag-e-k,   yad   (tmd)   nŋ-b-in.
    thunder sound-DS.PRIOR-PAST   I   ear   perceive-PF-IS
    'I heard thunder (with my own ears).' ('Thunder sounded, I heard with my ears.')

  c. Yad   gos   nŋ-sp-in.
    I   thought   perceive-PROG-IS
    'I am thinking.'

  d. Nad   wsn   nŋ-b-an.
    you   sleep   perceive-PF-2S
    'You had a dream.'

177

4.1.2 Sound-making processes, marked by *ag-*. The verb *ag-* 'sound, say' refers to making any kind of sound. The person or thing making the noise is always the Subject of *ag-*, regardless of whether the noise is made deliberately or involuntarily. The object of *ag-* can be a vocal product, such as *mnm* 'speech, *mukbel* 'belch', *guglum* 'snore', *suk* 'shout', *gullag* 'croak'.

(42) a. Yad   mnm a-sp-in.
        I   talk   say-PROG-IS
        'I am talking.'

     b Nad   si   a-sp-in.
        you   cry   say-PROG-2S
        'You are crying.'

     c. Guglum a-sp-an.
        snore   say-PROG-2S
        'You are snoring.'

4.1.3 Consuming, marked by *ñŋ-*. With the verb *ñŋ-* or *ñb-* 'consume, eat, drink, chew, smoke, suck, nick, nip', the person or thing doing the consuming is always the Subject.

(43) a. Yad   tap   ñ-b-sp-in.
        I   food   consume-PROG-IS
        'I am eating.'

     b. Nuk   snb   sat   ñb-s-a-p.
        he   ginger chewing   consume-PROG-3S
        'He is chewing ginger.'

     c. Yp   ña-pan-ŋaŋ puŋi ñb-s-a-p.
        me   baby   press eat-PROG-3S
        'I have labor pains.' (lit. 'Baby is pressing and eating me.')

(43c) has both an Experiencer Subject and Experiencer Object. The Subject is the baby, who is both 'pressing' (*puŋi-* 'pierce, poke, press against a point) and 'eating, nipping'. The Object is the mother in labor, represented here by *yp* 'me'.

4.1.4 Excreting and laying eggs, marked by *ki-*. The verb *ki-* 'excrete' is used for defecating, urinating, and laying eggs. The excreter is always the Subject, as in (1c) and the following:

(44) a. Nuk   sb   ki-a-k.
        He   faeces excrete-3S-PAST
        'He defecated.'

     b. Yakt   magi   omŋal   ki-p.
        bird   egg   two   excrete-PF.3S
        'The bird has laid two eggs.'

c. Np   yman   magi   ki-p.
you (OBJ) lice   egg   excrete-PF.3S
'You have nits' (lit. 'Lice have laid eggs on you.')

4.1.5 Sleeping and lying down, marked by *kn-*. The verb *kn-* has two senses: 'lie down, recline' and 'sleep'. These can be disambiguated by adding the verbal adjunct *wsn* 'sleep(ing).

(45) a. Kik   wsn   k-jp-ay.
     they sleep lie-PROG-3P
     'They are sleeping.'

b. Yad (wsn) kn tep   g-p-in.
    I   (sleep) lie good do-PF-IS
    'I slept well.'

c. Kik   wsn   k-jp-ay.
     they sleep lie-PROG-3P
     'They are sleeping.'

Feeling sleepy is expressed by an Object Experiencer construction (see 18a).

4.1.6 Dying or being nonfunctional, marked by *kum-*. The verb *kum-* has several senses: 'die', 'be nonfunctioning, unwell', 'have no feeling, be numb or paralyzed'. In the normal course of events dying and being nonfunctional are not done on purpose, but in Kalam these senses occur with an Experiencer Subject. Note the following contrast between (46a), where the people dying are the Subject and (46b) where those feeling like or in danger of dying are the Object:

(46) a. Bin-b   og-ok kum-ya-k.   Koyb   pk-ya-k.
     People the-PL die-IS-PAST witches   strike-3P-PAST
     'They died. Witches killed (them).'

b. Kuyp kum-eb   tek   ay-p.
    them die-NDR like form-PF.3S
    (i) 'They felt like dying.' (ii) 'They almost died.' (lit. 'Like-dying formed in them.')

4.2 Apparent Exceptions to the Principle That Experiencer Subject Equals Initiator/controller.

A sizeable minority of Experiencer Subject expressions denote acts or processes that are *not* initiated or controlled by the Experiencer. We referred above to the way dying or being nonfunctional is expressed. Forgetting is not, usually, a deliberately initiated or controlled process, but in Kalam it is expressed with an Experiencer Subject clause that means, literally, 'make (something) out of mind'.

(47)   Yad yb-nad   saki   g-p-in.
       I   name-your out-of-mind do-PF.IS.
      'I have forgotten your name.'

179

Feeling sorrow or sympathy, or being in a bad mood, can be expressed either with an Experiencer Object or with an Experiencer Subject construction, without apparent semantic difference, as in the pairs (48a, b) and (49a, b):

(48) a. Yp sb g-p.
    me guts act-PF.3S
    'I feel upset.'

  b. Yad mapn nŋ-b-in.
    I liver perceive-PF-IS
    'I feel sorrow/sympathy.'

(49) a. Yp mluk yow-p.
    me nose fall-PF.3S
    'I am in a bad mood.' (lit. 'Nose has fallen in me.')

  b. Yad nup mluk nŋ-b-in.
    I him nose perceive-PF-IS
    'I feel angry at him.' ('I perceive him with the nose.')

Sleeping and dreaming are both expressed by a Subject Experiencer construction.

(50) a. Kayn og-ok wsn kn-ya-k.
    dog the-pl sleep lie/sleep-3P-PAST
    'The dogs were asleep.'

  b. Yad wsn nŋd nŋ-sp-in.
    I sleep true perceive-PROG-IS
    'I had a true dream.'

Being startled or taken by surprise is usually expressed by an Experiencer Subject clause:

(51) a. Yad ju d-p-in.
    I withdraw hold-PF-IS
    'I was taken by surprise.'

  b. Monmon ap-ab, bin-b glk-p-ay.
    earthquake come-REC.3S people startle-PF-3P
    'When there's an earthquake, people are startled.'

Being afraid or frightened can be expressed by either an Experiencer Object or an Experiencer Subject construction, as in (52).

(52) a. Yp jel g-p.
    me fear/nervousness do.PF.3S
    'I'm afraid/nervous.'

  b. Yad ptk-sp-in.
    I afraid-PROG-IS
    'I'm afraid.'

Giving birth and being someone's parent are expressed by an Experiencer Subject construction, using the verb *tk-* 'give birth, be parent (father or mother) to an offspring'.

(53)   Bin   ak   ña-pan-ŋaŋ   tam-sek   tk-a-k.
          woman that   baby   joined   give.birth-3S-PAST
          'That woman gave birth to twins.'

4.3 Can the Exceptions Be Explained?

Experiencer Subject constructions are best regarded as a semantically disparate class with two main subclasses: (1) a majority of members in which the Experiencer is the initiator or semantic Agent, and (2) a fair number of exceptions—cases where the Experiencer is the involuntary Patient.

Can the exceptions be explained in a systematic way? We believe they can, at least to some extent. Subclass 2 does not consist of a wholly arbitrary set of exceptions to subclass 1, all of which are indistinguishable, semantically from Experiencer Object constructions. It seems that, in languages around the world, processes like sleeping and dreaming, seeing, hearing and feeling something accidentally by touch, belching and snoring, and forgetting and dying tend to be pushed into the same sorts of syntactic frames as wilful actions. In the Kalam case it would be stretching things to try to offer a semantic explanation for every apparent exception, but some sort of systematic explanation may be found for most cases. Some can be explained either in terms of speakers giving higher priority to particular semantic components of a verb and its arguments other than the involuntary component, or in terms of speakers drawing analogies with other uses of the verb where the Experiencer is plainly the initiator.

Seeing, hearing, and feeling accidentally, for instance, are perceptual processes, processes closely related to watching, listening, and feeling deliberately, which involve a focusing of the perceiver's attention. Accordingly, it is more natural to attribute control or agency to these processes than, say, to feeling itchy or hungry.

While sleeping is not an Experiencer-controlled process, it can be argued that it begins when the Experiencer decides to lie down and rest. In Papuan languages generally, 'lie down' and 'sleep' are expressed by the same verb. Dreaming is hardly Experiencer-controlled but it is expressed using the verbal adjunct *wsn* 'sleeping' plus the verb *nŋ-* 'perceive, cognize', and in Kalam that verb often refers to deliberate acts.

Why is dying normally encoded with an Experiencer Subject in Kalam and nearly every language, rather than, say, as 'death comes to one'? Perhaps because it is opposed to *md-* 'live, be alive, stay, remain', a stative verb that is also Experiencer Subject.

5. Bodily and Mental Processes Expressed by Clause Sequences or Serial Verb Constructions

Not all bodily and mental processes are expressed by a simple sentence with one verb. A number require a two- or three-clause construction, or a single "extended" clause containing a serial verb sequence. We must limit discussion to a few examples.

Consider the senses seeing, hearing, smelling, tasting, and feeling (by touch). In English we say *I saw the pig, I heard the pig, I smelled the pig*, etc. using single clause structures identical save for the verb of sensing. The earlier discussion of Kalam single clause constructions for BMPs and their correlation with semantic types may lead the reader to expect *seeing* and *hearing* to be treated similarly in Kalam. In fact, they are treated very differently. To say 'I saw the pig', Kalam need only mention (as English does) the Experiencer, the sensing, and the concrete Object that is sensed.

(54) Yad kaj (wdn) nŋ-b-in.
      I   pig (eye) perceive-PF-1S
      'I saw the pig.'

But to say 'I heard the pig', one must say 'I perceived the pig make a (certain kind of) sound', or 'I perceived the sound made by the pig' with an embedded clause as the Object of 'perceive'.

(55) Yad [kaj ag-e-k]   (tmd) nŋ-b-in.
      I   pig sound-DS.PRIOR.3S   (ear) perceive-PF-1S
      'I heard the pig.'

It is not necessary to mention the source Object (this may not be known, in some instances), but it is necessary to mention the sound that is heard, and specifically, the act of sound-making. Thus, in Kalam one cannot say *I heard you;* one must say *I heard you speak/call/cry,* etc. A partial explanation for this may reside in the commonsense view of visual and auditory perception. Only scientists think of seeing in terms of the eyes receiving light waves radiated by an object. The ordinary man thinks of the eyes perceiving the Object itself. The connection is a direct one. Not so with hearing. Sounds obviously come from a source, but exist independently of the source. We can hear a shout without seeing the shouter, and we can hear an echo reverberate as it travels across an enclosed space.

Similarly with smelling. Odor moves from source to smeller. The Kalam say 'odor comes to me, I perceive (with the nose)', although the second clause, denoting the act of perception, is optional. *I smelled the pig* is expressed in full as:

(56) [kaj kuy   ap-e-k],   (mluk) nŋ-b-in.
      pig   odor come-DS.PRIOR.3S   (nose) perceive-PF-1S

As with hearing, one can leave out the source object, but one cannot leave out the stimulus. Although hearing and smelling are expressed somewhat similarly in Kalam (and both are treated quite differently from seeing) they are not given identical syntax. Again, it can be argued that the syntactic differences reflect commonsense awareness of objective differences. Whereas a sound is normally produced by an event of sound-making—someone talking, a tree falling, water moving over rocks—a smell is not necessarily the result of an event. Instead, the odor coming from an object is often a persisting attribute, just as its shape and color are. However, the Kalam speak of an odor (but not shape and color) as something that moves through the air and is perceived with the nose.

Tasting and feeling are different again, in Kalam. As with hearing and smelling, the constructions are complex, but the complexities are different. Once more, the differences reflect real world differences. In Kalam, deliberate feeling and tasting take agent subjects and involve two successive actions by the subject.

(57)   Yad kaj d   ŋ-b-in.
          I   pig touch perceive-PF-1S
          'I felt the pig (intentionally).'

(58)   Yad kaj ñb   nŋ-b-in.
          I   pig consume perceive-PF-1S
          'I tasted the pig (intentionally).'

The use of two verbs here is typical of Kalam treatment of closely related event sequences that share a single actor while the use of a single clause with a single verb 'taste' or 'feel' in the English translation is typical of English treatments, where such event sequences are often compressed into a single semantically complex verb with its associated NP.

Accidental feeling and tasting are depicted differently from deliberate acts of feeling or tasting. In the former case, the focus is usually on the outcome, the resulting sensation. But, in both cases, the specific sensation experience can only occur if there are prior acts. Thus, in order to experience a certain kind of taste, the source must be consumed or at least put in the mouth, and, in order to feel an object, one must first come into contact with it. In Kalam, a full expression of *The axe felt sharp to me* involves saying, roughly, *I touch perceived the axe, it was sharp.*

(59)   Tu   d   nŋ-e-n,   ytk   g-p.
          axe touch perceive-DS.PRIOR-1S sharp do-PF-3S
          'The axe felt sharp to me.'

6. Conclusions

In Kalam expressions for bodily and mental processes there is high correlation between a semantic parameter—the animate experiencer's control vs lack of control over initiation of a process—and a syntactic parameter—choice of Experiencer Subject vs. Experiencer Object construction. There are, however, a number of exceptions. The most clearcut exceptions are cases where an uncontrolled process is expressed by an Experiencer Subject construction. In addition, when we come to examine variation within subtypes of Experiencer Subject constructions, we also find some apparently arbitrary pairings of meaning with grammatical type. In general, it appears that there are always more semantic contrasts than there are grammatical constructions. Speakers prefer to economize on the number of grammatical patterns, and to tolerate form-meaning mismatches, rather than to increase the number of grammatical types. However, the exceptions are not entirely arbitrary.

How much variation is there in the way different languages express bodily and mental processes? Here is a challenging research project for someone. Casual com-

parisons suggest that languages vary considerably in the kinds of semantic variables they make pivotal when assigning a particular semantic type of BMP to a particular grammatical construction. When we think of the complexity of variables involved in BMP processes it would be surprising if it were otherwise. Besides the opposition control vs. lack of control over the initiation or the entire course of a BMP, there are other oppositions—visible vs. invisible, transient vs. longlasting, stable vs. dynamic, pleasant vs. unpleasant, and bodily vs. mental, and so on—any of which might be chosen as the basis for a primary division into grammatical types. At the same time, basic patterns in BMP constructions tend to be strongly persistent within language families, such as the large Trans New Guinea Phylum to which Kalam belongs.

To what extent do different encodings of BMPs (and other events) present different perceptions of "the same" events? McElhanon (1978, 1992) takes up this question, drawing chiefly on data from a Trans New Guinea Phylum language, Selepet, and Tok Pisin. He concludes that these two languages present a radically different world view from that of English in respect of the causes of people's physical and psychological states and in respect of the degree of control that human beings exert over the elements in their world. He suggests, further, that the linguistic differences are commensurate with differences in cosmology and behavior—that Papua New Guinea people have different perceptions of and reactions to natural events from, say, Americans. But these are complex issues that lie beyond the scope of this essay.

REFERENCES

Farr, Cynthia. 1997. The interface between syntax and discourse in Korafe, a Papuan language of Papua New Guinea. Ph.D. dissertation, Department of Linguistics, Research School of Pacific and Asian Studies, Australian National University. In press in Pacific Linguistics.

Foley, William A. 1986. *The Papuan languages.* Cambridge: Cambridge University Press.

Givón, Talmy, 1990. Verb serialization in Kalam: A comparative study of temporal packaging. In *Typological studies in language,* ed. by John Verhaar, 19–55. Amsterdam: Benjamins.

Grace, George W. 1987. *The linguistic structure of reality.* London: Croom Helm.

Lane, Jonathan, 1991. Kalam serial verb constructions. M. A. thesis, Department of Anthropology, University of Auckland.

McElhanon, Kenneth A. 1978. On the origin of body image idioms in Tok Pisin. *Kivung* 11.1:3–25.

———. 1992. On the concept of person. In *The language game: Papers in memory of Donald C. Laycock,* ed. by Tom Dutton, Malcolm Ross, and Darrell Tryon, 241–255. Pacific Linguistics C-110. Canberra: Australian National University.

Pawley, Andrew, 1966. The structure of Kalam. Ph.D. dissertation, University of Auckland.

———. 1987. Encoding events in Kalam and English: Different logics for reporting experience. In *Coherence and grounding in discourse,* ed. by Russell S. Tomlin, 329–360. Amsterdam: John Benjamins.

———. 1993. A language that defies description by ordinary means. In *The role of theory in language description,* ed. by William A. Foley, 87–129. Berlin: Mouton de Gruyter.

Pawley, Andrew, and Jonathan Lane. 1998. From event sequence to grammar: Serial verb constructions in Kalam. In *Case, typology and grammar,* ed. by Anna Siewierska and Jae Jung Song, 201–227. Typological Studies in Language 38. Amsterdam: John Benjamins.

Pawley, Andrew, and Ralph Bulmer, with J. Kias, I. S. Majnep, and S. P. Gi. In press. A dictionary of Kalam with ethnographic notes. Pacific Linguistics Series C. Canberra: Australian National University.

Roberts, John. 1997. Switch reference in Papua New Guinea: A preliminary survey. In *Papers in Papuan linguistics*, no. 3, 101–241. Pacific Linguistics A-87. Canberra: Australian National University.

12:   ON NONVERBAL PREDICATES IN THAI

Kitima Indrambarya
Kasetsart University

1. Introduction[1]

Generally, every sentence consists of a subject and a predicate. Most predicates contain a verb as head of the construction, as in (1). However, not all predicates are verbs, as in (2) and (3).

(1) malii   láaŋcaan
    Malee   washing dishes
        +V
    'Malee is washing dishes.'

(2) wan níi   wansǎw
    day   this Saturday
        +N
    'Today is Saturday.'

(3) bâan   níi   sǎmràp thúkkhon
    house this for   everyone
        +P
    'This house is for everyone.'

Without a verb in sentences (2) and (3), each of these constructions has been referred to as a "verbless predication" (Haas 1964:xx), an "equation construction" (Noss 1964:75), and a "verbless sentence" (Warotamasikkhadit 1969:74).

Within Transformational Grammar, Warotamasikkhadit (1969:75–76) explored similar constructions and concluded that all sentences in Thai need to have an underlying verb in the deep structure. Hence, for sentence (2), he posits the verb *khii* 'to be' in the underlying structure, as shown in (4). Similarly for (3), there is the verb *chây* 'to be true' with a repeated noun in the underlying structure, as shown in (5).

(4) wan níi   khii wansǎw
    day   this be   Saturday
    'Today is Saturday.'

(5) bâan   níi   chây bâan sǎmràp thúkkhon máy
    house this true house for   everyone or not
    'Is this house the house for everyone?'

According to Warotamasikkhadit (1969:75–76), the underlying verbs *khɨ̀ɨ* and *chây* can be deleted by a transformation. Even if we were to increase the expressive power of the theory to allow for an extra abstract level of structure, Warotamasikkhadit's claim would still be suspicious. First of all, *khɨ̀ɨ* may not occur with any negation word, as illustrated in (6), whereas the predicate noun in (2) may occur with the negation word *mâychây*, as shown in (7). This fact suggests that (2) and (4) simply bear different syntactic structures and that (4) cannot be the underlying structure of (2). Besides, the noun phrase that occurs after the verb *chây* does not necessarily have to be a repeated noun; a semantically coreferential noun phrase is also acceptable, as shown in (8).

(6) *wan níi   mây/mâydây/mâychây khɨ̀ɨ wansǎw
    day this NEG   be   Saturday
    'Today is not Saturday.'

(7) wan níi   mâychây wansǎw
    day this NEG   Saturday
    'Today is not Saturday.'

(8) lǎŋ níi chây bâan sǎmràp thúkkhon máy
    CL   this true house for   everyone or not
    'Is this house the one for everyone?'

Moreover, to derive the surface structure (3), Warotamasikkhadit has to apply a deletion transformation to *chây* and an erasure transformation to the noun *bâan*, which by his analysis is the head noun of an NP constituent, with the prepositional phrase *sǎmràp thúkkhon* 'for everyone' as its dependent. This latter transformation is illegitimate even in a powerful framework such as the conventional transformational grammar, in that it deletes only a part of a constituent. Hence, (5) cannot be the underlying structure of (3).

Working within Lexicase dependency grammar, a monostratal grammar in which such abstractness is not possible, I contend that sentences (2) and (3) contain no verbs. While most sentences in Thai contain verbal predicates, other sentences contain nonverbal predicates.

The remainder of this essay is divided into four sections. Section 2 provides a background of the term "predicate" in Lexicase, followed by tests for different types of predicates in section 3. Section 4 explores the occurrence of nonverbal predicates in Thai. The final section presents the conclusion.

2. Predicate

Within Lexicase dependency grammar, a predicate is the lexical head of a sentence or a clause. It can be a verb, a noun, or a preposition. There is, however, no instance of a predicate adverb or a predicate adjective.[2] This predicate analysis has an advantage over other theories in that there is no need to postulate a null main verb for a verbless sentence (Starosta 1988:235).
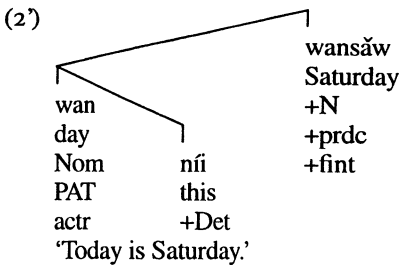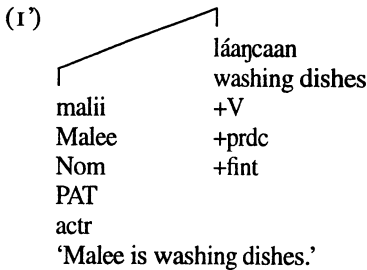
Most sentences contain a verbal predicate, that is, a predicate that is a verb. Verbs are prototypical predicates. Hence, a verb is redundantly marked as predicate (prdc), as

shown in RR-1. Moreover, the unmarked value of a predicate is intransitive. Thus, any nonverbal predicate is marked with the feature [−trns] (intransitive). A transitive verb, on the other hand, is lexically marked [+trns] and hence, not affected by the following RR-2.

RR-1:   [+V]   →   [+prdc]
RR-2:   [+prdc] →   [−trns]

The feature [+/− prdc] (predicate) is optional for a preposition and a noun. The term "nonverbal predicate" refers to prepositions and nouns when functioning as predicates. Notice that it is the function as predicate, that is, position in the structure, that enables the N or P to be marked [+prdc], and nothing inherent in its properties. A nominal predicate is semantically defined as a nominal that does not have its own referent (Starosta 1988:237). Examples (1'), (2'), and (3') illustrate the tree structures of sentences containing verbal and nonverbal predicates in Thai.

(1')

             láaŋcaan
             washing dishes
malii   +V
Malee   +prdc
Nom   +fint
PAT
actr
'Malee is washing dishes.'

(2')

             wansǎw
             Saturday
wan   +N
day   +prdc
Nom   níi   +fint
PAT   this
actr   +Det
'Today is Saturday.'

---

2. In Thai, words that correspond to the class of English adjectives such as *wǎan* 'sweet' are analyzed as stative verbs, as shown in the following example:

    sôm   wǎan   caŋ
    orange sweet much
          +V
          +sttv
    'The orange is very sweet.'

    Based on Savetamalya's analysis (1989:163–168), adjectives in Thai are dependents of classifiers only. With the given definition of predicate, I found none of the adjectives to function as predicate.

(3')

```
                            ┌──────────────
              ┌───────────  │ sămràp
              │             │ for          │
    ┌─────────╲            │ +P            │
    │ bâan     ╲           │ +prdc          thúkkhon
    │ house     ╲  ┐       │ +fint          everyone
    │ Nom      níi         │                Acc
    │ PAT      this         │               COR
      actr     +Det
```

'This house is for everyone.'
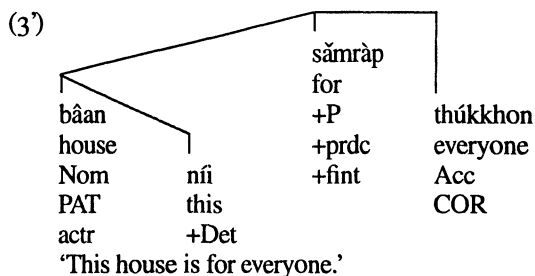
3. Types of Predicates

This section discusses tests used to identify different types of predicates. It is divided into tests for predicates and tests for types of predicates.

3.1 Tests for Predicates in Thai

This section investigates two types of categories: nouns and prepositions. The two tests determining predicates are the root predicate and the regular Patient-to-Actor Control Chaining Rule (P2a).
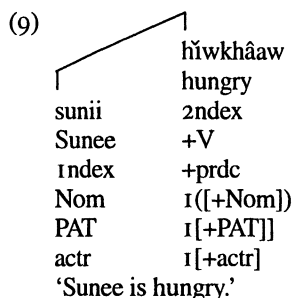
3.1.1 Root predicate. Generally, a predicate is able to occur in a root clause, that is, a finite matrix clause. A root predicate clause is defined as a clause that contains only one predicate, that is, a simple sentence. Unlike English, Thai has no affixes marking the finiteness of a predicate. A finite predicate is identified by the presence of a Nominative actant. As a pro-drop language, Thai may manifest a clause without a subject or an object of the verb, if it is retrievable within context. The optionality in the presence of a subject and/or an object can be shown by the parentheses in the lexical matrix of the "regent predicate." Hence, in (9), the verb $hiwkhâaw$ has the feature $[1([+Nom])]$ in its lexical matrix, indicating that it expects $sunii$ as its nominative actant. According to the Patient Centrality Hypothesis, all verbs need a Patient in their case frames. A nominative actant is the Patient actor of an intransitive verb. This can be shown formally in the lexical matrix of the verb as $[1[+PAT]]$ and $[1[+actr]]$.

(9)

```
          ┌───────────┐
          │           │ hǐwkhâaw
    ┌─────╱            hungry
    │ sunii            2ndex
    │ Sunee            +V
    │ 1ndex            +prdc
    │ Nom              1([+Nom])
    │ PAT              1[+PAT]]
      actr             1[+actr]
```
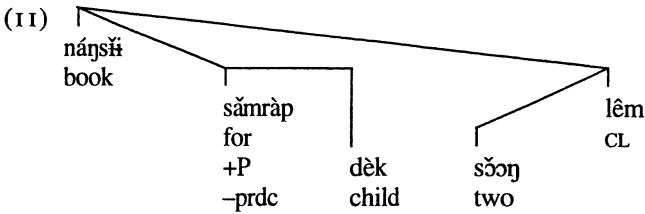
'Sunee is hungry.'

Since verbs always function as predicates, they are considered lexical predicates and hence, can occur in a root clause, as illustrated above. A lexical predicate refers to

189

a category whose sole function is that of being a predicate. Neither nouns nor prepositions are lexical predicates. While prepositions like *sămràp* 'for' may function as a predicate, as in (3), others such as *kàp* 'with', as illustrated in (10), cannot serve the same function. Again, *sămràp* is not a lexical predicate because the prepositional phrase headed by it may also function as a noun modifier, as in (11).
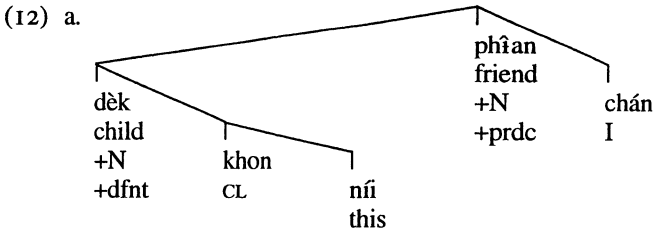
(10) *dèk khon níi kàp phîan
    boy CL   this with friend
              +P
              −prdc
    'This kid is with a friend.'

(11)
    náŋsŭ
    book
            sămràp   lêm
            for   CL
            +P   dèk   sɔ̌ɔŋ
            −prdc   child two
    'two books for children'

On the other hand, all nouns may potentially function as predicates, provided that in certain specified structures, the nominative or the subject must be definite, as illustrated in (12a–b).

(12) a.
            phîan
            friend
    dèk   +N   chán
    child   I
    +N   khon   +prdc
    +dfnt   CL
                    níi
                    this
    'This kid is my friend.'

    b. *dèk   phîan   chán
       child   friend I
       +N   +N
       −dfnt   −prdc
       'A kid is my friend.'

Because adverbs and adjectives in Thai may not occur as a predicate in a simple sentence, as illustrated in (13b) and (14b), respectively, the test indicates that they may not function as predicate. Example (14b) would be acceptable only as a noun phrase, not a sentence.

(13) a. kháw waaŋ náŋsɨ̌ wáy thîi nîi
      he   lay book lying at here
                       +Adv
                       −prdc
      'He left his book here.'

    b. *kháw wáy   náŋsɨ̌ thîi nîi
      he   lying book at here
              +Adv
              −prdc
      'He left his book here.'

(14) a. kháw mii   náŋsɨ̌ sǎam lêm
      he   have book three CL
                  +Adj
                −prdc
      'He has three books.'

    b. *náŋsɨ̌ sǎam lêm
      book   three CL
      'There are three books.'

3.1.2 The regular patient-to-actor control rule (P2a). The regular P2a control rule in Lexicase can be used to find a predicate in a nonfinite complement. It is a general rule accounting for the missing actor of a nonfinite complement clause. It is applicable for all nonfinite complements that meet the structural description of the rule (Starosta 1988:235). In other words, a nonfinite predicate, whether it be a noun or a verb, must undergo this rule if the structural description is met.[3]

Since a predicate is head of each clause, a sentence with more than one predicate, one being embedded under another, must be a complex sentence. That is, a "regent" or a higher predicate takes another predicate as its complement. The complement can

---

3. In Lexicase, there are three types of control chaining rules for complements: Patient-to-Actor control rule (P2a), Patient-to-Patient control rule (P2P), and Actor-to-Actor control rule (a2a). The last two rules are more specific than the P2a in that they are triggered by a [+cntn] and [+mnnr] verbs, as illustrated by examples a. and b., respectively. Hence, by convention, they take precedence over the regular P2a rule, which is more general (Indrambarya 1994:297–306).
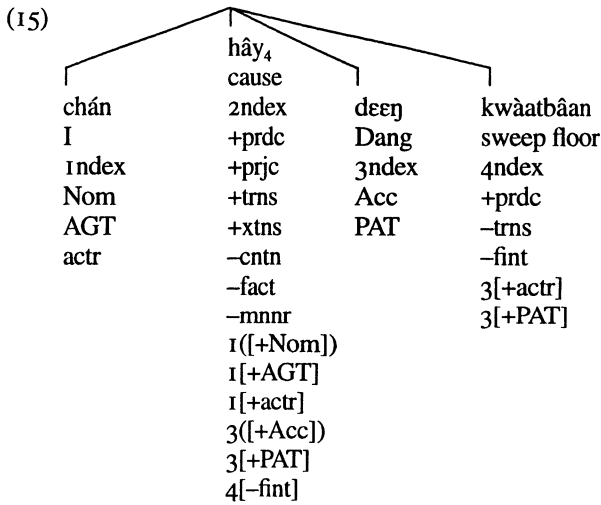
    a. dεεŋ thùuk   mɛ̂ɛ   dùʔ
       Dang undergo mother reproach
       Nom **+cntn** Acc   −fint
       PAT −trns   COR
       'Dang was reproached by his mother.'

    b. kháw dəən   kin   ʔaytim
       he   walk   eat   ice cream
       Nom **+mnnr** −fint
       PAT −trns
       actr
       'He walked eating ice cream.'

be either finite or nonfinite. A finite predicate is defined as a predicate with a nominative actant in its case frame. Without a nominative actant, the embedded predicate must be nonfinite. All nonfinite predicate complements must undergo a control chaining rule to find and interpret their missing actants. The actor of the nonfinite nominal predicate can be identified with the help of the P2a control chaining rule.

The following Patient-to-Actor control chaining rule states that a nonfinite predicate complement finds its missing actor by copying the index of the Patient of the regent verb (Starosta 1998:323–324).

$$\text{CR-1} \begin{bmatrix} ?[+\text{actr}] \\ -\text{fint} \\ \text{nndex} \end{bmatrix} \rightarrow [m[+\text{actr}]] \quad / \quad \begin{bmatrix} m[+\text{PAT}] \\ n[-\text{fint}] \end{bmatrix}$$

(15)

chán	hây₄ cause	deɛŋ	kwàatbâan
I	2ndex	Dang	sweep floor
1ndex	+prdc	3ndex	4ndex
Nom	+prjc	Acc	+prdc
AGT	+trns	PAT	−trns
actr	+xtns		−fint
	−cntn		3[+actr]
	−fact		3[+PAT]
	−mnnr		
	1([+Nom])		
	1[+AGT]		
	1[+actr]		
	3([+Acc])		
	3[+PAT]		
	4[−fint]		

'I made Dang sweep the floor.'

As shown, the compound verb *kwàatbâan* 'to sweep the floor', as a nonfinite complement, looks for an actor nominal rather than a nominative nominal. The compound verb then finds its actor by copying the index 3 of [3[+PAT]] on the regent verb, or more specifically the matrix verb *hây*, resulting in [3[+actr]], satisfying the requirement for the P2a control rule. The verb *kwàatbâan* is thus shown to be predicate. According to the Patient Centrality hypothesis, which states that every verb needs a Patient in its case frame, a Nominative or the actor of an intransitive verb is interpreted as Patient (Starosta 1988:182–183). Hence, the higher Patient *deɛŋ* is interpreted as the Patient actor of the verb *kwàatbâan*. It should be noted that this control chaining rule is applicable to both verbal and nonverbal nonfinite complements. Examples of nonverbal nonfinite complements are discussed in section 4.2.2. Because the compound verb *kwàatbâan* may satisfy the requirement for this control chaining rule, it is a predicate.

3.2 Tests for Identifying Types of Predicates

After having identified the predicate, one needs to know what type of predicate it is. Is it a verb, a noun, or a preposition? The following tests are employed to identify the type of a predicate: choice of negation word, determiner insertion, and potential subject.

3.2.1 Choice of negation word in a root predicate. In Thai, the negation word *mây* 'not' may precede a verb or a sentence-final adverb. Even though *mây* may not differentiate verbs from adverbs when they occur after another verb, as shown in (16) and (17), the negation adverb *mây* may distinguish verbs from adverbs when it precedes the forms in question in a root predicate clause.

3.2.1.1 *Mây* **before an adverb occurring sentence-finally**

(16) kháw pay₃⁴ thîi₄ nân mây bɔ̀y
      he go at there NEG often
      +V +Adv
      'He does not often go there.'

3.2.1.2 *Mây* **before an embedded verb**

(17) kháw kradòot mây khâam rúa
      he jump NEG cross fence
      +V +V
      'He did not succeed in jumping over the fence.'

3.2.1.3 Negation in root clauses. In a root predicate clause, only a verb may occur in construction with the negation word *mây*, as shown in (18–20). Both root predicate nouns and root predicate prepositions occur in construction with *mâychây* 'not true', rather than *mây* 'not'. An adverb, on the other hand, may not occur at all as a root predicate in a sentence, as shown in (21a), and hence cannot occur with the negation word *mây* in (21b).

(18) kháw mây pay₁
      he NEG go
      +V
      −trns
      'He is not going.'

(19) náŋsɨ̌ níi *mây/mâychây khɔ̌ɔŋ thəə
      book this NEG POSS you
      +N
      +prdc
      'This book is not yours.'

---

4. Subscripts are used as labels for homophonous words.

(20)   náŋsɨ̌ɨ nɪ́ɪ   *mây/mâychây   sǎmràp thəə
        book this   NEG   for   you
                                          +P
                                        +prdc
    'This book is not for you.'

(21) a. *kháw   wáy   náŋsɨ̌ɨ thîi nîi
       he   lying   book   at   here
              +Adv
              −prdc
    'He left a book here.'

     b. *kháw mây   wáy   náŋsɨ̌ɨ thîi nîi
        he   NEG   lying   book   at here
                   +Adv
                   −prdc
    'He did not leave a book here.'

3.2.2 Determiner insertion. Thai is a head-initial language and thus allows a modifier to follow its head. Only a noun allows a determiner such as *nîi* 'this' or *nán* 'that' as its dependent, as in (22); a verb and a preposition do not, as shown in (23) and (24).

(22)   rûup   nîi   sǔay
        picture this   beautiful
        +N   +Det
    'This picture is beautiful.'

(23)   *chán wâat nîi   rûup
        I   draw this picture
              +V   +Det
    *'I this draw picture.'

(24)   *kháw kinkhâaw dûay nîi[5]
        he   eat   with this
        +P   +Det
    *'He eats this with.'

3.2.3 Potential subject. The ability to occur as subject of a sentence is another characteristic of a noun. As illustrated in (25), a noun may occur as subject of a sentence, whereas a verb and a preposition, or even a prepositional phrase, cannot, as in (26–28), respectively.

---

5. Note that sentence (24) would be acceptable if the preposition were used with the demonstrative pronoun *nîi* 'this' instead.

    kháw kinkhâaw dûay nîi
    he   eat   with this
                      +P   +N
    'He eats with this.'

(25) náŋsʉ̌ nîi dii
book this good
+N +Det +V
'This book is good.'

(26) *thamŋaan dii
to work good
+V +V
'Work is good.'

(27) *kàp dii
with good
+P +V
*'With is good.'

(28) *kàp chán dii
with me good
+P +V
*'With me is good.'

A summary of tests identifying the types of predicates is given in table 1.

Table 1. Tests Identifying Types of Predicates

	CHOICE OF NEGATION WORD		DETERMINER INSERTION	POTENTIAL SUBJECT
	mây	mâychây		
VERBS	+	−	−	−
NOUNS	−	+	+	+
PREPOSITIONS	−	+	−	−

4. Nonverbal Predicates in Thai

Only two types of nonverbal predicates are found in Thai. They are prepositional predicates and nominal predicates.

4. 1 Prepositional Predicates

While all verbs are lexical predicates, prepositions may be predicate or nonpredicate [+/−prdc] (Starosta 1988:236). In the course of investigating nonextension prepositions in Thai, subcategorized in Indrambarya (1994:41–48), only sǎmràp 'for' is able to function as predicate.[6] In other words, the preposition sǎmràp 'for' is the only predicate preposition found. Consider (29).

---

6. Within Lexicase dependency grammar, nonextension prepositions are prepositions that take a noun phrase as cohead. Extension prepositions, on the other hand, are those taking a clause as cohead. The latter are commonly referred to as subordinate clauses elsewhere.

(29) a.

```
                            ┌─────────────────────┬─────────┐
                            │  sǎmràp             │         │
                            │  for                │         │
┌─────────┐                 │  +P                 │  wítthayakɔɔn
│ dɔɔkmáy ├──────────┐      │  +prdc              │  guest lecturer
│ flower  │  ┌───────┴──┐   │                     │  +N
│ Index   │  │ chɔɔ     │   │                     │  Acc
│ Nom     │  │ bouquet  ├──┐│                     │  COR
│ PAT     │  │ 2ndex    │  │└──────┐
└─────────┘  │ CL       │  │  níi  │
             └──────────┘  │  this │
                           │  3ndex│
                           │  +Det │
                           └───────┘
```
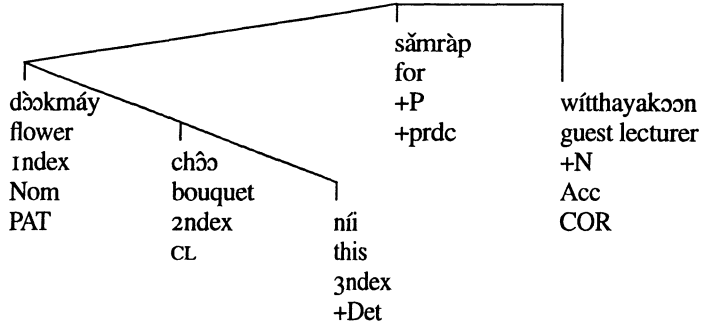
'This bouquet of flowers is for the guest lecturer.'

As shown in (29a), *sǎmràp* can function as predicate of the sentence. While *sǎmràp* may not be identified absolutely as a preposition in (29b), because a noun can also occur with the negation *mâychây,* the fact that it does not allow a dependent determiner in (29c) confirms its not being a noun, but a preposition.

(29) b. dɔɔkmáy chɔɔ   níi mâychây/*mây/*mâydây sǎmràp
       flower   bouquet this NEG   for

       wítthayakɔɔn
       guest lecturer
       'This bouquet of flowers is not for the guest lecturer.'

    c. *dɔɔkmáy chɔɔ   nii mâychây/mây/mâydây sǎmràp níi
       flower   bouquet this NEG   for   this
       'This bouquet of flowers is not for this.'

4.2 Nominal Predicates

Nouns have two functions: nonpredicate [–prdc] and predicate [+prdc]. A predicate noun functions as a predicate and does not have its own referent. A nonpredicate, on the other hand, bears a case relation: Agent, Patient, Locus, Correspondent, or Means. There is no lexical predicate noun, because no noun functions as a predicate in all circumstances.

Nominal predicates may occur in root clauses (matrix clauses), nonroot clauses (finite embedded clauses), and nonfinite embedded complements. The feature [+/– root] is used with a finite clause. A finite predicate can be subcategorized into [+root] and [–root], as shown in figure 1.

Figure 1. Subcategorization of a finite predicate

```
        [+fint]
        /     \
   [–root]   [+root]
```
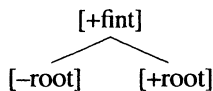
(30) Wichian bòn wâa kháw mây yàak klàp bâan
Wichian complain that he NEG want return home
+V +V +V
+prdc +prdc +prdc
+fint +fint –fint
+root –root
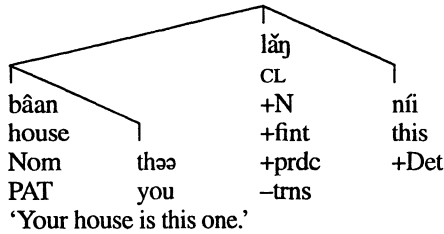'Wichian complained that he did not want to return home.'

Within the Lexicase analysis of (30), *bòn* and *yàak* are both finite verbs. While *bòn* occurs in a root clause, *yaàk* is a finite verb occurring in a nonroot clause. In other words, *bòn* is a verb in the matrix clause taking a finite verb complement in which *yaàk* is the head. The verb *klàp*, on the other hand, is a nonfinite verb in an embedded clause.

The following subsections discuss the occurrence of the predicate nouns in root clauses and as nonfinite complements.

4.2.1 Predicate nouns in a root clause. This section discusses the occurrence of classifiers, the possessive noun *khɔ̌ɔŋ*, color words beginning with *sǐi*, and other potential nominal predicates.

4.2.1.1 Classifiers. Following Savetamalya (1989:160), this analysis considers classifiers, as exemplified in (31), to be nominal predicates. As shown below, the classifier is the regent of the sentence.[7]

(31) a.

lǎŋ
CL
bâan +N nǐi
house +fint this
Nom thəə +prdc +Det
PAT you –trns
'Your house is this one.'

In sentence (31a), the nominative nominal *bâan* does not function as a predicate and is assigned the Patient case relation. The classifier *lǎŋ* functions as the predicate and does not have a case relation. Within an appropriate context, sentence (31a) may be negated with *mâychây*, as in (31b), giving evidence that the classifier is definitely not a verb. The acceptability of the sentence having the determiner *nǐi* as a dependent of the

7. Savetamalya also analyzes classifiers that function as attributes of a noun to be predicate nouns like those in a nominative relative clause, based on the semantic interpretation and on the linking rule. That is, the classifier *lêm* finds its higher actor from the noun *náŋsǔɨ*, as with any other nonfinite predicate.

náŋsǔɨ sǎam lêm
book three CL
+N +Adj +N
PAT +prdc
'books that are three volumes'

classifier *lǎŋ* illustrates that it must be a noun. Supporting evidence can also be shown in (31c), in which *lǎŋ*, the classifier of *bâan*, can occur as subject of the sentence.

(31) b. bâan thəə mâychây lǎŋ nîi
house you NEG CL this
'Your house is not this one.'

c. lǎŋ nîi lék mâak
CL this small much
'This one (house) is very small.'

4.2.1.2 The possessive noun *khɔ̌ɔŋ₂*. Following Savetamalya (1989:192), this analysis considers the possessive noun *khɔ̌ɔŋ₂* to be a predicate noun.[8] As illustrated in (32a), the possessive noun *khɔ̌ɔŋ₂* has the noun *sîa* as its nominative Patient. The cooccurrence with the negation word *mâychây* in example (32b) shows that *khɔ̌ɔŋ₂* is not a verb, and the fact that *khɔ̌ɔŋ₂* can occur as a subject in (32c) illustrates further that it is a noun.

(32) a.

khɔ̌ɔŋ₂
POSS

sîa		3ndex	thəə
shirt	nii	+N	you
Nom	this	+fint	COR
PAT		+prdc	

'This shirt is yours.'
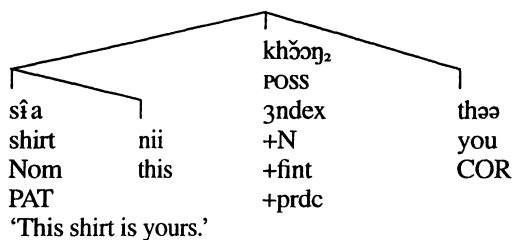
b. sîa nîi mâychây/*mây khɔ̌ɔŋ₂ thəə
shirt this NEG POSS you
'This shirt is not yours.'

c. khɔ̌ɔŋ₂ chán ʔan nîi
POSS I CL this
'Mine is this one.'

4.2.1.3 Color words beginning with *sǐi-*. Color words beginning with *sǐi-* may have a predicate function, as shown in (33a). Sentence (33b) shows that *sǐikhǐaw* can be negated only with *mâychây*, suggesting that it is not a verb. Sentence (33c) then tells us that *sǐikhǐaw* is a noun, because it can occur as subject of the sentence.

It should be noted that a color word alone, if possible, is a stative verb meaning 'to become . . .', as can be tested with the word *lέεw* 'already' and the comparative degree *kwàa* 'more . . . than'. However, not all color words beginning with *sǐi-* have corresponding stative verb forms.

---

8. Savetamalya (1989:192) considers there to be two homophonous words *khɔ̌ɔŋ*. The one, *khɔ̌ɔŋ₁* 'thing', is a nonpredicate noun, while the other, the possessive noun *khɔ̌ɔŋ₂*, is a classifier and a predicate noun.

(33) a.

		sĭikhĭaw
		green color
rót		+N
car		+fint
Nom	chán	+prdc
PAT	I	−trns

'My car is green.'

b. rót chán mâychây/*mây/*mâydây sĭikhĭaw
car I NEG green color
'My car is not green.'

c. sĭikhĭaw mòt
green use up
'The green color is used up.'

(34) sîa thəə dam/khăaw/dɛɛŋ/khĭaw/lĭaŋ/*fáa/*námtaan/*mûaŋ lɛ́ɛw
shirt you black/white/red/green/yellow/blue/brown/purple already
'Your shirt became black/white/red/green/yellow/blue/brown/purple already.'

(35) sîa tua níi dam/khăaw/dɛɛŋ/khĭaw/lĭaŋ/*fáa/*námtaan/*mûaŋ
shirt CL this black/white/red/green/yellow/blue/brown/purple

kwàa tua nán
than CL that
'This shirt became more black/white/red/green/yellow than that shirt.'

As illustrated in (34) and (35), while the predicate nouns *sĭidam* 'black color', *sĭikhăaw* 'white color', *sĭidɛɛŋ* 'red color', *sĭikhĭaw* 'green color', and *sĭilĭaŋ* 'yellow color' have corresponding stative verb forms without the prefix *sĭi-*, with the meaning 'to become . . . (the color indicated by the root word)', the others do not.

4.2.1.4 Potential nominal predicates. All nouns are potential predicates when cooccurring with a definite nominative actant to express identification. Predicate nouns can be the copy of the nominative subject (semantically-related to the nominative actant); nouns indicating time such as *mooŋ* 'hour', *wan* 'day', *dian* 'month, *pii* 'year'; a specific time or date; or certain other entities.

4.2.1.4.1 Predicate noun as a copy of the nominative. In (36a), the first nominal *bâan* is a nonpredicate noun that is assigned a nominative Patient. The second *bâan* is a predicate noun, and as expected, may only occur with the negation word *mâychây*.
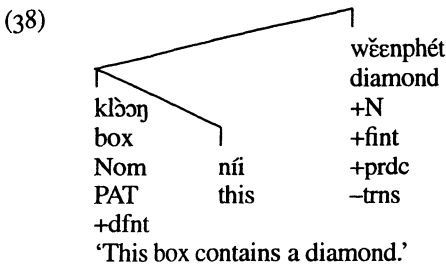
(36) a.

		bâan	
		house	
bâan		+N	fɛ̂ɛt
house		+fint	twin
Nom	níi	+prdc	
PAT	this	−trns	
+dfnt			

'This house is a twin house.'

b. bâan   níi   mâychây/*mây/*mâydây   bâan   fɛ̀ɛt
house   this   NEG   house   twin
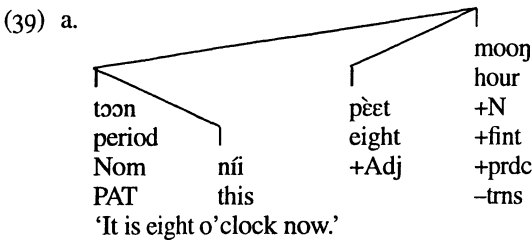'This house is not a twin house.'

4.2.1.4.2 Predicate noun as semantically related to the nominative. In both (37) and (38), the predicate nouns and the nominative actant are semantically related.

(37)

```
                          rot
                          car
khan                      +N        chán
CL                        +fint     I
Nom       níi             +prdc
PAT       this            –trns
+dfnt
```
'This one is my car.'
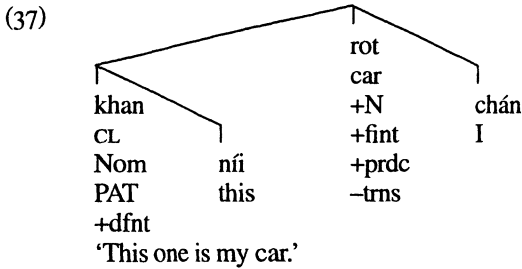
The predicate noun *rot* 'car' in (37) has its corresponding classifier *khan* as a nominative Patient actant, while the predicate noun *wɛ̌ɛnphét* 'diamond' in (38) has a container, *klɔ̀ɔŋ* 'box', as its nominative Patient actant.

(38)

```
                          wɛ̌ɛnphét
                          diamond
klɔ̀ɔŋ                    +N
box                       +fint
Nom       níi             +prdc
PAT       this            –trns
+dfnt
```
'This box contains a diamond.'

4.2.1.4.3 Nouns indicating time. In (39a), *mooŋ* 'hour' functions as a predicate noun. The fact that it may only occur with *mâychây* as illustrated in (39b) and that it may occur as a subject in (39c) confirms that it is a noun.

(39) a.

```
                                    mooŋ
                                    hour
tɔɔn                      pɛ̀ɛt      +N
period                    eight     +fint
Nom       níi             +Adj      +prdc
PAT       this                      –trns
```
'It is eight o'clock now.'

b. tɔɔn   níi   mâychây/*mây/*mâydây   pɛ̀ɛt   mooŋ
period   this NEG   eight   hour
'It is not eight o'clock now.'

c. pɛ̀ɛt   mooŋ   pen₂ weela thîi chán tùin
eight   hour   be   time   that I   wake up
'Eight o'clock is the time I wake up.'

Now consider (40).[9]

(40) a.
```
            ┌──────────┐
           /           wansùk
          /            Friday
     phrûŋníi    +N
     tomorrow    +prdc
     Nom         +fint
     PAT         −trns
     actr
```
'Tomorrow is Friday.'

b. phrûŋníi mâychây/*mây/*mâyday wansùk
   tomorrow NEG   Friday
   'Tomorrow is not Friday.'

c. wansùk níi   pen₂ wankàət chán
   Friday this be   birthday I
   'This Friday is my birthday.'

That the predicate *wansùk* may only coocur with *mâychây*, as in (40b), and can function as subject, as in (40c), shows that it is a nominal predicate.

4.2.1.4.4 Other predicate nouns. Any noun may also function as a predicate provided that the nominative actant is definite, as in (41) and as illustrated earlier in (12).

(41)
```
            ┌──────────┐
           /           wěɛnphét
          /            diamond
     níi         +N
     this        +fint
     Nom         +prdc
     PAT         −trns
     +dfnt
```
'This is a diamond."

4.2.2 Predicate nouns as nonfinite complements. Predicate nouns, but not predicate prepositions, are found within a nonfinite complement. The complement can be either finite or nonfinite. Without a Nom(inative) feature in its case frame, the embed-

---

9. It should be noted that the nominative Patient *phrûŋníi* 'tomorrow' in (40) is different from *phrûŋníi* in the following sentences, in which 'tomorrow' does not function as a subject and may occur either at the beginning or at the end of the sentence.

   a. phrûŋníi rooŋrian chán yùt
      tomorrow school   I   stop
                              +V

   b. rooŋrian chán yùt phrûŋníi
      school   I   stop tomorrow
                   +V

ded predicate must be nonfinite. A predicate noun functioning as a complement of a verb is analyzed as a nonfinite predicate because it does not allow the presence of a subject. Example (42b) illustrates that the sentence becomes unacceptable with the presence of the nominative actant *Nida.* Hence, the predicate noun *khruu* is a nonfinite complement.

(42) a. nidaa   pen₂   khruu
        Nida   be   teacher
        1ndex   2ndex   3ndex
        Nom   +V   +N
        PAT   +fint   +prdc
        actr   −trns   −fint
        'Nida is a teacher.'

    b. *nidaa   pen₂   nidaa   khruu
        Nida   be   Nida   teacher
        1ndex   2ndex   3ndex   4ndex
        Nom   +V   +N   +N
        PAT   +fint   +prdc
        actr   −trns   *3([+Nom])
                1([+Nom])
        'Nida is a teacher.'

This predicate noun complement has no independent reference, but rather adds information about the controller in the regent clause (cf. Starosta 1988:237). On the other hand, a nonpredicate noun is perceived as a separate entity encoding a separate case relation (cf. Starosta 1988:115).

Nonfinite noun complements are required to satisfy the regular P2a control rule, in which the lower actor must be coreferential with the higher Patient. Predicate nominal complements are found to be embedded within two verb classes: extension nonfact nonprojection intransitive verbs and extension nonfact nonprojection transitive verbs.[10]

Nonfact nonprojection verbs are verbs that expect a nonfinite predicate noun complement, as can be shown by the following redundancy rule.

$$\text{RR-3} \quad \begin{bmatrix} -\text{fact} \\ -\text{prjc} \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} ?[\ +\text{N}\ \ ] \\ [\ +\text{prdc}\ ] \\ [\ -\text{fint}\ \ ] \end{bmatrix}$$

4.2.2.1 Nonfact nonprojection simple intransitive verbs. These verbs require a predicate noun phrase as their dependent. Verbs of this class include (a) noncopula verbs such as *chây* 'to be true', *chîi* 'to be named', *naamsakun* 'to be surnamed', *mii₅* 'to include', *ruam* 'to include', *thâwkàp* 'to equal', *thaa₅* 'to paint', etc.; and (b) the copula verbs *pen₂* 'to be' and *khîi* 'to be'.

4.2.2.1.1 Noncopula verbs. Consider the nonfinite predicate noun embedding under a noncopula verb in (43).

(43) a.

		chɨ̂ɨ	
		name	
khon		3ndex	kaan
person		+fint	Karn
ɪndex	nɨ́i	–cpla	4ndex
Nom	this	–fact	+N
PAT	2ndex	–trns	+prdc
actr	+Det	–prjc	–fint
		ɪ([+Nom])	–trns
		ɪ[+PAT]	ɪ[+actr]
		ɪ[+actr]	
		4[–fint]	

$$4 \begin{bmatrix} +N \\ +prdc \\ -fint \end{bmatrix}$$
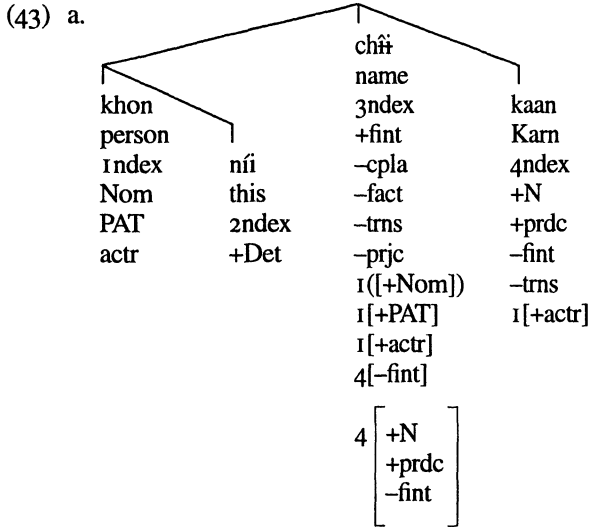
'This person is named Karn.'

b. *khon   nɨ́i   kháw   kaan
person this   he   Karn
'This person is Karn.'

---

10. Indrambarya (1994:221) classifies extension verbs or verbs that require a sentential complement into either fact or nonfact and either projection or nonprojection, as in SR-2. Fact verbs expect a finite verb complement, as shown in RR-9. As RR-10 shows, projection verbs expect a nonfinite verb complement. RR-3 shows that nonfact nonprojection verbs expect a nonfinite nominal predicate complement.

RR-8   [+xtns]   →   [?[+prdc]]

SR-2   [+xtns]   → $\begin{bmatrix} \pm fact \\ \pm prjc \end{bmatrix}$

RR-9   [+fact]   → $\begin{bmatrix} ?[+V \quad] \\ [+fint\ ] \end{bmatrix}$

RR-10   [+prjc]   → $\begin{bmatrix} ?[+V \quad] \\ [-fint] \end{bmatrix}$

RR-3 $\begin{bmatrix} -fact \\ -prjc \end{bmatrix}$ → $\begin{bmatrix} ?[+N \quad] \\ [+prdc\ ] \\ [-fint\ ] \end{bmatrix}$

203

The features [4[+N, +prdc, −fint]] in the lexical matrix of the verb *chîi* indicate that *chîi* expects a nonfinite predicate noun as complement. The predicate noun *kaan* is a nonfinite complement because it does not allow the presence of the nominative actant *kháw*, as in (43b). Being a nonfinite complement, the predicate noun *kaan* is required to undergo the P2a control chaining rule. Hence, *kaan* finds its missing actor by copying the index 1 from the higher Patient of the regent verb *chîi*, resulting in having [1[+actr]] on the lexical matrix of *kaan*. When testing with the root predicate test, the noun *kaan* may occur as a predicate, as illustrated in (43c). Notice that the brackets in [4[−fint]] in (43a) indicate that the embedded clause is a complement. The parentheses such as [4([−fint])] instead would indicate that the embedded clause is an adjunct.

(43) c.
khon	níi	kaan
person	this	Karn
Nom		+N
PAT		+fint
actr		+prdc

'This person is Karn.'

d. *dèk   kaan
   child   Karn
   'A child is Karn.'

Note, however, that the selectional restriction requires only a definite nominative to cooccur with root predicate nouns in order for it to be interpreted as being identified. Hence, an indefinite noun, as in (43d), is unacceptable.

There remains a problem, as some nonfinite predicate nouns may occur as a root predicate only in the presence of the negation word *mâychây,* but not when occurring alone, as exemplified in (b) and (c) of sentences (44) and (45). Further investigation is needed to shed light in this area. Notice also that the sentence is still acceptable with the reversed order in (45d). However, the selectional restriction forbids the summing number to occur as a nominative of a root predicate noun, as in (44d).

(44) a.
sɔ̌ɔŋ	bùak	sɔ̌ɔŋ	thâwkàp	sìi
two	plus	two	equal	four

'Two plus two equals four.'

b.
*sɔ̌ɔŋ	bùak	sɔ̌ɔŋ	sìi
two	plus	two	four

'Two plus two is four.'

c.
sɔ̌ɔŋ	bùak	sɔ̌ɔŋ	mâychây/*mây/*mâydây	hâa
two	plus	two	NEG	five

'Two plus two is not five.'

d.
*hâa	mâychây/*mây/*mâydây	sɔ̌ɔŋ	bùak	sɔ̌ɔŋ
five	NEG	two	plus	two

'Five is not two plus two.'

(45) a.
khɔ̌ɔŋᵢ	thîi	síi	maa	mii₅	phàk	phǒnlamáy	lɛ́ʔ	nom
thing	that	buy	thither	include	vegetable	fruit	and	milk

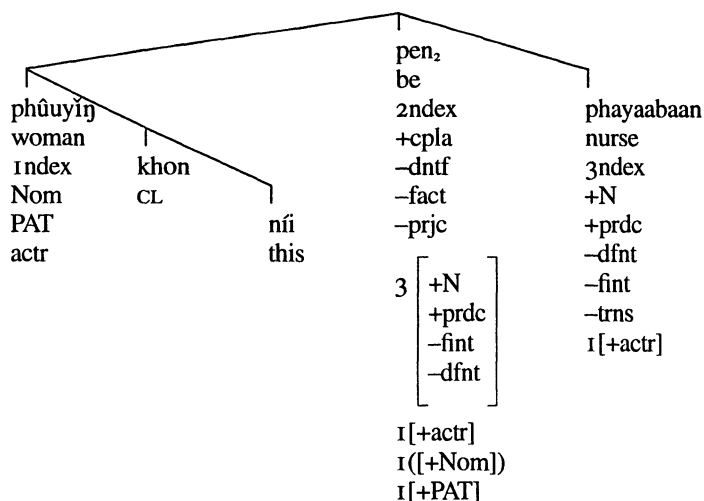'What I bought included vegetables, fruit, and milk.'

b. *khɔ̌ɔŋᵢ thîi sɨ́ɨ maa phàk   phǒnlamáy lɛ́ʔ nom
   thing that buy thither vegetable fruit   and milk
   'What I bought was vegetables, fruit, and milk.'

c. khɔ̌ɔŋᵢ thîi sɨ́ɨ maa   mâychây/*mây/*mâydây phàk
   thing that buy thither NEG   vegetable

   phǒnlamáy lɛ́ʔ nom
   fruit   and milk
   'What I bought was vegetables, fruit, and milk.'

d. phàk   phǒnlamáy lɛ́ʔ nom mâychây/*mây/*mâydây
   vegetable fruit   and milk NEG

   khɔ̌ɔŋᵢ thîi sɨ́ɨ maa
   thing that buy thither
   'Vegetables, fruit, and milk are not what I bought.'

4.2.2.1.2 Nonidentifying copula verb. Now consider example (46) in which a predicate noun is embedded under the nonidentifying copula verb *pen₂* 'be'. RR-4 states that a nonidentifying copula expects an indefinite noun predicate as complement. RR-5 shows that the feature [±dfnt] is used only with a noun.

RR-4: $\begin{bmatrix} +\text{cpla} \\ -\text{dntf} \end{bmatrix}$ → $\begin{bmatrix} ?[+\text{prdc}] \\ [-\text{dfnt}] \end{bmatrix}$

RR-5: [±dfnt] → [−trns]

(46) a.

```
                                  pen₂
                                  be
   phûuyîŋ                        2ndex         phayaabaan
   woman                          +cpla         nurse
   ɪndex      khon                −dntf         3ndex
   Nom        CL                  −fact         +N
   PAT               níi          −prjc         +prdc
   actr              this                       −dfnt
                                    ⎡ +N    ⎤   −fint
                                  3 ⎢ +prdc ⎥   −trns
                                    ⎢ −fint  ⎥   ɪ[+actr]
                                    ⎣ −dfnt ⎦

                                  ɪ[+actr]
                                  ɪ([+Nom])
                                  ɪ[+PAT]
```

'This woman is a nurse.'

b. *phûuyĭŋ khon níi pen₂ phayaabaan khon nán
   woman CL this be nurse CL that
   Nom   +cpla +prdc
              –dntf +dfnt
   'This woman is a nurse.'

The features [3[+N, +prdc, –fint, –dfnt]] on the lexical matrix of the nonidentifying copula verb *pen* indicate that it requires a nonfinite indefinite predicate noun as its complement. The presence of a definite noun as predicate complement results in the unacceptability of (46b). As a nonfinite predicate complement, the noun *phayaabaan* finds its missing actor from the higher Patient *phûuyĭŋ*, according to the P2a control rule.

4.2.2.1.3 Identifying copula verb. Unlike the situation in which *pen* is the verb, the predicate noun dependent of the identifying copula *khɨɨ* refers to the entity in the context of discourse that is identified with its subject and thus must be interpreted as definite, as shown in RR-6. *Khɨɨ* in (47a) identifies *dɛɛŋ* as a particular nurse, and not just any nurse. The features [3[+prdc, +dfnt]] show that *khɨɨ* expects a definite noun predicate as complement. Hence, the indefinite noun in (47b) produces an unacceptable sentence.

RR-6: $\begin{bmatrix} +cpla \\ +dntf \end{bmatrix} \rightarrow \begin{bmatrix} ?[+prdc] \\ [+dfnt] \end{bmatrix}$

(47) a.

```
              khɨɨ
              be
dɛɛŋ         2ndex      phayaabaan
Dang         +cpla      nurse
Index        +dntf      3ndex        khon
Nom          –fact      +N           CL      nán
PAT          –prjc      +dfnt        +N      that
actr         1[+PAT]    +prdc                +Det
             3[–fint]   –fint
             1[+actr]   1[+actr]
             1([+Nom])
```

$3 \begin{bmatrix} +prdc \\ +dfnt \end{bmatrix}$

'Dang is that nurse.'

b. *dɛɛŋ khɨɨ phayaabaan
   Dang be nurse
   'Dang is a nurse.'

4.2.2.2 Nonfact nonprojection simple transitive verbs. Simple transitive verbs expecting a nonfinite predicate noun complement are limited in number. They include *rabaay* 'to paint', *thaa₆* 'to paint', *thaasĭi₆* 'to paint', and *tèŋ₅* 'to decorate'. These verbs require their predicate complement nouns to be nouns expressing color only, as shown in RR-7.

$$
\text{RR-7:} \quad
\begin{bmatrix} \text{+trns} \\ \text{-fact} \\ \text{-prjc} \end{bmatrix}
\rightarrow
\begin{bmatrix} \text{?[+prdc]} \\ \text{[+colr]} \end{bmatrix}
$$

(48) a.

naaraa	thaa₆	rúa		sĭidɛɛŋ
	paint			
Nara	2ndex	fence	nán	red color
ɪndex	+trns	3ndex	that	5ndex
Nom	+xtns	Acc	4ndex	+N
AGT	-fact	PAT	+Det	+prdc
actr	-prjc			+colr
	ɪ([+Nom])			-fint
	ɪ[+AGT]			3[+actr]
	ɪ[+actr]			
	3([+Acc])			
	3[+PAT]			

$$
5\begin{bmatrix} \text{+N} \\ \text{+prdc} \\ \text{-fint} \\ \text{+colr} \end{bmatrix}
$$

'Nara painted that fence red.'

In (48a), the features [5[+N, +prdc, -fint, +colr]] indicate that the regent needs a nonfinite predicate color noun as a complement. The predicate noun phrase *sĭidɛɛŋ* 'red color' satisfies this requirement. According to the P2a control chaining rule, the nonfinite predicate noun finds its actor from the higher Patient *rúa* 'fence'.

b. *naaraa thaa₆ rúa nán sĭinámman
   Nara   paint fence that   enamel[-colr]
   'Nara painted that fence with enamel.'

While the color *sĭidɛɛŋ* 'red color' may cooccur with *thaa₆* 'to paint', a material object used in painting may not occur with *thaa₆*, as shown in the unacceptable (48b). Examples (49a–b) illustrate that both *sĭidɛɛŋ* and *sĭinámman* may occur as root predicates with the negation word *mâychây*. Hence both can function as predicates. Yet only *sĭidɛɛŋ*, which is specified with the semantic feature [+colr] in the lexicon, may occur as a complement of *thaa₆* in (48).

(49) a. rúa   nán   *mây/mâychây   sĭidɛɛŋ
        fence   that   NEG   red color
        Nom   +N
        PAT   +colr
        actr   +prdc
                                       −trns
        'That fence is not red.'

    b. krapɔ̌ɔŋ nán   *mây/mâychây   sĭinámman
        tin   that   NEG   enamel
        Nom   +N
        PAT   +prdc
        actr   −colr
        'That tin is not enamel.'

5. Conclusion

By employing the root predicate test, the regular Patient-to-Actor control chaining rule (P2a), choice of negation words, determiner insertion, and potential subject, we have found that nonverbal predicates are of two types: nominal and prepositional. Nominal predicates may occur in both root clauses and in nonfinite complements, whereas prepositional predicates may occur in root clauses but not in nonfinite complements.

REFERENCE

Haas, Mary R. 1964. *Thai–English student's dictionary*. Stanford: Stanford University Press.

Indrambarya, Kitima. 1994. Subcategorization of verbs in Thai: A lexicase dependency approach. Ph.D. dissertation, University of Hawai'i.

———. 1996. Are there prepositions in Thai? In *Papers for the Third Annual Meeting of the Southeast Asian Linguistic Society*, ed. by Mark Alves, 101–117. Tempe: Arizona State University.

Noss, Richard B. 1964. *Thai reference grammar*. Washington, D.C.: Foreign Service Institute.

Savetamalya, Saranya. 1989. Thai nouns and noun phrases: A lexicase analysis. Ph.D. dissertation, University of Hawai'i.

Starosta, Stanley. 1988. *The case for lexicase: An outline of lexicase grammatical theory*. London: Pinter Publishers

———. 1998. Lexicase revisited. In *Studies in linguistics in honour of Béla Korponay: A special issue*, ed. by József Andor, Béla Hollósy, Tibor Laczkó, and Péter Pelyvás, 279–334. Debrecen: Kiadta a Kossuth Lajos Tudományegyetem, Angol-Amerikai Intézete.

Warotamasikkhadit, Udom. 1969. Verbless sentences in Thai. *Linguistics: An International Review*. 47:74–79.

13: DOUBLE OBJECT CONSTRUCTIONS IN THAI REVISITED

Supriya Wilawan
Thammasat University

1. Introduction[1]

This study revisits the issue of double object constructions in Thai using a Lexicase dependency-grammar analysis. This issue goes back at least to Indrambarya 1994. There she proposes that in sentence (1), the NP 'dessert' has a Patient case relation, that is, that it is a direct object. The NP 'children', on the other hand, is assigned a Correspondent case relation; it is an indirect object.

(1) kháw hây   khánŏm dèkdèk
    he   give   dessert children
    'He gave the children some dessert.'

An essay arguing against these proposed case relations appeared in Wilawan 1996. Both these works determine the syntactic structures of double object constructions by identifying the similarities between each object and some elements found in other constructions. Both objects can be compared with the object of a transitive verb, the NP dependent on another NP, and the complement NP of an intransitive verb.

In this essay, I consider evidence from four different constructions that suggests that the second NP after verbs like 'give' in example (1) bears a Patient case relation. The first NP bears a new case relation, and therefore, a new type of case relation is proposed. Section 2 presents the syntactic characteristics of the ditransitive verbs in question and compares them with those found in other constructions. The results of the comparison provide evidence that a new case relation should be assigned to the NP immediately following the verb in instances like (1). Section 3 discusses some additional pieces of evidence that support this proposal, and considers its theoretical implications. Section 4 gives the conclusion.

2. Double Object NPs Compared with Four Other Constructions

Indrambarya (1994:190–194) proposes that the first NP after ditransitive verbs such as 'give' bear the Patient case relation, similar to the single object of simple transitive verbs, while the second NP bears a Correspondent case relation, as do the complement nouns of intransitive verbs. The syntactic structure of the double object constructions, according to this proposal, is shown in (2).

(2)

kháw	hây	khánŏm	dèkdèk
he	give	dessert	children
		+N	+N
		+Pat	+Cor

'He gave the dessert to the children.'
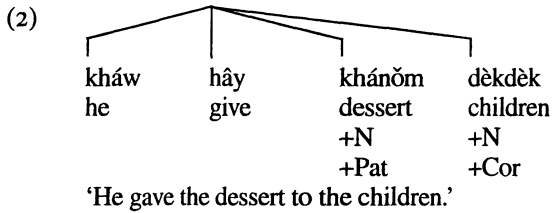
Indrambarya argues that only the object of transitive verbs can be topicalized, and because the NP 'dessert' can be topicalized, it has the same case relation as that carried by the objects of transitive verbs. It is thus a Patient noun.

Wilawan (1996) proposes a different structure for the verb 'give'. The syntactic structure of sentence (1) according to her proposal is as shown in (3).

(3)

kháw	hây	khánŏm	dèkdèk
he	give	dessert	children
		+N	+N
		+Cor	+Pat

'He gave the dessert to the children.'
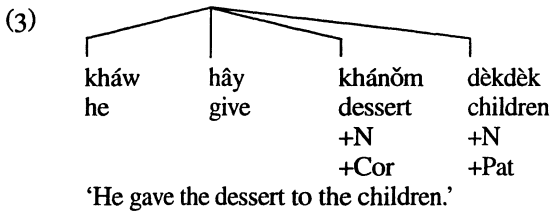
The evidence she gives is that the second NP after the verb 'give', *dèkdèk* 'children', has syntactic characteristics that are similar to those of the objects of transitive verbs.[2] Thus, it should be considered as having the same case relation, namely, Patient. On the other hand, the first object 'dessert' does not have all the syntactic properties that are similar to those found in the objects of simple transitive verbs. As a result, 'dessert' is assigned the Correspondent case relation, and 'children', the Patient relation (Wilawan 1996:43). However, the claims made about the first NP after ditransitive verbs by Wilawan and Indrambarya are open to some question. Why does the first NP of ditransitive verbs share only some syntactic characteristics of NPs following transitive verbs? What determines these properties?

In the following subsections, I reconsider the characteristics of the first and second NPs of ditransitive verbs such as 'give', as well as other heretofore unnoticed syntactic properties of nouns found in four other parallel constructions, namely, the object of transitive verbs, the complement noun of intransitive verbs, the incorporated NP in intransitive verbs, and the NP dependent of another NP.

2.1 Object of Transitive Verbs

The question to reconsider is whether the first NP or the second NP of ditransitive verbs should have the Patient case relation. That is, which of them has the syntactic characteristics most similar to those of the objects found in simple transitive sentences? The examples below indicate that the syntactic properties of the second NP are similar to those found in transitive verbs. Both nouns can be topicalized, clefted,

---

2. See 2.1 for a fuller description.

relativized, and quantified. On the other hand, the first NP of ditransitive verbs can only be topicalized and clefted; it cannot be relativized or quantified.

The syntactic characteristics of the transitive verb 'hit' are shown in sentences (4a–d).

(4)   kháw tii dèk
         he   hit child
         'He hit the child.'

         TOPICALIZATION
      a. dèk nán ná? kháw tii
         child that TOP he   hit
         'That child, he hit.'

         CLEFTING
      b. nân khɨ̀ɨ dèk   thîi kháw tii
         this is   child this he   hit
         'This is the child that he hit.'

         RELATIVIZATION
      c. kháw tii dèk   thîi nít chɔ̂ɔp
         he   hit child that Nit like
         'He hit the child that Nit liked.'

         MODIFICATION WITH QUANTIFIER
      d. kháw tii dèk$_i$ thúk khon$_i$
         he   hit child every CL
         'He hit every child (that is in this room).'

In sentence (4a), 'that child' is topicalized by adding the morpheme *ná?* and in (4b), 'this child' is clefted by adding the subject *nân* 'this' and an existential verb *khɨ̀ɨ* 'be'.

The syntactic properties of the first NP and the second NP of the ditransitive verb in example (1) are illustrated in sentences (5a) to (5h).

(5)   kháw hây khánǒm dèkdèk
         he   give dessert   children
         'He gave the dessert to the children.'

         TOPICALIZATION
      a. khánǒm nán ná? kháw hây dèkdèk
         dessert   that TOP he   give children
         'That dessert, he gave to the children.'

      b. dèkdèk nán ná? kháw hây khánǒm
         children that TOP he   give dessert
         'Those children, he gave the dessert to.'

         CLEFTING
      c. nân khɨ̀ɨ khánǒm thîi kháw hây dèkdèk
         this is   dessert that he   give children
         'This is the dessert that he gave to the children.'

    d. nân khɨɨ dèkdèk thîi kháw hây khánŏm
       this is   children that he   give dessert
       'These are the children that he gave the dessert to.'

    RELATIVIZATION
    e. *kháw hây khánŏm thîi nít chɔ̂ɔp dèkdèk
       he   give dessert that Nit like children
       'He gave the dessert that Nit liked to the children.'

    f. kháw hây khánŏm dèkdèk thîi nít chɔ̂ɔp
       he   give dessert children that Nit like
       'He gave the dessert to the children that Nit liked."

    MODIFICATION WITH QUANTIFIER[3]
    g. *kháw hây khánŏm thúk chín dèkdèk
       he   give dessert every CL   children
       'He gave every dessert to the children.'

    h. kháw hây khánŏm dèkdèk thúk khon
       he   give dessert children every CL
       'He gave the dessert to every child.'

These characteristics are summarized in table 1. From this table, it can be observed that it is the second NP of ditransitive verbs that should be assigned the Patient case relation, that is, the direct object, because it shares the syntactic characteristics of objects of transitive verbs. The fact that the first NP of ditransitive verbs does not manifest all these properties leads us to question the case relation it bears.

Table 1. Comparison of Transitive and Ditransitive Verbs

SYNTACTIC CHARACTERISTICS	OBJECT OF TRANSITIVE VERBS	DITRANSITIVE VERBS	
		NP 1	NP 2
TOPICALIZATION	yes	yes	yes
CLEFTING	yes	yes	yes
RELATIVIZATION	yes	no	yes
QUANTIFIER	yes	no	yes

---

3. In Thai, a quantifier can float, as in the following example.
    a. khruu thamthôot nákrian thúk khon mɨ̂akhɨɨnníi
       teacher punish   student every CL   last.night
       'The teacher punished every student last night.'

    b. khruu thamthôot nákrian mɨ̂akhɨɨnníi thúk khon
       teacher punish   student last.night   every CL
       'The teacher punished every student last night.'

Thus, the following sentence is acceptable.
    c. kháw hây khánŏm$_i$ dèkdèk thúk chín$_i$
       he   give dessert   children every CL
       'He gave every dessert to the children.'

2.2 Complement NP of Intransitive Verbs

The question is how the first NP 'dessert' in (1) is related to 'sugar', a complement noun of the intransitive verb *láʔlaai* 'dissolve', as exemplified in (6).[4]

(6)   kháw láʔlaai   náamtaan nán
       he   dissolve sugar   that
       'He dissolved that sugar.'

'Dissolve' is an intransitive verb that requires a complement noun as its dependent. In Lexicase analysis, 'sugar' is analyzed as having the Correspondent case relation. It is not a Patient noun, as is the object of a transitive verb. In Thai, a Correspondent noun such as this cannot be topicalized or clefted. However, it can be relativized and it can cooccur with a quantifier noun, as shown in the examples below.

(6)   TOPICALIZATION
    a. *náamtaan nán náʔ kháw láʔlaai
         sugar   that TOP he   dissolve
         'That sugar, he dissolved.'

       CLEFTING
    b. *nân khɨ̵ náamtaan thîi kháw láʔlaai
         this be   sugar   that he   dissolve
         'This was the sugar that he dissolved.'

       RELATIVIZATION
    c. kháw láʔlaai   náamtaan thîi nít sɨ̵
         he   dissolve sugar   that Nit buy
         'He dissolved the sugar that Nit bought.'

       MODIFICATION WITH QUANTIFIER
    d. kháw láʔlaai   náamtaan thúk kɔɔn
         he   dissolve sugar   every CL
         'He dissolved every sugar.'

The syntactic characteristics of the NP complement of intransitive verbs, the object of transitive verbs, and the first and second NPs of ditransitive verbs are summarized in table 2.

Table 2. Comparison of Intransitive, Transitive, and Ditransitive Verbs

SYNTACTIC CHARACTERISTICS	COMPLEMENT OF INTRANSITIVE VERBS	OBJECT OF TRANSITIVE VERBS	DITRANSITIVE VERBS	
			NP 1	NP 2
TOPICALIZATION	no	yes	yes	yes
CLEFTING	no	yes	yes	yes
RELATIVIZATION	yes	yes	no	yes
QUANTIFIER	yes	yes	no	yes

---

4. See Indrambarya (1994:146–148) for details of the analysis of intransitive verbs of this type.

It would seem that the first NP of the ditransitive verb 'give' does not fit the type of NP we have just considered. Could it be that the so-called ditransitive verb 'give' is not a verb that can take an NP as its dependent? Can the verb 'give' in (1), restated as (7), be an intransitive verb with the noun 'dessert' incorporated into it as a compound unit?

(7) kháw hây khánǒm dèkdèk
    he give dessert children
    'He gave dessert to the children.'

In 2.3, we consider the syntactic properties of compound verbs to see whether they share the same properties as those found in double object constructions.

2.3 Incorporated Intransitive Verbs

There are quite a number of incorporated intransitive verbs in Thai (Wilawan 1993:75). An example is 'fall work' in (8), shown together with its syntactic structure.

(8) kháw tòk ŋaan
    he   fall work
    'He lost his job.'

    kháw   tòk ŋaan
    he   fall work
            +V
            −trns (intransitive)
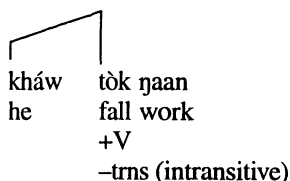
At first glance, the incoporated intransitive verb 'lose one's job' may look like a transitive verb with an object. That is, 'work' may look like the object of the verb 'fall' or a complement noun of the intransitive verb. However, its syntactic characteristics show that this noun does not behave like an ordinary single object of a transitive verb. Neither does it behave like a complement noun of an intransitive verb. It cannot be topicalized, clefted, relativized, or modified by a quantifier, as can be seen in (9).

(9)   TOPICALIZATION
  a. *ŋaan ná? kháw tòk
      work TOP he   fall
      'This job, he lost.'

      CLEFTING
  b. *nân khɨɨ ŋaan thîi kháw tòk
      this is   job that he   fall
      'This is the job that he lost.'

      RELATIVIZATION
  c. *kháw tòk ŋaan thîi nít hǎa háy
      he   fall work that Nit find for
      'He lost the job that Nit found for him.'

      MODIFICATION WITH QUANTIFIER
  d. *kháw tòk ŋaan thúk ŋaan
      he   fall job every CL
      'He lost every job.'

Thus, 'fall work' is considered a compound verb. Here the noun 'work' has syntactic properties that are different from the ordinary object NP of transitive verbs or the complement NP of intransitive verbs because it is incorporated into a verb, thus forming a single unit. This lexicalization explains why we cannot topicalize and cleft the noun after the intransitive verb 'fall', and it determines the fact that 'work' cannot occur with a relative clause or a quantifier. In addition, no adverb may be inserted between the verb 'fall' and ŋaan 'work', as in the unacceptable sentence in (10).[5] 'Work' behaves as an integral part of the verb.

(10) *kháw tòk thúkwan lǝǝi ŋaan
       he   fall   every.day EMPH work
      'He lost his job every day.'

Now if we consider the ditransitive verb 'give', we may wonder whether it is possible that sentence (1), restated as (11), has the same syntactic structure as that of sentence (8), that is, with the first noun after the verb incorporated into the verb, as in (11):

(11)

kháw   hây khánǒm   dèkdèk
he   give dessert   children
        +V   +N
        −trns
'He gave the children some dessert.'

However, if we consider the previous examples, especially sentences (5a) and (5c), which show that the first NP of the verb 'give' can be topicalized and clefted, we have to rule out the tree structure in (11). The syntactic characteristics of the first NP after the verb 'give' and the incorporated NP 'work' of (8) are different, as summarized in table 3.

Table 3. Four-way Comparison of Thai Verbs

SYNTACTIC CHARACTERISTICS	INTRANSITIVE INCORPORATED NP	INTRANSITIVE COMPLEMENT	TRANSITIVE OBJECT	DITRANSITIVE NP 1	NP 2
TOPICALIZATION	no	no	yes	yes	yes
CLEFTING	no	no	yes	yes	yes
RELATIVIZATION	no	yes	yes	no	yes
QUANTIFIER	no	yes	yes	no	yes

5. In Thai, adverbs of time cannot be placed between a transitive verb and its object NP.

    a. kháw kɔ̀ɔt lûukchaai thúkwan   lǝǝi
        he   hug son   every.day EMPH
        'He hugged his son every day.'

    b. *kháw kɔ̀ɔt thúkwan   lǝǝi   lûukchaai
        he   hug every.day EMPH son
        'He hugged his son every day.'

215

Having ruled out the two possible occurrences of an NP with two types of intransitive verbs, what other explanation can be offered for the first NP 'dessert' of the verb 'give' in (1) and (5)? In the following subsection, we test another possibility.

2.4 NP Dependent on Another NP

Another possibility is that the one NP is dependent on the other NP. An example of this is shown in sentence (12) below.

> (12) nít chɔ̂ɔp rót   phûuchaai khon nán
>   Nit like   car (NP1) man (NP2) CL   that
>   'Nit liked that man's car.'

The NP 'man's car' has the following syntactic structure:

> (12) a.
>
> | rót | phûuchaai |
> |---|---|
> | car | man |
> | +N | +N |
> | 'man's car' | |

The NP 'man' is a dependent of the NP 'car'. As a result, the NP 'man's car' behaves as one unit, as can be observed in sentences (13a) to (13c).

> (13) TOPICALIZATION
>
> a. rót phûuchaai khon nán ná? nít chɔ̂ɔp
>   car man   CL   that TOP Nit like
>   'That man's car, Nit liked.'
>
> b. *rót nán ná? nít chɔ̂ɔp phûuchaai khon nán
>   car that TOP Nit like   man   CL   that
>   *'That car, Nit liked that man's.'
>
> c. *phûuchaai khon nán ná? nít chɔ̂ɔp rót
>   man   CL   that TOP Nit like   car
>   *'That man, Nit liked the car of.'

Since 'man's car' behaves as a single unit, we expect that clefting each NP creates some unacceptable sentences, and this is what we find in the sentences below.

> (14) CLEFTING
>
> a. *nân khɨɨ rót thîi nít chɔ̂ɔp phûuchaai khon nán
>   this is   car that Nit like   man   CL   that
>   *'This is the car that Nit liked the man's.'
>
> b. *nân khɨɨ phûuchaai khon nán thîi nít chɔ̂ɔp rót
>   this is   man   CL   that that Nit like   car
>   *'This is the man that Nit liked the car.'

Only the second NP 'man' can be modified by a relative clause.

216

(15) RELATIVIZATION

    a. *nít chɔ̂ɔp rót$_i$ khan$_i$ thîi nát sɨ́i phûuchaai
       Nit like   car CL   that Nat buy man
       *'Nit liked the car of the man that Nat bought.'

    b. *nít chɔ̂ɔp rót$_i$ phûuchaai khan$_i$ thîi nát sɨ́i
       Nit like   car man   CL   that Nat buy
       *'Nit liked the man's car that Nat bought.'

    c. nít chɔ̂ɔp rót phûuchaai$_i$ khon$_i$ thîi lɔ̀ɔ
      Nit like   car man   CL   that handsome
      'Nit liked the car of the man who was handsome.'

Since the NP 'man's car' is a unit, we cannot put a quantifier between these two NPs.

(16) MODIFICATION WITH A QUANTIFIER

    a. *nít chɔ̂ɔp rót thúk khan phûuchaai khon nán
       Nit like   car every CL   man   CL   that
       *'Nit liked every car of that man.'[6]

    b. nít chɔ̂ɔp rót phûuchaai thúk khon
      Nit like   car man   every CL
      'Nit liked the car of every man.'

After considering the syntactic characteristics of an NP depending on another NP that are summarized in table 4, we have to rule out the possibility that 'dessert', the first NP of the ditransitive verb in (1), restated here as (17), has 'children' as its dependent.

(17) kháw hây khánŏm dèkdèk
    he   give dessert children
    'He gave dessert to the children.'

Table 4. Five-way Comparison of Thai Verbs

SYNTACTIC CHARACTERISTICS	NP DEP. ON NP		INTRANS. INCORP. NP	INTRANS. COMPL.	TRANS. OBJECT	DITRANSITIVE	
	NP 1	NP 2				NP 1	NP 2
TOPICALIZATION	no	no	no	no	yes	yes	yes
CLEFTING	no	no	no	no	yes	yes	yes
RELATIVIZATION	no	yes	no	yes	yes	no	yes
QUANTIFIER	no	yes	no	yes	yes	no	yes

2.5 Conclusions from the Comparison

After observing the incongruities between both nouns in double object constructions and those nouns in four different constructions summarized in table 4, we are left

---

6. Since a quantifier can float in Thai, the following sentence is acceptable.

    a. kháw chɔ̂ɔp rót$_i$ phûuchaai$_j$ khon$_j$ nán thúk khan$_i$
      he   like   car man   CL   that every CL
      'He liked every car of that man.'

with nothing but the basic claim that both nouns after ditransitive verbs are objects. I propose then that there are two types of Patient case relation in double object constructions in Thai, namely, Patient 1, which is a direct object, and Patient 2, which is an indirect object.

(18) kháw cèɛk   ŋaan lûuknɔ́ɔŋ
     he   distribute work subordinate
                   Pat2 Pat1
     'He distributed his work to his subordinates.'

Thus, 'work' is an indirect object and 'subordinate' is a direct object.

2.5.1 Patient 1. Unlike Indrambarya's analysis (1994), which claims that the second NP of ditransitive verbs is an indirect object having the Correspondent case relation, I propose that it should be the direct object having the first type of Patient case relation, represented as Patient 1. This NP shares four syntactic properties of the object of transitive verbs as shown in previous tables. Thus, in (19a), the NP 'subordinate' is assigned the Patient 1 case relation.

(19) a. kháw cèɛk   ŋaan lûuknɔ́ɔŋ   thúk khon
        he   distribute work subordinate every CL
        'He distributed his work to every subordinate.'

Besides four syntactic characteristics that the object of transitive verbs and the second NP of ditransitive verbs share, we find that the second NP of ditransitive verbs passivizes as easily as the direct object of transitive verbs, as can be seen below.

(19) b. lûuknɔ́ɔŋ   thúk khon doon   kháw cèɛk   ŋaan
        subordinate every CL   undergo he   distribute work
        'Every subordinate was assigned work by him.'

On the other hand, it is not easy to passivize the first NP 'work', which is analyzed as an indirect object and is assigned the Patient 2 case relation in the present study.

(19) c. *ŋaan doon   kháw cèɛk   lûuknɔ́ɔŋ   thúk khon
        work undergo he   distribute subordinate every CL
        *'That work was distributed to every subordinate.'

The fact that the first NP cannot passivize does not indicate an absolute argument against treating it differently from the second NP, due to the fact that not all objects of transitive verbs themselves passivize in Thai.[7] However, (19b) shows that the second NP, Patient 1, does behave in the same manner as the object NP of transitive verbs in terms of passivization.

2.5.2 Patient 2. Contrary to Indrambarya's analysis of double object constructions, which proposes that the first NP is a direct object bearing a Patient case relation, and unlike the standard Lexicase analysis of double object constructions, which assigns

---

7. See Wilawan (1996:49) for a fuller description.

the Correspondent case relation (the case relation of the NP complement of intransitive verbs) to an indirect object,[8] I propose that the first NP is the indirect object and is assigned the Patient 2 case relation. There are two reasons to support this analysis. First, the syntactic properties of the first NP are different from those of the object of transitive verbs. The object of transitive verbs can be relativized and quantified, but the first NP of ditransitive verbs cannot, as seen in sentences (4c) and (4d) and sentences (5e) and (5g). Thus, it should not be analyzed as a direct object but as an indirect object. Second, I propose that the indirect object of ditransitive verbs in Thai should not bear a Correspondent case relation. The reason for this lies in the fact that this NP does not behave syntactically like a Correspondent noun of intransitive verbs, as seen from sentences (5a, c, e, g) and sentences (6a–d).

In section 3, I consider the other syntactic evidence that determines the assignment of these two types of Patient case relation and the theoretical consequence of my analysis.

3. The Patient Nouns of Double Object Constructions

By considering their syntactic behavior, we will see how the two Patient nouns of ditransitive verbs are related, and as two types of Patients, how they are different. Following this is a discussion of the theoretical consequence of this analysis.

3.1 Evidence for Two Patients

To determine the reasonableness of the proposal that both direct and indirect object nouns of ditransitive verbs in Thai bear the same Patient case relation rather than that of a Patient and a Correspondent, we need to see whether they share the syntactic characteristics of Patient nouns. After considering their syntactic characteristics in sentences (4a–d) and (5a–h), we find that the direct and indirect objects of ditransitive verbs and the objects of transitive verbs have two properties in common. They can all be topicalized and clefted. However, having two different types of Patient nouns, as we do in this analysis, the first NP and the second NP of ditransitive verbs, there are two properties that distinguish the one from the other.

First, the second NP of ditransitive verbs, which is assigned the Patient 1 case relation in this analysis, has one additional property that it shares with the Patient noun of transitive verbs, but not with the first NP of ditransitive verbs. The former two NPs can provide a referent for an extracted object in relation to infinitival adjuncts, something that the first NP of ditransitive verbs cannot do.

> (20) TRANSITIVE VERB : 'help'
>   kháw chûai chaawkhǎw$_i$ phîa hâi Ø$_i$ mii   thǎanádii
>   he   help hilltribes   for to   have wealth
>   'He helped the hilltribes so that they would be wealthy.'

---

8. In English, an indirect object, 'John' in the sentence below, is assigned a Patient case relation, while a direct object, 'a book', a Correspondent (Starosta 1988:130).

   a. Anne gave John a book.

(21)   DITRANSITIVE VERB: 'teach'
a. kháw sɔ̌ɔn sàatsanǎa chaawkhǎwᵢ phɨ̀a hâi Ø̧ᵢ láʔ?   kilèt
   he   teach religion hilltribes   for to   relinquish defilement
   'He taught religion to the hilltribes so that they would relinquish
   defilement.'

b. *kháw sɔ̌ɔn sàatsanǎaᵢ chaawkhǎw phɨ̀a hâi Ø̧ᵢ phrêelǎai nai chonnabòt
   he   teach religion hilltribes   for to   spread in rural.area
   *'He taught the hilltribes religion so that it would spread in the rural area.'

In (20), 'hilltribes', which is the object of 'help' and bears the Patient case relation, can be interpreted as being the subject of 'have'. This is also true for sentence (21a), where 'hilltribes', the direct object bearing the Patient 1 case relation, can also be interpreted as being the subject of the verb 'relinquish'. But the indirect object 'religion', bearing the Patient 2 case relation, cannot be referred to as the subject of the infinitival verb in the adjunct clause in (21b). These sentences indicate that the two Patient case relations are different. Patient 1 manifests a syntactic behavior similar to the Patient object of transitive verbs, whereas Patient 2 does not.

The inability of the first NP of ditransitive verbs to take a relative clause or a quantifier as modifiers is the second syntactic characteristic that indicates that the first and second NPs of ditransitive verbs in Thai are different. As we have seen in sentences (5e) to (5h), and as summarized in the tables above, only the second NP can occur with a relative clause or a quantifier. Why are sentences (22) and (23) unacceptable?

(22) *kháw cɛ̀ɛk   khúkkîi thúk chín dèkdèk
       he   distribute cookies every CL   children
      *'He distributed evey cookie to the children.'

(23) *kháw cɛ̀ɛk   khúkkîi thîi nít chɔ̂ɔp dèkdèk
       he   distribute cookies that Nit like   children
      *'He distributed the cookies that Nit liked to the children.'

Given the modifications in the first NP in (22) and (23), the main reason for their ungrammaticality lies in the fact that the direct object, Patient 1, which behaves similarly to the Patient NP of transitive verbs, is shunted farther from the head verb. There is a tendency for a Patient noun to be close to the head verb (Starosta, pers. comm.). Let us look at some additional sentences that involve adverbs of time and the right-position topicalization marker *nà?* in order to make more obvious the constraint that the direct object noun needs to be close to the head verb.

Sentence (24) below contains an adverb of time occurring between the head verb and the object of the transitive verb. The sentence is unacceptable.

(24) *nít cɛ̀ɛk   thúk khɨɨn lǝǝyná? khúkkîi dèkdèk
       Nit distribute every night EMPH   cookies children
      *'Nit distributed the children cookies every night.'

In addition, the basic order of double object constructions is fixed. Pat2 has to precede Pat1, as can be seen in (25).

220

(25) *kháw cὲεk   dèkdèk khúkkiî thúk khɨɨn lɔɔyná?
     he   distribute children cookies every night EMPH
     *'He distributed the children cookies every night.'

Now consider the sentences below, which involve the right-position topicalization marker and an adverb of time.

(26) *nít cὲεk   khúkkiî thúk khɨɨn lɔɔyná?   dèkdèk nà?
     Nit distribute cookies every night EMPH   children TOP
     *'Nit distributed cookies to the children every night.'

Sentence (26) indicates that when Pat1, 'children', is topicalized by a right-position topicalization marker, it is also far from the head. There are two elements, namely, the Pat2 and the adverb of time, interrupting the head verb and the direct object; hence, sentence (26) is unacceptable. However, when the Pat2 is topicalized by a right-position topicalization marker, as in (27) below, the construction is acceptable. The constraint on the order of the Patient nouns in the double object constructions is relaxed when Pat2 is topicalized.

(27) nít cὲεk   dèkdèk thúk khɨɨn lɔɔyná?   khúkkiî nà?
     Nit distribute children every night EMPH   cookies TOP
     'Nit gave the children some cookies every night.'

This sentence is acceptable because there is nothing between the head verb and the noun 'children', Pat1.

If we return to the unacceptable sentences (22) and (23), which involve quantifier and relative clause modifications, we see that they become acceptable when the right-position topicalization marker is applied and there is nothing between the head verb 'distribute' and the direct object 'children'.

(28) nít cὲεk   dèkdèk thúk khɨɨn lɔɔyná? khúkkiî thúk chín nà?
     Nit distribute children every night EMPH   cookies every CL   TOP
     'Nit distributed every cookie to the children every night.'[9]

(29) nít cὲεk   dèkdèk thúk khɨɨn lɔɔyná? khúkkiî thîi nít chɔɔp nà?
     Nit distribute children every night EMPH   cookies that Nit like   TOP
     'Nit distributed cookies that Nit liked to the children every night.'

On the other hand, when the direct object 'children' is not close to the head verb because of the presence of the adverb of time and the right-position topicalization marker, the sentences are unacceptable.

(30) *nít cὲεk   khúkkiî thúk chín thúk khɨɨn lɔɔyná? dèkdèk nà?
     Nit distribute cookies every CL   every night EMPH   children TOP
     'Nit distributed every cookie to the children every night.'

(31) *nít cὲεk   khúkkiî thîi nít chɔɔp thúk khɨɨn lɔɔyná? dèkdèk nà?
     Nit distribute cookies that Nit like   every night EMPH   children TOP
     'Nit distributed cookies that Nit liked to the children every night.'

---

9. There are some speakers who do not like this sentence, but all of them agree that sentence (29) is acceptable. A fuller description of the quantifier 'every' in Thai is needed.

3.2 Theoretical Consequence: One per Sent

One of the several assumptions in Lexicase dependency grammar is the "one per sent" constraint. It states that no case relation may occur more than once in any given clause. At first glance, proposing two Patient nouns for the double object constructions seems to violate this constraint. However, two similar case relations are allowed if both case relation nouns refer to endpoints of the path traveled by the Patient, as we see in the English sentence below.

(32) John carried the dog from the barn to the mailbox.
                      Pat   Loc   Loc

Both 'barn' and 'mailbox' refer to the endpoints of the path traveled by the Patient 'dog' (Starosta 1988:138). Proposing two Patient nouns for the double object constructions in Thai does not seem to violate this constraint if it is relaxed in parallel fashion. Consider sentence (19a), restated below as (33) .

(33) kháw cèek   ŋaan lûuknɔ́ɔŋ   thúk khon
      he   distribute work subordinate every CL
                      Pat2 Pat1
      'He distributed his work to every subordinate.'

We can consider the endpoint of 'work' to be 'subordinate'. In other words, the Patient noun 'work' is transferred to the other Patient noun 'subordinate'. It should be pointed out that this generalization excludes certain verbs that look like ditransitive verbs, such as yɨɨm 'borrow', sɨ́ 'buy', and *khamoy* 'steal', for two reasons. First, the syntactic characteristics of the NP objects of these verbs are different from those analyzed in this study. We cannot topicalize the second NP after these verbs.[10] Second, we cannot interpret both nouns after these verbs as Patient nouns. They do not refer to endpoints of the path the Patient travels. In (34) below, 'book' is transferred from 'friend' to the Agent noun 'he', not from the Agent to the noun 'friend'.

(34) kháw yɨɨm   nǎŋsɨ̌ phîan
      he   borrow book   friend
      'He borrowed a book from his friend.'

With these look-alike verbs excluded, and with sufficient relaxation of the "one per sent" constraint, this analysis adds to our understanding of ditransitive verbs in relation to other Thai verbs.

---

10. Consider the look-alike ditransitive verb 'to borrow'. When the second NP after this verb is topicalized, the construction is unacceptable, as in (b).

    a. kháw yɨɨm   nǎŋsɨ̌ phîan
       he   borrow book   friend
       'He borowed a book from his friend.'

    b. *phîan khon nán ná? kháw yɨɨm   nǎŋsɨ̌
       friend CL that TOP he   borrow book
       *'That friend, he borrowed a book.'

4. Conclusion

In this essay, I argue that two classes of the Patient case relation are needed to account for the NP objects of ditransitive verbs such as 'give' and 'distribute'. The evidence lies in the fact that the first NP next to these verbs has syntactic properties that are different from those found in four other relevant constructions. By proposing a new type of case relation for this NP, the puzzling syntactic characteristics of this type of double object construction are clarified.

REFERENCES

Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure.* Chicago: University of Chicago Press.

→ Hudson, Richard. 1992. So-called double objects and grammatical relations. *Language* 68:251–276.

Indrambarya, Kitima. 1994. Subcategorization of verbs in Thai: A lexicase dependency approach. Ph.D. dissertation, University of Hawai'i.

Starosta, Stanley. 1988. *The case for lexicase: An outline of lexicase grammatical theory.* London: Pinter Publishers.

Wilawan, Supriya. 1993. A reanalysis of so-called serial verb constructions in Thai, Khmer, Mandarin Chinese, and Yoruba. Ph.D. dissertation, University of Hawai'i.

———. 1996. Double object constructions in Thai (in Thai). *Journal of Language and Linguistics* 15:43–51.

14: SOME REMARKS ON THE GRAMMATICAL FUNCTIONS OF THE NONABSOLUTIVE AGENT IN TAGALOG

Videa P. De Guzman
The University of Calgary

1. Introduction[1]

Much has been published on the topic of "subject" and "object" in Tagalog and in the other Philippine languages. However, the conclusions reached in these studies vary and range from one extreme to the other. Some conclude that there are no subjects, (Schachter 1976; Cena 1995) nor are there (direct) objects (Gil 1984; Naylor 1980); others maintain the relevance of one or more grammatical relations or functions (Hoekstra 1986; Byma 1986; Gerdts 1988; De Guzman 1988; 1995; Shibatani 1988; Kroeger 1993; Brainard 1996). The approaches adopted in these studies are, of course, different from each other, and consequently, they cannot be compared strictly using the same set of criteria. Could the disparity in the conclusions lie on the differences in defining these grammatical terms (e.g., structural/configurational versus nonstructural/nonconfigurational) or is it the attempt to universalize the application of each grammatical function that brings us to this dilemma?

Most of the previous studies have concentrated on the grammatical functioning of the ang-phrase, argued in some as the subject of Tagalog structures. In this essay, my purpose is to show that the two nominal phrases—semantically, the AGENT/EXPERIENCER and the PATIENT/OBJECT/THEME—in Tagalog basic verbal structures, when marked identically with the particle ng, are different in their grammatical functions.[2] In Naylor 1980:37, both of these ng-phrases, and likewise the ng-instrument phrase, are considered as verbal attributes that share the same attributive relation as the possessive and the adverbial structures, also marked by the particle ng.[3] Here, I will restrict my arguments only to the distinctions that need to be made between the ng-agent and the ng-patient. It will be shown that the distinctions are syntactically relevant. In the exposition, it is deemed simpler and less confusing if the two nominal phrases to be compared are referred to according to the markers they take. Each nominal phrase when marked by the particle ng corresponds to a different grammatical relation. (I will refer to their corresponding grammatical relations based on the approach adopted to emphasize their distinctions.)

It will be argued that the ng-agent compared with the ng-patient actively performs certain grammatical operations that are exclusively its own, that is, even regardless of

---

1. An earlier version of this essay was read at the Eighth International Conference on Austronesian Linguistics, Taipei, December 1997.

2. The marker ng is pronounced [naŋ].

3. Popular orthography uses nang instead of ng in adverbial phrases.

particle marking.[4] In some cases, however, these functions are shared with the recipient or benefactive phrase, traditionally labeled indirect object. Furthermore, there are other grammatical processes that the *ng*-agent may share with the *ang*-phrase. These pieces of evidence all point to the grammatical significance of the *ng*-agent phrase in contrast with the failure of the *ng*-patient to operate under similar conditions. This difference leads to the inevitable conclusion that these two identically-marked nominal phrases must be grammatically distinguished.

To account for the differences, the presentation will employ the ergative analysis based on a modified version of the Relational Grammar (RG) framework (See De Guzman 1988 and 1995). The distinctions between the primary grammatical relations referred to as term 2 (or "direct object"; the *ang*-PATIENT nominal in Tagalog basic verbal clauses), term 1 (or "subject"; the *ng*-AGENT nominal), and the nonterm 2-chomeur ("unemployed," the *ng*-PATIENT nominal in clauses where an advancement to term 2 applied) may be delineated. First, the semantic and syntactic differences with respect to particle marking and the substitute forms of the *ng*-agent and the *ng*-patient nominals will be shown. Second, I will present the markers of phrases corresponding to these two in the other major Philippine languages to provide further evidence that these phrases are not identical. Then, I will take a closer look at two specific structures—a special control structure and simple conjoined structures for pronominal interpretation—to show further the operativeness of the *ng*-agent, almost competing with the *ang*-phrase, and the clear inability of the *ng*-patient to participate in any of these grammatical processes. How can this difference be accounted for? The RG account is then compared with the Lexicase Grammar account of the *ng*-patient nominal. It will be proposed that the alignment suggested in RG whereby the semantic relation agent/experiencer is the "Subject" and the semantic relation patient/object/theme (or goal in other works) is the "Direct Object" (following the classical tradition) may better unify and may cause less confusion in the description of these grammatical relations across languages.

2. *Ng*-agent versus *Ng*-patient in RG

Following the ergative analysis (Payne 1982; Gerdts 1988) in a modified version of Relational Grammar (De Guzman, 1988), the patient-phrase of a given verb is aligned initially with term 2 (Direct Object), higher in the hierarchy of grammatical relations (GRs) than term 1 (Subject), which is aligned with the agent-phrase. Limiting our analysis to these two primary GRs and touching tangentially on the Oblique (for recipient, location, beneficiary, etc.), we follow the relational hierarchy: 2 > 1 > Obl. In the accusative analysis, the hierarchy of grammatical relations is: 1 > 2 > 3 > Obl (Bell 1983; Gerdts 1988). Note that in both accusative and ergative type languages, the same semantic relations are aligned with the same corresponding terms or GRs. The difference is in the choice of which GR is more central or salient in a given language. Briefly, the term relations may be illustrated in Tagalog as follows with corresponding sketches of networks of relations:

---

4. Schachter 1977 refers to these as role-related functions in contrast to reference-related ones.

CHAPTER 14

(1) ibinigay ng mestra ang lapis sa bata   V   A   P   R[5]
gave-OV 1 teacher 2   pencil Obl child Pred 1   2   Obl
'The teacher gave the pencil to the child.'   bigay mestra lapis bata

Particle-marking assigns *ng* before common nouns in Final 1, *ang* in Final 2, and *sa* in Final Obl. Where a rule of advancement to term 2 applies, in which a term lower in the hierarchy takes over the primary relation 2, this initial term 2 becomes "unemployed" and is called a 2-chomeur, a nonterm. For example, when term 1 in (1) advances to term 2, the resulting structure shown in (2) below is what has been referred to as an ANTIPASSIVE:

(2) nagbigay ang mestra ng lapis sa bata   V   A P   R
gave-AV 2   2-CHO Obl   Pred 1 2   Obl
'The teacher gave a pencil to the child.'   Pred 2 2-cho Obl(1>2Adn)

Note that in (1), the final GRs corresponding to the semantic roles required by the verb 'give' are terms 1, 2, and Obl, a transitive clause. In (2), the only final term relations cooccurring are 2 and Obl, that is, without a 1. Without a cooccurring term 1, the clause structure with only a term 2 is intransitive. Compare the following basic intransitive clause:

(3) bumagsak ang bata/niyog   V   P
fell-AV   2   child/coconut   Pred 2
'The child/coconut fell.'

On the other hand, a different type of transitive clause results when an Oblique term advances to 2. As in sentence (2) above, the initial 2 becomes a 2-chomeur (CHO), while term 1 remains unchanged. Note that with each structure undergoing an advancement, depending on which semantic relation assumes the primary grammatical relation with the verb, a corresponding change in the voice affix on the verb is registered, as for example *m(pag-)* for agent; *-an* for recipient, location; *i-* for instrument, benefactive; and so forth.

(4) binigyan ng mestra ng lapis ang bata   Pred 1 2   Obl
gave-BV 1   2-CHO 2
'The teacher gave the child a pencil.'   Pred 1 2-CHO 2(Obl>2Adn)

Thus, in terms of GRs, the *ng*-agent that occurs in a transitive clause is a FINAL TERM 1 (Subject), whereas the *ng*-patient that occurs in an antipassive or in a nonbasic tansitive clause, that is, with an Oblique to 2 advancement, is a FINAL 2-CHOMEUR, a nonterm. At this point, it may be emphasized that the GRs being compared here are those manifested by *ng*-agents and *ng*-patients.

---

5. The following abbreviations are used: A, agent; ABS, absolutive; Adn, advancement; AV, active voice; BV, benefactive voice; CHO, chomeur; com, common noun; COMP, complement; def, definite; DemPro, demonstrative pronoun; ERG, ergative; EV, experiencer voice; GEN, genitive; LKR, linker; P, patient; pers, personal; PersPro, personal pronoun; Pred, predicate; OBL, Oblique; OV, objective voice; R, recipient; SPCF, specific; V, verb.

2.1 Formal Differences and Similarities

First, let us look at the various forms that the *ng*-agent and the *ng*-patient take as term
1 and as 2-chomeur, respectively. The (a) sentences that follow show the Agent as a
final term 1, while the (b) sentences show the patient as a final 2-chomeur.

(5) a. tinanong   ni   Pedro   si   Pablo
      asked-OV   1, ERG   2, ABS
                   [+pers name]   [+pers name]
      'Pedro asked Pablo.'

    b. nagtanong si   Pedro   kay/*ni   Pablo
       asked-AV   2, ABS   2-CHO, OBL
                  [+pers name]   [+pers name]
       'Pedro asked Pablo.'

(6) a. tinanong   ng   bata   ang   kapit-bahay
      asked-OV   1, ERGchild 2, ABS   neighbor
                [+com]   [+com]
      'The child asked the neighbor.'

    b. nagtanong ang bata   sa/*ng   kapit-bahay
       asked-AV   2, ABSchild   2-CHO, OBL neighbor
                [+com]   [+com], [+spcf]
       'The child asked the neighbor.'

(7) a. tinanong   niya   siya
      asked-OV   1, ERG   2, ABS
                 [+PersPro]   [+PersPro]
      'He/She asked her/him.'

    b. nagtanong siya   sa kaniya/*niya
       asked-AV   2, ABS   2-CHO, OBL
                [+PersPro]   [+PersPro]
       'She/He asked (of/from) him/her.'

In (5a), where the nominals are personal names, term 1 takes the ergative *ni*, whereas
term 2 is marked with the absolutive *si*. Note that in the antipassive structure (5b), the
2-CHO personal name takes the oblique marker *kay*; using the ergative or genitive
marker *ni* is ungrammatical. Being personal names, the agent and the patient nominals
are both definite or specific. One observable difference then is that personal names
bearing term 1 and 2-CHO are marked differently, the former takes the ergative particle,
whereas the latter, the oblique. Compared with (6a) and (6b), where both nominals are
common nouns, but still referring to specific people or animate beings, we see a differ-
ence once again in the particles that mark the agent (term 1) with the ergative *ng*, and
the patient (2-CHO), with the oblique *sa*. Correspondingly, when personal pronouns are
used in the same relations, the agent takes the ergative form but the patient takes the
oblique. When the patient is nonhuman or an object, which is typical of a patient/
theme/goal semantic role, the scenario becomes still different. To illustrate:

227

(8) a. babasahin   ng   bata   ang   libro
     will.read-OV   1, ERGchild   2, ABS   book
     'The child will read the book.'

   b. magbabasa   ang   bata   ng   libro
     will.read-AV   2, ABSchild   2-CHO, OBL   book
                                     [−def/−spcf]
     'The child will read a book.'

   c. magbabasa   ang   bata   sa   libro
     will.read-AV   2, ABSchild   2-CHO, OBL   book
                                     [+spcf]
     'The child will read of/from the book.'

The very common particle that marks this type of 2-CHO patient is *ng*, and it connotes nondefiniteness or nonspecificity. However, there are verbs that allow the use of the oblique particle *sa*, instead of *ng*, to refer to a specific item from a class.[6] Although this alternation may be allowed in the case of the 2-CHO patient, it is prohibited for the term 1 agent. In contrast, the usual interpretation of the *ng*-agent is specific or definite, except when modified by a qualifying indefinite word such as *isa* 'one' or *ilan* 'some'. Furthermore, when a demonstrative pronoun is used in place of the agent and the patient in question, there is confirmation of the similarity and difference observed previously, as shown in the following:

(9) a. babasahin   nito   ito
     will.read-OV   1, ERG   2, ABS
                   [+DemPro] [+DemPro]
     'This (one here) will read this (thing here).'

   b. magbabasa   ito   nito
     will.read-AV   2, ABS   2-CHO, OBL
                   [+DemPro] [+DemPro]
     'This (one here) will read this (thing here).'

   c. magbabasa   ito   dito
     will.read-AV   2, ABS   2-CHO, OBL
                   [+DemPro] [+Dem Pro]
     (i) 'This (one here) will read of/from this.'[7]
     (ii) 'This will read here.'

Being demonstrative pronouns, the referents must be specific. It is interesting to note that both the agent and the patient take the identical form *nito*. The difference, however, similar to that in the preceding example, is that the 2-CHO *nito* may be replaced by the OBL *dito* to mean 'of/from this'. Thus far, we have shown that the term 1 agent is invariably marked by the ergative particle, whereas the 2-CHO patient reveals a less uniform

---

6. It may be mentioned that there are verbs in the active voice that mark their patient only with *sa*, e.g., *tumulong* 'help'.

7. The same demonstrative pronoun is usually used as a locative OBL with the meaning as given in (9c)(ii).

marking, leaning more towards the oblique forms. It may be suggested that since the 2-CHO is a nonterm relation, the marker *ng* for the corresponding nominal must be identified as oblique with the added feature [-spcf], an alternant of the oblique *sa*-form. The term I relation, on the other hand, is notably ergative in all its forms. In sum, the proposed account is that since the two *ng*-phrases differ in the range of forms that can occur in their respective grammatical positions, these two nominals must be distinguished. I agree with Naylor (1980) that on the surface level there is no apparent distinction that can be made between them. That is, given a phrase *ng* + N, we cannot determine its function. (Similarly in English, given Det + N, the structural context must be known to be able to determine its specific function.) Likewise, I share the notion that it is the meaning of the verb that requires the semantic role agent and patient to be specified, but I maintain that these two roles are assigned their respective GRs accordingly and are marked with particles corresponding to their designated GRs.

2.2 Corresponding *Ng*-phrases in other Philippine Languages

Table 1 gives a bird's eye view of the types of markers in the other major languages of the Philippines that correspond to the forms of the *ng*-agent and the *ng*-patient in Tagalog. A cursory inspection of the markers for Final 2-CHO shows that even when one form may be identical to the marker for Final I, another marker will distinguish it from the corresponding marker in Final I. More often than not, one or two forms marking Final 2-CHO are identical to the Oblique markers of Final Oblique. This identity is indicated by the forms in boldface in the last two columns. It is interesting that Ilokano stands alone in marking the two primary relations, term I and term 2,

Table 1. Common Noun/Personal Name Markers of the Various Grammatical Relations in the Eight Major Languages of the Philippines

LANGUAGE	FINAL 2 ABS	FINAL I ERG	FINAL 2-CHO OBL	FINAL OBL OBL
TAGALOG	ang/si	ng/ni	ng [–spcf] sa/kay [+spcf]	**sa/kay**
BIKOL	an/si	kan/ni	ning [–spcf] kan/**ki** [+spcf]	sa/**ki** or kay
KAPAMPANGAN	ing/i	ning/(na)ng	-ng	king/kang
HILIGAYNON	ang/si	sang/ni	sang/**kay**	sa/**kay**
ILOKANO	ti or diay/ni	ti or diay/ni	iti* [–spcf] **idiay/kenni** [+spcf]	**idiay/kenni**
PANGASINAN	su/si	na/nɨn	na/**id kinɨn**	**id/id kinɨn**
SEBUANO	ang/si	sa/ni	ug [–spcf] **sa/kang** [+spcf]	**sa/kang**
WARAY	an/hi	han/ni	hin [–spcf] **ha/kan** [+spcf]	**ha/kan**

*   Many times in ordinary speech, even /iti/ is contracted to /ti/.

identically. But as soon as personal pronouns are used, the Absolutive form can only be used with Final 2, the Ergative for Final 1, and the Oblique for both Final Oblique and Final 2-CHO. Waray is the only one that manifests the four-particle distinctions before common nouns marking the four different GRs. Note, however, the identity between the alternant forms of Final 2-CHO and those of Final Obl.

From the above, it appears reasonable to align the nonterm 2-CHO with the Oblique term, similar to that of Final Obl, since these two relations are nonprimary. Notice the strong tendency of Final 1 to be distinctively marked from Final Obl, which is quite the opposite of the marking of Final 2-CHO.

3. Grammatical Functions of the Agent, Term 1

In previous works on Tagalog and the other Philippine languages, it has been shown that syntactic properties or processes commonly associated with the "subject" (the nominative phrase usually the agent in accusative languages) divide between the *ang*-phrase (the absolutive phrase or Final 2 in the present analysis) and the "nonsubject" or nonabsolutive *ng*-agent phrase (Final 1). These processes drawn from the different studies are listed below for easy reference and recall (Schachter 1976, 1977; O'Grady 1987; Gerdts 1988; De Guzman 1988; Kroeger 1993):

A. *Ang*-phrase		B. *Ng*-Agent phrase	
Relativization		Reflexivization	
Question formation		Control or Equi-target*	(*ang*-phrase)
Quantifier floating		Imperative addressee	
Clefting		Relevance to word order	
Topicalization*	(*ng*-agent)	Pronoun interpretation*	(*ang*-phrase)
Raising*	(*ng*-agent)	Quantifier scope	
Possessor ascension			
Conjunction reduction clause			

While the processes listed in Column A are unequivocally attributable to the *ang*-phrase, there are specifiable conditions where the *ng*-agent phrase may also participate in some of them. These two processes, which will be considered below, are Topicalization and Raising, both marked with an asterisk. In Column B, the processes involve the agent-phrase, regardless of its marker. However, I will illustrate two cases where a nonagent *ang*-phrase (i.e., patient) may function in the processes marked with an asterisk, namely, Control or Equi-target, and Pronoun interpretation.[8] It may be stressed once again that the focus in this section is to determine whether the processes that apply to the *ng*-agent may equally apply to the identically marked *ng*-patient.

---

8. The first set of two processes were cited in my previous works 1988 and 1995. For an extensive discussion of counter-arguments to other processes in column A based on some nonbasic structures, see Cena 1995.

3.1 Topicalization

I have shown before (De Guzman 1988, 1995) that, with respect to topicalization (or focus construction), the *ng*-agent/experiencer (term 1) may also be topicalized, provided its particle *ng* is replaced by *ang* in initial position. This particle gives the nominal the necessary focus or highlight, which the topicalization process bestows upon it, in addition to its relocation in initial position. In contrast, the *ng*-patient (Final 2-CHO) is not allowed to topicalize in exactly the same manner. If it does at all, it must occur without the obligatory particle *ang*. Moreover, this structure functions to show contrastive meaning more than anything else, which is uncharacteristic of a topicalized structure. For example:

(10)   binigyan   ng   madre   ng   tinapay   ang   pulubi
        gave-BF   1, ERG   nun   2-CHO, OBL   bread   2, ABS beggar
        'The nun gave the beggar (some) bread.'

    a. ang pulubi # binigyan (siya) ng madre ng tinapay   (Final 2)
    b. ang madre # binigyan (niya) ng tinapay ang pulubi   (Final 1)
    c. *ang/*ng tinapay # binigyan ng madre ang pulubi   (Final 2-CHO)
    d. tinapay # binigyan ng madre ang pulubi
        (hindi pera 'not money')   (Final 2-CHO)

In (10a), the general rule for topicalization applies to the *ang*-phrase, final term 2. In (10b), the topicalized phrase is the Final term 1 *ng*-agent, with *ng* replaced by *ang*. It will be noted that the topicalized phrase may optionally leave a corresponding pronoun copy. (10c) shows that when the *ng*-patient is topicalized, even when the *ng*-marker is replaced by the usual marker *ang*, the result is still ungrammatical. In (10d), with the preposed patient being unmarked and the contrastive phrase almost always being present, the structure appears to be virtually different from the first two. And, therefore, this last structure is considered to be semantically and syntactically different from the typical topicalized construction.

3.2 Raising

Raising normally applies to the *ang*-phrase (Final 2) of a clause complement of the matrix verb. But the *ng*-agent is not completely restricted from undergoing the same, provided it is marked appropriately according to the voice marking in the matrix verb. For example:

(11)   inasahan   ko-ng   ibibigay   ng   babai
        expected-OV I, ERG-COMP will.give-OV 1, ERG   woman

        ang   saging   sa   bata
        2, ABS banana R, OBL child
        'I expected that the woman would give the banana to the child.'

    a. inasahan ko ang babai-ng ibibigay (niya) ang saging sa bata
        (*ng*-agt to Abs raising)
    b. umasa ako sa babai-ng ibibigay (niya) ang saging sa bata
        (*ng*-agt to Obl raising)

> c. *inasahan ko ang saging na magbibigay ang babai (nito) sa bata
> (*ng*-pat to Abs raising)
> d. *umasa ako sa saging na magbibigay ang babai (nito) sa bata
> (*ng*-pat to Obl raising)

It is obvious that the matrix verb dictates the kind of marking the raised agent should take, as for example, Abs *ang* in (11a) and Obl *sa* in (11b). When tried on the *ng*-patient/object as in (11c) and (11d), even with the proper marking, the result is still ungrammatical. It will be noted further that with the *ng*-agent raising, an optional pronoun copy is permissible, which is not necessary with the usual raising of a Final 2 Abs. The occurrence of this pronoun copy of the raised constituent identifies the structure as being marked.

In (11c) and (11d), it may be countered that the reason the *ng*-patient cannot raise is due to its semantic content. But even when a nominal with a feature [+human], for example, *katulong* 'helper', is substituted for *saging* 'banana', the result is still ungrammatical. The immediate reading of the raised nominal is that of an agent instead of a patient. For example:

> (11) e. *inasahan ko ang katulong na magbibigay ang nanay (nito) sa amin
> (to mean: 'I expected the helper that mother will give to us.')

From the above illustrations, the *ng*-agent phrase is seen to be an active and an apparently favored alternative, even in those grammatical processes that are normally attributed to the syntactically relevant *ang*-phrase. However, there is an observable degree of markedness in the structure involved in these particular occurrences. In contrast, the *ng*-patient is virtually inoperative in such processes.

3.3 Control or Equi Structures

Kroeger (1993) has a whole chapter devoted to an extensive discussion of "equi or control constructions."[9] The usual "controllee" (a gap, indicated by Ø in the examples below, or a resumptive pronoun) in this type of construction refers to the agent or experiencer of the verb in the complement clause, be it conforming to a *ng* or an *ang* (Final 1 or Final 2) null nominal or its corresponding pronoun. As for the controller, it is the semantics of the matrix verb that determines which nominal performs this function. It may be the agent/experiencer, the recipient/addressee, or the patient/object/ goal. For example:

> (12) binalak   ng   tatay   na   magtanim Ø   ng   saging
> planned-OV 1, ERG father COMP plant-AV   2-CHO, OBL banana
> sa   likod ng   bahay
> OBL back   GEN house
> 'Father planned to plant bananas in back of the house.'

---

9. For a comparison of this type of constructions in the GB framework, using an ergative analysis, see Byma 1986.

(13) sinikap ni   Maria-ng   tulungan Ø   ang   estudyante
     tried-OV I, ERG Maria-COMP help-OV   2, ABS student
     'Maria tried to help the student.'

Besides these regular equi-constructions, Kroeger (1993:97ff) identifies a set of struc-
tures where the controllee is a nonagent (Final 2) whose agreement with its verb is man-
ifested on the voice affix. In these structures, (14) to (16) below, the controllee has the
semantic role of recipient, patient, and patient, respectively. Although Kroeger drew
them from various sources, there is reason to doubt the acceptability of these struc-
tures:[10]

(14) ?nagpilit   si   Maria-ng   bigyan   ng   pera   ni   Ben[11]
     insisted-AV ABS Maria-COMP to.give-RV OBL money ERG Ben
     'Maria insisted on being given money by Ben.'

(15) ?hinimok   ni   Maria   si   Juan-ng   suriin
     persuaded-OV ERG Maria ABS Juan-COMP examine-OV

     ng   bago-ng   doktor
     ERG new-LKR   doctor
     'Maria persuaded Juan to be examined by the new doctor.'

(16) ?natatakot   si   Pedro-ng   tanungin ng   pulis
     be.afraid ABS Pedro-COMP ask-OV   ERG police
     'Pedro is afraid to be questioned by the police.'

At this point, I cannot ascertain how other unbiased native speakers, that is, those
who have not been exposed to elicitation sessions on this topic or those who have not
been immersed in investigating these structures (which I believe might have colored
the judgment given to the above sentences) will react to these sentences when com-
pared with the following:

(14) a. nagpilit si Maria-ng bigyan **siya** ng pera ni Ben
        'Maria insisted on being given money by Ben.'
     b. nagpilit si Maria-ng **ma**bigyan Ø ng pera ni Ben

(15) a. hinimok ni Maria si Juan-ng suriin **siya** ng bago-ng doktor
        'Maria persuaded Juan to be examined by the new doctor.'
     b. hinimok ni Maria si Juan-ng **ma**suri Ø ng bago-ng doktor

---

10. To confirm the preference for the two suggested structures, I tested three other native speak-
    ers of Tagalog: Neri Cachero (Bataan), Luz Einsiedel (Batangas), and Albert Einsiedel
    (Manila). Without any prior questions or background information, each one was simply
    instructed to rate the three renditions of (14) to (16) as good, OK, or questionable. All of
    them invariably judged the sentences to be questionable that had the missing pronoun and a
    verb that was not of the "ability-involuntary-accidental" type.

11. The morpheme glosses have been revised to conform to the presentation in this paper. The
    same is done in succeeding examples.

(16) a. natatakot si Pedro-ng tanungin **siya** ng pulis
    'Pedro is afraid to be questioned by the police.'
  b. natatakot si Pedro-ng **ma**tanong Ø ng pulis

The modified structures in (a) obligatorily manifest an Abs pronoun (Final 2) as the controllee that is not an agent, referring to the controller marked Abs (Final 2) in the matrix clause; in (b), the nonagent controllee is a gap, allowed by the marked verb form in *ma-*, which Dell (1981) calls "ability-involuntary-accidental" or AIA form. These modified renditions of (14) to (16) are judged to be the more acceptable structures because they sound natural and effective.

The explanation that may be offered to account for the preference toward these two modified structures is that when the controllee/gap is not an agent, which is a rather unexpected turn, the resulting aberration must find some clue to guide the syntactic processing. By filling in the gap with the Abs pronoun referring to the preceding Abs (Final 2) phrase in the matrix clause or changing the complement verb into an AIA *ma-*form, which ordinarily indicates a nonagentive Final 2, the interpretation of the null or the pronominal controllee certainly takes it away from the usual agent. These two devices, being absent from the regular form of control structures, focus the attention on the patient or the recipient controllee. Similar to the less common target in raising constructions, the irregular nonagentive controllee also appears more marked in terms of the obligatory occurrence of a coreferential pronoun or of the use of the AIA verb form. This being the case, we can expect a patient to be a controllee, in such instances as above, except that it must be a Final 2. Consequently, this automatically disqualifies a *ng*-patient, which is a 2-CHO, from functioning as a controllee. Again, we see the disparity between the active participation of the *ng*-agent in control structures and the lack of it on the part of the *ng*-patient.

3.4 Pronoun Interpretation

For this process, I present only the interpretation of pronouns in conjoined structures using psych and action verbs. Consider the following sentences:

(17) a. naiinggit si   Ben$_i$ kay Ed$_j$ at   nayayamot   siya$_{i,*j}$ kay Lito
    envy-EV 2, ABS Ben OBL Ed and disgusted-EV he, 2, ABS OBL Lito
    'Ben envies Ed and he is disgusted with Lito.'

  b. kinaiinggitan ni   Ben$_i$ si   Ed$_j$ at   nayayamot   siya$_{i,j}$
    envy-OV   I, ERG Ben 2, ABS Ed and disgusted-EV he, 2, ABS
    kay Lito
    OBL Lito
    'Ben envies Ed and he is disgusted with Lito.'

In both coordinate structures, the Abs pronoun is an experiencer, based on the semantics of the verb. Normally, we would expect it to refer to the same semantic role in the preceding clause, hence in (17a), it is easy to predict that it picks out the Abs phrase *si Ben* as its antecedent. And indeed this is correct. It does not refer to the Obl phrase *kay Ed*, a specific object role, as antecedent because it does not match the semantic

role of experiencer. Moreover, the verb in the second clause has its own Obl object phrase that parallels that of the preceding clause. Compared with (17b), however, the Abs pronoun is ambiguous. It may refer to the experiencer noun marked Erg, whose semantic role matches it, or to the object phrase marked Abs. Somehow, the occurrence of the object phrase in the neutral form Abs, as a Final 2, overrides its semantic role feature and licenses it for possible antecedentship. Let us see what happens when the pronoun form is non-Abs, as in (18).

(18) a. naiinggit si   Ben$_i$ kay Ed$_j$ at nayayamot
envy-EV 2, ABS Ben OBL Ed   and disgusted-EV

si   Lito sa kaniya$_{i, j}$
2, ABS Lito OBL him
'Ben envies Ed and Lito is disgusted with him.'

b. kinaiinggitan ni   Ben$_i$ si   Ed$_j$ at   nayayamot
envy-DV   I, ERG Ben 2, ABS Ed   and disgusted

si   Lito sa kaniya$_{*i, j}$
2, ABS Lito OBL him
'Ben envies Ed and Lito is disgusted with him.'

The pronoun in the Obl form, a specific object role, has an ambiguous interpretation in (18a). It may refer to the Obl phrase, its object role match, in the first clause, or it may be interpreted as referring to the Abs-marked experiencer. In (18b), where the object role phrase in the initial clause is also the Abs phrase, the Obl pronoun object is unambiguous, because its semantic role match is at the same time the Abs phrase. Earlier, we observed that *ng*-patients, when realized as pronouns (being definite), may take only the Obl form. Let us compare this observation with the following conjoined structures consisting of forms manifesting the verb affixes *-in* and *m(pag)*-:

(19) a. minamahal ni   Ben$_i$ si   Lito$_j$ at   sinasamba
love-OV I, ERG Ben 2, ABS Lito and   worship-OV

niya$_{i, j}$ si   Maria
he, I, ERG 2, ABS Maria
'Ben loves Lito and he worships Maria.'

b. nagmamahal si   Ben$_i$ kay   Lito$_j$ at   sinasamba
love-AV   2, ABS Ben 2-CHO, OBL Lito and worship-OV

niya$_{i, *j}$ si   Maria
he, I, ERG 2, ABS Maria

(20) a. minamahal ni   Ben$_i$ si   Lito$_j$ at   sumasamba
love-OV I, ERG Ben 2, ABS Lito and worship-AV

si   Maria sa kaniya$_{*i, j}$
2, ABS Maria he, 2-CHO, OBL
'Ben loves Lito and Maria worships him.'

235

b. nagmamahal si   Ben$_i$ kay   Lito$_j$ at   sumasamba
   love-AV   2, ABS Ben 2-CHO, OBL Lito and worship-AV
   si   Maria **sa kaniya**$_{i,j}$
   2, ABS Maria he, 2-CHO, OBL

In the preceding illustrations, there is a similarity of pronoun interpretation with the first set, even when there is a difference in semantic role participation. The Erg agent pronoun in (19a) may refer either to the Erg agent *ni Ben* or to the Abs phrase *si Lito*. In (19b), however, the same pronoun refers only to the agent Abs phrase. To refer to the Obl phrase would result in a semantic role contradiction. The patient pronouns in (20a) and (20b), being in Obl form, are identical to the object ones shown in (18). Thus, in (20a), only the Abs patient *si Lito* is taken as the antecedent, whereas in (20b), the pronoun may refer to either the Obl patient phrase or the Abs agent.

We can sum up the observations above by saying that the pronoun in the second clause of a conjoined structure, whatever form it may have—ergative, oblique, or absolutive—makes use of semantic role-matching with a given noun phrase in the first clause for coreference, and at the same time, the cooccurring Final 2, Abs phrase exerts its grammatical influence on the pronoun by being an "all-around" antecedent. The pronoun is unambiguous when the nominal bearing its semantic role-match is at the same time in the Abs form.

It may be recalled that in control structures where the Abs pronoun has a nonagent role, it can only refer to the Abs phrase in the matrix clause. This confirms the syntactic significance of Final 2, the Abs phrase.

4. Semantic/Case Relations and Grammatical Relations

From the preceding, we see that even as the *ng*-agent is rather active in syntactic processes, the identically marked *ng*-patient is not. If, as has been shown in previous studies, the patient is the most salient of all semantic roles in Tagalog basic clause structures (and most likely in the other Philippine languages as well), there must be an explanation for why the *ng*-patient does not perform any remarkable function. Why is it not as significant as when it is a Final 2 marked with the Abs *ang*? Is it because it has a different nonnuclear GR? If, as some linguists would maintain for Tagalog, the important feature in the syntax is not the GR of the nominals but the semantic relation, why then does it not compare favorably with the *ng*-agent? Most accounts dealing with "subject" and "object" cited here are silent on this issue. I can only mention two frameworks that offer some explanation.

In RG, this is accounted for by a change of grammatical relation, term 2 (Direct Object) to a nonterm 2-chomeur when another cooccurring semantic relation or nominal usurps term 2, via an advancement rule. As a nonterm, this patient is usually marked with the particle *ng* (which in other types of constructions may alternate with the particle *sa* to convey a [+definite] or [+spfc] feature). Being now a nonterm, the *ng*-patient has become virtually an adjunct, or an oblique phrase at best, one that comes lower in the hierarchy of GRs. It is a nonparticipant in a number of syntactic

operations, unlike the nuclear terms 2 and 1. It may be mentioned that certain verbs require their final 2-CHO patients to be marked only with the Obl *sa*, in which case it follows the functions of other oblique phrases, as for example, adjunct fronting and WH-question.

In the Lexicase framework, Starosta (1986:92) accounts for the properties of the "nonsubject 'direct objects' . . . as having a nonnuclear case relation, Means" via some lexical derivation rules. This account, however, is deemed problematic on the grounds that changing the semantic roles in derivationally-related verb forms runs counter to the conceptual notions associated with the verb forms as expressed by their cooccurring case relations.[12] Instead of a semantic modification or change in semantic perception, because it is believed that the same participant or semantic role still obtains in the derived verb form (indicated only by a change in voice affix), the change is a morphosyntactic one. The grammatical function of the patient goes from a nuclear term 2 (direct object) to a nonterm or oblique, and as such it takes a different particle marking. Thus, we observe the consequent "disability" of the *ng*-patient to participate in just about any syntactic process (De Guzman 1997:315–318).

It is interesting to note Starosta's reference to the *ng* patient as the "nonsubject 'direct object'" from which we may infer that it is still a "direct object." Based on this, will it be off the mark to construe that the patient nominal when marked as an *ang*-phrase may be referred to as the "subject 'direct object'"? Or can the latter GR be simply called "Direct Object" for the sake of uniformity across languages, thereby avoiding the use of contradictory terms in the designation of GRs?

For Kroeger (1993:47–48), *ng*-patients are "terms," that is, "Objects," (versus nonterms, in RG) based on their negative occurrence in the process he labels "Adjunct Fronting," and also on their ability to control the gap in a participial *ng* clause. The former ground is shaky because it is not only *ng*-agents and *ng*-patients that do not undergo "Adjunct Fronting"; even the oblique *ng*-instrument and *ng*-measure phrases may not either. Yet, he would not bestow termhood to instruments or to measures. It could very well be that the restriction here has to do with the marker *ng* of these phrases, rather than with their grammatical functions. As for the illustrations used as evidence for the second ground, they need further verification to be admissible as support for the *ng*-patient's status as a term.[13]

As for the *ng*-agent, when approached in an ergative analysis in RG, where term 2 is most central, it is also a nuclear term ranking second to the *ang*-phrase (usually, the patient) term 2 (Direct Object). Kroeger (1993) decries the previous RG passive analysis of this nominal because it ends up as a 1-chomeur, and rightly so, for it is easy to show counter-evidence for the *ng*-agent being a term rather than a nonterm. In a foot-

---

12. It is possible to conceive of similar lexical derivation rules in which it is the case forms (or grammatical relations) of the nominals bearing some specified case relations that change from a nonterm to a term, and vice versa.

13. The judgment on the ambiguity or nonambiguity in examples (50b) and (51) is to me questionable. It appears that the rule given above on pronoun interpretation applies to these structures as well.

note, he mentions that the grammatical relation of this "nonsubject Actor" should technically be something like "OBJ$_{agt}$" (1993:100). This suggestion derives from the argument that the *ng*-agent is an "object" (not an oblique) in the grammar. And since the GR "subject" has already been allocated to the *ang*-phrase, his only recourse is to refer to it as the OBJ agent. This move, however, results not only in confusion but also in the noncomparability of the supposed general labels "subject" and "(direct) object." Again, for the sake of uniformity and consistency in description, could the *ng*-agent be simply referred to as "subject," if we grant that the *ang*-phrase is "(direct) object?" As for the *ng*-patient, it would join the nonnuclear GR called "obliques."

Based on previous studies where linguists are divided in their views concerning the existence or nonexistence of grammatical relations in Tagalog, particularly subject and object, or where the assignment of these two relations are in opposite semantic roles, we have to admit that the problem is alive but not well. To many, it is still baffling. As we continue to reflect further on the arguments for or against the individual proposals and to continue our quest for some tenable solutions, we may expand our horizon beyond the properties attributed to the *ang*-phrase and the *ng*-agent and consider the *ng/sa*-patient as well, if indeed Tagalog could be governed by semantic relations instead of grammatical relations.

5. Conclusion

Upon reviewing the various accounts of "subject" and "object" in Tagalog (and the other Philippine languages) and the contradictory views that have ensued, I still find it appropriate to refer to the Direct Object (term 2) in an ergative analysis of Tagalog as the most central grammatical relation for three reasons: (i) it can host the other nominals in this function, (ii) it shows a distinct marker *ang*, labeled Abs, (or its corresponding Abs substitutes), and (iii) it has a specified range of syntactic processes associated primarily with it. Right next to the Object in importance is the Subject (term 1) whose morphological properties, as well as grammatical functions, distinguish it clearly from the *ng*-patient, a nonterm. Its grammatical functions are still largely different from those of the Direct Object (especially if these two are compared using the same basic verbal constructions), and even when it begins to encroach on some processes that are clearly attributed to the Direct Object (term 2), the resulting structure becomes more marked. The *ng*-patient, in contrast, as a nonterm or oblique, has lost its previous central GR. It is only when it occurs with a definite marker *sa* that it can participate in some identifiable processes shared with other *sa* obliques. This explanation, provided by the present account for the inoperativeness of the *ng*-patient, has yet to be dealt with in other accounts.

Abandoning the two primary grammatical relations, we cannot unify the characteristic features of the various semantic roles that may manifest the Direct Object relation. We will not be able to account for the formal and syntactic differences between the *ng*-agent and the *ng*-patient. But by keeping them, much more is gained than lost in terms of a more coherent description.

REFERENCES

Bell, Sarah J. 1976. Cebuano subjects in two frameworks. Ph.D. dissertation, Massachusetts Institute of Technology.

———. 1983. Advancements and ascensions in Cebuano. In *Studies in relational grammar I,* ed. by David M. Perlmutter, 143–218. Chicago: University of Chicago Press.

Brainard, Sherri. 1996. Why the "focused NP" is not the subject in Philippine languages: Evidence from Karao. *Philippine Journal of Linguistics* 27:1–47.

Byma, Gary. 1986. An ergative analysis of Tagalog. M.A. thesis, University of Calgary.

Cena, Resty M. 1977. Patient primacy in Tagalog. Paper read at the Linguistic Society of America Winter Meeting, Chicago, Illinois.

———. 1995. Surviving without relations. *Philippine Journal of Linguistics* 26:1–32.

De Guzman, Videa P. 1988. Ergative analysis for Philippine languages: An analysis. In *Studies in Austronesian linguistics,* ed. by Richard McGinn, 323–345. Athens, Ohio: Ohio University Press.

———. 1995. Experiencers of psych verbs in Tagalog. In *Grammatical relations: Theoretical approaches to empirical questions,* ed. by Clifford S. Burgess, Katarzyna Dziwirek, and Donna Gerdts, 45–62. Stanford, Calif.: Center for the Study of Language and Information.

———. 1997. Verbal affixes in Tagalog: Inflection or derivation? In *Proceedings of the Seventh International Conference on Austronesian Linguistics,* ed. by Cecilia Ode and Wim Stokhof, 303–325. Amsterdam: Rodopi, B.V.

Dell, Francois C. 1981. On certain sentential complements in Tagalog. *Philippine Journal of Linguistics* 12.1:11–28.

Gerdts, Donna. 1988. Antipassives and causatives in Ilokano: Evidence for an ergative analysis. In *Studies in Austronesian Linguistics,* ed. by Richard McGinn, 295–321. Athens, Ohio: Ohio University Press.

Gil, David. 1984. On the notion of 'Direct Object' in patient prominent languages. In *Objects: Towards a theory of grammatical relations,* ed. by Frans Plank, 87–108. London: Academic Press.

Hoekstra, Eric. 1986. On the structural subject position in Tagalog. *Lingua* 70: 41–56.

Kroeger, Paul. 1993. *Phrase structure and grammatical relations in Tagalog.* Stanford, Calif.: Center for the Study of Language and Information.

Naylor, Paz Buenaventura. 1980. Linking, relation-marking, and Tagalog syntax. In *Austronesian studies: Papers from the Second Eastern Conference on Austronesian Languages,* ed. by Paz Buenaventura Naylor, 33–50. Ann Arbor, Mich.: Center for South and Southeast Asian Studies, University of Michigan.

O'Grady, William. 1987. The nature of Philippine ergativity. Ms.

Pawley, Andrew, and Lawrence A. Reid. 1980. The evolution of transitive constructions in Austronesian. In *Austronesian studies: Papers from the Second Eastern Conference on Austronesian Languages,* ed. by Paz Buenaventura Naylor, 103–130. Ann Arbor, Mich.: Center for South and Southeast Asian Studies, University of Michigan.

Payne, Thomas. 1982. Role and reference-related subject properties and ergativity in Yup'ik Eskimo and Tagalog. *Studies in Language* 6.1:75–106.

Schachter, Paul. 1976. The subject in Philippine languages: Topic, actor, actor-topic, or none of the above? In *Subject and Topic,* ed. by Charles N. Li, 419–518. New York: Academic Press.

———. 1977. Reference-related and role-related properties of subjects. In *Syntax and semantics 8: Grammatical relations,* ed. by Peter Cole and Jerrold M. Sadock, 279–306. New York: Academic Press.

Shibatani, Masayoshi. 1988. Voice in Philippine languages. In *Passive and Voice,* ed. by Masayoshi Shibatani, 85–142. Amsterdam: John Benjamins.

Starosta, Stanley. 1982. Case relations, perspective, and patient centrality. *University of Hawai'i Working Papers in Linguistics* 14.1:1–34.

————. 1986. Focus as recentralization. In FOCAL I: *Papers from the Fourth International Conference on Austronesian Linguistics,* ed. by Paul Geraghty, Lois Carrington, and S. A. Wurm, 73–95. Pacific Linguistics C-93. Canberra: Australian National University.

Travis, L., and E. Williams. 1982. Externalization of arguments in Malayo-Polynesian languages. *Linguistic Review* 2:57–77.

15: NOTES ON A POSSESSIVE CONSTRUCTION IN THE FORMOSAN LANGUAGES

Elizabeth Zeitoun
Institute of Linguistics, Academia Sinica

1. Introduction[1]

Three types of possessive constructions can be distinguished in the Formosan languages, as exemplified in (1–3).[2]

(1) Tsou (Tfuya)
    a. pan to peisu-si
       exist OBL money-his
       'He has money.'

    b. pan to peisu ta moʔo
       exist OBL money OBL Mo'o
       'Mo'o has money.'

(2) Tsou (Tfuya)
    a. ʔe   tposɨ eni zou nu-taini
       NOM book this be POSS-his
       'This book is his.'

    b. ʔe   tposɨ eni zou nu ta moʔo
       NOM book this be POSS OBL Mo'o
       'This book is Mo'o's.'

---

2. Abbreviations in the glosses include: ACC, accusative; AF, Agent Focus; ART, article; COP, copula; GEN, genitive; LF, Locative Focus; LIN, linker; LOC, locative; NOM, nominative; OBL, oblique; P, preposition; POSS, possessive; RED, reduplication; S, singular. In some cases, irrelevant glosses that may appear in the original works in which the data are found have been intentionally omitted.

(3) Tfuya Tsou
    a. mo tuyo ʔo   ʔo-ʔoko-si
       AF   three NOM RED-child-his
       'He has three children.'
    b. mo tuyo ʔo   ʔo-ʔoko to moʔo
       AF   three NOM RED-child OBL Mo'o
       'Mo'o has three children.'

What makes these examples different from their corresponding nonpossessive counterparts is the fact that the theme has to be "formally" possessed. (1a–b) are structurally identical to existential sentences: these are headed by the existential verb *pan* 'exist', which is followed by a theme NP (possessed entity) and eventually a locative phrase (possessor). (2) and (3) are characterized by the nonoccurrence of the existential *pan:* (2) is an equational/nominal sentence; (3) is, on the other hand, a verbal sentence with the quantifier/numeral *tuyo* 'three' used as the main lexical verb of the sentence.

In the present essay, I deal only with the possessive type of structure found in (1) and compare it in the following Formosan languages (dialects): Central Amis (Changpin), Squliq Atayal (Wufeng), Bunun (Isbukun), Kavalan (Hsinshe), Northern Paiwan, Puyuma (Nanwang), Rukai (Mantauran and Budai), Saisiyat (Tungho), Seediq (Paran and Truku). Remaining problems in Tsou are outlined in the conclusion. For ease of comparison, I restrict myself to the study of affirmative sentences. Remarks on negative sentences are made just in passing, when relevant to the discussion. Taking Freeze's (1991) study as a starting point (section 2) but departing from him in some respects, I try to show in section 3 that the Formosan languages exhibit either one or two *have*-structures that differ not only semantically but also syntactically. My aim is twofold. On the one hand, I examine the syntactic variation among languages that share the same structure and on the other hand, I analyze the semantic differences yielded by the use of these two structures in languages where they both coexist. In so doing, careful attention is paid to (i) the marking of the verb and (ii) the marking of the nominal arguments.

Though this essay is not framed in any particular formal theory, I hope that the typological analysis provided here will contribute to a better understanding of this type of sentential structure in the Formosan languages.

2. Freeze's (1991) Proposal

Working in GB theory, Freeze (1991; 1992) proposes a cross-linguistic analysis[3] that relates existential, possessive, and locative constructions.[4] While I do not agree with all his assumptions and conclusions, his two articles contain some interesting impli-

---

3. Freeze's (1991) study is illustrated mainly with examples taken from Austronesian languages other than Formosan, while his 1992 investigation includes languages belonging to different families.

4. This hypothesis is not new. It is also found in Lyons 1967, Kuno 1971, Clark 1978, and Huumo 1996, among others.

cations concerning language universals. To me, they shed light on syntactic and semantic variation that can be observed in the Formosan languages but that has never been systematically compared and accounted for. In this section, I briefly summarize his study, indicate those respects in which I disagree with him, and present the line of reasoning I will pursue in section 3.

Freeze's main purpose is to demonstrate that existential, possessive, and locative constructions are all derived from the same D-structure, a structure that involves a preposition at the head of the predicate phrase, a theme, and a locative. For him, variation among languages is "very restricted" and "highly predictable."[5] He argues that

---

5. There are two cross-linguistic generalizations to be found in the Formosan languages that I will not develop further in the course of this study. One concerns the fact that there is no restriction concerning the nature of the possessed NP: it may refer to a [±human] or [±animate] entity, as shown in (i. a–c):

  (i) Truku Seediq (Tsukida 1999:626)
    a. niqan bubu-na   ka   laqi nii
        exist mother-3S.GEN NOM child this
        'This child has a mother.'

    b. niqan kiŋal katin-na   ka   senaw nii
        exist one   cow-3S.GEN NOM man   this
        'This man has a cow.'

    c. niqan patas-na   ka   laqi nii
        exist book-3S.GEN NOM child this
        'This child has a book.'

On the other hand, as pointed out by Freeze (1991, 1992), the possessor is usually [+human]. In examples of this sort, there is no restriction concerning the relation of possession. That is, it may refer to an inalienable or to an alienable relationship. If the possessor is [−human], the possessed entity must bear some kind of inherent relationship with it. Hence, while (ii. a–b) are grammatical examples, only (iii. a) but not (iii. b) is acceptable. The use of the locative preposition *i* in (iii. c) gives back the example its grammaticality but prevents it from getting a possessive interpretation.

  (ii) Central Amis
    a. ira   ku   kahəŋaŋay a   ŋosoʔ ni   aki
        exist NOM red   LIN nose GEN Aki
        'Aki has a red nose.'

    b. ira   ku   pusi ni   aki
        exist NOM cat   GEN Aki
        'Aki has a cat.'

  (iii) Central Amis
    a. ira   ku   kahəŋaŋay a   papah nu kilaŋ
        exist NOM red   LIN leave GEN tree
        'The tree has red leaves.'

    b. *ira   ku   pusi nu kilaŋ
         exist NOM cat   GEN tree

    c. ira   ku   pusi i   kilaŋ
        exist NOM cat LOC tree
        'There is a cat on the tree.'

the difference between existential and locative constructions can be reduced to a sim-
ple movement in the Spec IP (left empty at D-structure) of one of the two NPs as a
result of definiteness: if the theme is definite, it moves into the subject position, and
the sentence turns into a locative construction. If, on the other hand, the theme is
indefinite, the locative phrase moves into the subject position, and the sentence is
given an existential interpretation. In this context, Freeze shows that two *have*-struc-
tures can be distinguished in the Austronesian languages. The first (denoted by H_1) is
locative in nature and thus shares the same structure as the existential. That is, the loc-
ative phrase, usually introduced by a preposition, is the subject of the sentence. The
second (denoted by H_2) is not locative and differs from the former in that (i) it has a
DP predicate instead of a PP predicate, (ii) the theme is the subject of the sentence,
and (iii) the copula may agree with the theme argument. Despite these differences,
Freeze posits that the second structure shares many properties with the first, because
of the presence of a locative copula—usually the same as the one occurring in exis-
tential sentences—and the fact that the possessor is semantically interpreted as a loca-
tion. The following two pairs of examples illustrate the parallelism between
existential (a) and possessive (b) sentences in Kiribati and Palawan. Kiribati exem-
plifies the first possessive structure and Palawan, the second. The subject in each sen-
tence is enclosed in brackets for the sake of clarity.

> (4) Kiribati (adapted from Freeze 1991:5)
>   a. iai   booki [iaon te   taibora]
>   COP book on   ART table
>   'There are some books on the table.'
>
>   b. iai   tabeua boki [irou-n te   mʷmʷaane]
>   COP some   book to-him the man
>   'The man has (some) books.'
>
> (5) Palawan
>   a. ŋ-ŋar-ŋii a bilis [er a sers-ek]
>   3s-be-it   dog P   garden-my
>   'There is a dog in my garden.' (Freeze 1992:563)
>
>   b. ŋ-ŋar [a berrul aʔad]
>   3s-be   raft-his man
>   'The man has a raft.' (lit. the man's raft is) (from Freeze 1991:8)

Some languages, such as Palawan, exhibit only one of these two structures, while
others, like Kiribati, may display both. Compare (6a) and (6b).

> (6) Kiribati (adapted from Freeze 1991:5)
>   a. iai   tabeua boki [irou-n te   mʷmʷaane]
>   COP some   book to-him the man
>   'The man has (some) books.'
>
>   b. iai [am boki]
>   COP your book
>   'Do you have a book?'

To account for the coexistence of the two different *have*-structures in the same language, Freeze (1991:12) suggests that H_2 be interpreted as "an assertion of ownership by the 'possessor'. The locative H_1 does coexist with such a structure: the H_2 for inalienable possession and the H_1 for alienable possession."

In my opinion, one of the main contributions of Freeze's (1991) study lies in the distinction he makes between these two *have*-structures, which exhibit not only syntactic variation but may also yield a different interpretation (alienable vs. inalienable relation) in languages where they coexist. Below, I will also argue that the Formosan languages exhibit two different *have*-structures. Though we seem to reach the same conclusion, my own interpretation of the data differs from Freeze's in at least three fundamental respects, however.

Concerning the marking of nominal arguments, Freeze makes the following assumptions: (i) in H_1, the locative phrase is the subject of the sentence, as a result of the "definiteness effect"; (ii) the locative phrase/possessor is usually preceded by a preposition, as in Kiribati; (iii) there are languages such as Tagalog where a preposition occurs in existential but not in possessive sentences, because they make a distinction between [+human] vs. [–human] subjects. Compare for instance (7a) and (7b).

(7) Tagalog (from Freeze 1992:585)
    a. may gera sa ewropa
       COP war in Europe
       'There is a war in Europe.'

    b. may relos aŋ naanai
       COP watch ART mom
       'Mom has a watch.'

Contrary to what Freeze demonstrates, in most, if not all, Formosan languages, the "definiteness effect" does not play any role in the selection of the theme or the locative phrase as subject in existential vs. locative constructions. (For details, see Zeitoun et al. 1999.) This claim is mainly supported by case assignment properties. That is, it is the theme that is usually marked as nominative.[6] On the other hand, there are languages where the locative phrase is the subject of the sentence, but in this case, no preposition precedes the possessor, because it is marked as nominative. Finally, in many languages, existential sentences are, to some extent, parallel to possessive sentences, but while the locative phrase is preceded by a locative preposition in the former, it is marked by the genitive in the latter. In such constructions, the theme is the subject of the sentence, as the following examples taken from Central Amis reveal.

(8) Central Amis
    a. ira [ku wawa] i paputal
       exist NOM child LOC outside
       'There are children outside.'

---

6. This analysis also departs from what has been commonly assumed regarding the Formosan languages, for example, that existential/possessive sentences are subjectless (cf. Li 1973:178ff).

b. ira   [paysu ni   panay]
exist money GEN Panay
'Panay has money.'

Regarding the marking of the copula, Freeze establishes a distinction between 'have' (H₁) and 'be' (H₂): while the first is unmarked, the second is said to usually agree with the theme. Though I am not at all familiar with the languages Freeze deals with, it seems clear, looking at the data provided to support his analysis, that (i) in many languages (e.g., Kiribati as in [6b], K'ekchi', etc.), there is no such agreement on the copula and (ii) in languages (e.g., Palawan, Tongan) where agreement is found, it also occurs in existential sentences. Compare for instance (5a) and (5b). This may lead us to wonder about the well-foundedness of the distinction established between 'have' and 'be'. In the Formosan languages, there seems to be no such pronominal agreement in this type of possessive construction. What we do find, on the other hand, is that in possessive sentences where the locative phrase is the subject, the verb may be marked by the -*an* suffix, whereas in sentences where the theme is subject, the occurrence of -*an* is prohibited. Thus while (9a) is correct, (9b) is totally ungrammatical.

(9) Isbukun Bunun
    a. ʔaið(a)-an-ik   ʔuvað
       exist-*an*-IS.NOM child
       'I have a child.'

    b. *ʔaið(a)-an   ʔinak ʔuvað
       exist-*an*   my   child

Freeze emphasizes that the variation between the two structures is reflected at D-structure: H₁ involves a prepositional phrase as the head of the predicate, H₂ a noun. I have several comments regarding this claim: (i) based on a fair amount of evidence, I have shown (Zeitoun et al. 1999) that, in most Formosan languages, the existential morpheme heading existential, possessive, and locative constructions should be treated as a (full lexical) verb;[7] (ii) if the distinction established between these two structures is made in terms of "transitivity," there is no need to posit different structures of this sort at D-structure; and (iii) in most Formosan languages, the two *have*-structures exhibit exactly the same syntactic characteristics as those found in existential constructions.

In conclusion, I believe, like Freeze but for different reasons, that two *have*-structures can be distinguished in the Formosan languages and that this distinction should be based on subject selection. In the first structure, the existential/locative verb behaves as a two-place predicate with the theme/possessed entity marked as accusative or oblique and the locative phrase/possessor as nominative. This construction can be glossed as 'Y has X'.[8] In the second, the existential/locative verb is (in most cases) a one-place predicate. The theme/possessed entity is, in any case, marked as nominative, while the syntactic role of the possessor varies among languages. This second

---

7. In a recent paper, Yeh (1998) has provided additional data on Saisiyat supporting this claim.

8. The distinction established in the glosses 'have' vs. 'exist' is made only for the sake of convenience.

construction corresponds to 'Y's X exists'. In the next section, I present a typological study of these two structures in the Formosan languages.

3. Two *have*-Structures in the Formosan Languages

The main purpose of this section is to provide an overview of the possessive construction by highlighting the major characteristics and idiosyncracies of each language and showing in what respects they differ from one another.

My presentation of the data requires a few explanations. First, in order to be able to compare the data at hand, I have been careful to choose the same kind of examples in each language (e.g., "he has money" or "John has a child"). Second, in the linguistic networks/families (e.g., Rukai or Seediq) where two or more dialects display interesting divergences, I have included the two most representative ones for comparative purposes. Third, only semantic alternations will be discussed in section 3 when comparing languages that display the two *have*-structures.

I first consider languages where the structure 'Y has X' is found. I then turn to languages where the structure 'Y's X exists' occurs, before dealing with a set of languages where the two structures coexist. I pay particular attention to the variation that exists among these languages in the marking of the verb and its nominal arguments.

3.1 Y Has X

In Paran Seediq, Bunun Isbukun, Saisiyat, and Kavalan, the first structure 'Y has X' is found. The verb is a two-place predicate, with the possessor marked as nominative and the theme as oblique/accusative.[9]

    (10) a. Paran Seediq (from Chang 1997a)
            niqan-<u>ku</u>   laqi
            exist-1S.NOM   child
            'I have a child.'

          b. Isbukun Bunun
            ʔaið(a)-an-<u>ik</u>   ʔuvað
            exist-LF-1S.NOM child
            'I have a child.'

    (11) a. Tungho Saisiyat (Marie Yeh, pers. comm.)
            <u>yako</u> hayðaeh ka   rayhil
            1S.NOM have   ACC money
            'I have money.'

          b. Hsinshe Kavalan (from Lee 1997:125)
            yau-<u>iku</u>   tu   keɾisiw
            exist-1S.NOM ACC money
            'I have money.'

---

9. Some languages do not make any distinction between accusative, locative, and genitive, but exhibit an oblique case marker that takes over these functions.

The most striking difference between these two sets of languages concerns the marking of the verb. In Bunun and Seediq, the verb is marked with the suffix *-an*,[10] whereas in Saisiyat and Kavalan, the verb is unmarked. Chang (1997a) claims that in Paran Seediq, the verb has been grammaticalized and that the suffix *-an* is already fossilized, and thus that it has nothing to do with focus/voice. Chang's analysis is based, in fact, on what I believe to be a wrong assumption, that the theme should be analyzed as the subject of the sentence, and thus bear the nominative case. Because the theme NP is (usually) unmarked for case in such a structure, he argues in support of his point of view that if *niqan* were really a verb marked as LF, the pronoun occurring on the verb should be in the genitive case. However, it is not, as the ungrammaticality of (12) reveals.

(12) Paran Seediq (from Chang 1997a)
      *niqan-<u>mu</u>   kiŋan laqi
      exist-1S.GEN one   child

Lin (1996), on the other hand, was the first to call attention to the occurrence of the LF *-an* suffix in Bunun possessive sentences. Like him, I believe that in Bunun and Paran Seediq, the verb is marked as LF. This analysis is supported by (i) subject selection and (ii) coreference restrictions. The grammatical variation exemplified in Bunun in (13a, b) shows that the verb suffixed by *-an* can only select a locative phrase (i.e., the possessor) as subject of the sentence, as is the case in Paran Seediq. On the other hand, if it remains unmarked, as in (13c), it cannot take the locative phrase as its subject. In Paran Seediq, the nominative pronoun *yaku*, preceded by the nominative marker *ka*, is coreferent with the pronominal suffix *-ku* attached to the verb, and thus the theme cannot be marked simultaneously as nominative.

(13) Isbukun Bunun
    a. ʔaið(a)-an-ik   ʔuvað
        exist-LF-1S.NOM child
        'I have a child.'

    b. *ʔaið(a)-an ʔinak ʔuvað
        exist-LF   my   child

---

10. In negative sentences, Isbukun Bunun differs from Paran Seediq, in that the former, but not the latter, exhibits a negator marked by the suffix *-an* in this type of construction. Compare the grammaticality of (i) and (ii):

(i) Isbukun Bunun
    ʔuka-an-ik   ʔuvað
    not.exist-LF-1S.NOM   child
    'My child does not exist.'

(ii) Paran Seediq
    a. uka-ku   laqi
        not.exist-1S.NOM   child
        'I have no child.'

    b. *uka-an-ku   laqi
        not.exist-LF-1S.NOM child

    c. *ʔaið(a)-ik   ʔuvað
        exist -1S.NOM   child

(14) Paran Seediq (based on Chang 1997a)
       niq-an-ku   laqi (ka yaku)
       exist-LF-1S.NOM child NOM 1S
       'I have a child.'

Other variation among these four languages can be observed, most notably the position of sentential arguments—Saisiyat is turning into an SVO language—and the use of bound/free pronouns. Abstracting word order and pronominal variation, these two pairs of examples can be schematized as in (15).

(15) a. Paran Seediq and Isbukun Bunun
       exist-*an*   [Y]$_{NOM}$ [X]$_{ACC}$ (lit. 'Y is the place where X exists')

      b. Saisiyat and Kavalan
       exist/has   [Y]$_{NOM}$ [X]$_{ACC}$ (lit. 'Y has X')

A word should be said about Puyuma, where, for reasons still unaccounted for, this structure is found only in negative and not in affirmative sentences.

(16) Nanwang Puyuma (from Tan, 1997:86ff)
    a. *ulaya-ku   ḍa paisu
       exist-1S.NOM OBL money

    b. unian-ku   ḍa paisu
       NEG-1S.NOM OBL money
       'I have no money.'

3.2 Y's X Exists

The second and most prominent structure 'Y's X exists' is found in various languages, including some of those mentioned above. However, the variation among these languages, as well as among dialects belonging to the same language, is extensive.

What characterizes these languages mostly is that the possessive construction is headed by an existential verb unmarked for focus that introduces a single (nominative) NP. The theme (or possessed entity) is the head of this constituent; the possessor may be expressed by means of a pronoun or a noun as an attribute. Variation among these languages includes the following:

    (a) VERB FORMATION. In Amis and Paiwan, the existential verb can be analyzed as the fusion of the locative preposition *i* and the demonstrative 'that'. In Budai,[11] the existential verb is made almost exactly on the same pattern,

---

11. While the verb formation of *yakai* is quite transparent in Budai, it is rather opaque in Mantauran Rukai. For one thing, I am quite unable to account for the occurrence of *om-* (> *om-iki*) in Mantauran Rukai, a language that, like the other Rukai dialects, is characterized by the absence of focus marking. Besides, the root *iki* has nothing to do with the demonstratives found in this language (cf. *ʔina* 'this', *ana* 'that', *ðona* 'that').

except that *i-* is, in this language, used as a verbalizer; *kai* means 'this' and the verb is inflected for tense/aspect (cf. realis *a-*; for details, see Zeitoun et al. 1999). In Kavalan, *yau* also means 'that'.

(b) NOMINAL CASE MARKING OF THE THEME. The Wufeng dialect of Atayal is part of the Squliq branch, where nominative NPs have become unmarked for case. In Paiwan, no case marker seems to occur before the noun phrase. This is due to the elision of one of the two juxtaposed identical vowels (*izua a > izua*).[12] In Puyuma, Tan (1997) has shown that the nominative case does not occur before a possessed noun. In Mantauran Rukai, nominative NPs are usually unmarked for case.

(c) MARKING AND POSITION OF THE POSSESSOR INDICATED THROUGH PRO-NOMINAL MEANS. In some languages, the bound pronoun may precede the head noun, as in Paiwan and Puyuma; in others (e.g., Amis, Atayal, Kava-lan, Bunun, and Rukai) it can only follow.

(d) MARKING AND POSITION OF THE POSSESSOR INDICATED THROUGH NOMI-NAL MEANS. The possessor may be preceded by (i) a genitive case marker, as in Amis, Atayal, Paiwan, and Kavalan, or (ii) by an oblique case marker, as in Puyuma and Budai Rukai. In both instances, word order variation is rather restricted, with the possessor usually following the head noun. In Mantauran Rukai, the possessor is unmarked and may either precede or follow the head noun. In Nichols's (1986) terms, most languages (except Puyuma and Man-tauran Rukai) behave like "dependent-marking" languages, that is, the pos-sessive relation is marked on the modifier by means of nominal case marking. Puyuma exhibits a double marking: the two nominal arguments are marked. Mantauran Rukai displays, on the other hand, a head-marking pat-tern in which the possessive relation is marked on the modifiee/theme.

Variation of these types can be seen in the following examples.

(17) Central Amis
    a. ira   ku   wawa-aku
       exist NOM child-my
       'I have (a) child.'

    b. ira   ku   wawa ni   panay
       exist NOM child GEN Panay
       'Panay has a child.'

---

12. The case marker *a* is not elided if the existential verb is followed by a morpheme ending in a consonant (e.g., *izua-anan a*), as (i) shows:

    (i) Northern Paiwan (Anna Chang, pers. comm.)
       izua-<u>anan</u> <u>a</u>   su-paysu
       exist-still NOM your-money
       'Do you still have money?'

(18) Wufeng Atayal
    a. cyux laqiʔ-mu
       exist child-my
       'I have a child.'

    b. cyux laqiʔ naʔ pilay
       exist child GEN Pilay
       'Pilay has a child.'

(19) Northern Paiwan (from Anna Chang, pers. comm.)
    a. izua ku-paysu
       exist my-money
       'I have money.'

    b. izua paysu ni palaŋ
       exist money GEN Palang
       'Palang has money.'

(20) Isbukun Bunun
    a. ʔaiðaʔ ʔinak ʔuvað
       exist my child
       'I have a child.'

    b. ʔaiðaʔ ʔalaŋ tuʔ ʔuvað
       exist Alang LIN child
       'Alang has a child.'

(21) Hsinshe Kavalan
    a. yau sunis-ku
       exist child-my
       'I have children.' (based on Chang 1997:46)

    b. yau keɾisiw ni abas
       exist money GEN Abas
       'Abas has money.' (from Dorinda Liu, pers. comm.)

(22) Nanwang Puyuma
    a. ulaya ku-paisu
       exist my-money
       'I have money.' (from Tan, pers. comm.)

    b. ulaya tu-tiɭil kan sigimuɭi
       exist his-book OBL Sigimuli
       'Sigimuli has a book.' (from Tan 1997:85)

(23) Budai Rukai
    a. yakai ku paisu-li
       exist NOM money-my

    b. yakai ku paisu ki takanaw
       exist NOM money OBL Takanaw
       'Takanaw has money.'

251

(24) Mantauran Rukai

    a. omiki   paiso-li
       exist   money-my
       'I have money.'

    b. omiki   paiso-ni   taotao
       exist   money-his Taotao
       'Taotao has money.'

    c. omiki   taotao   paiso-ni
       exist   Taotao   money-his
       'Taotao has money.'

Mantauran Rukai exhibits another structure, one that is exemplified in (25), where the possessor is marked by the oblique case.[13] Though unmarked for case, the theme must be treated as the subject of the sentence. This claim is based on coreference restrictions illustrated in (25b) and (25d, e), which show that the oblique pronoun occurring on the verb does not corefer with the possessed entity.

(25) Mantauran Rukai

    a. omik-iaə   [paiso]$_{NOM}$
       exist-1S.OBL money
       'I have money.'

    b. *omik-iaə   paiso-li
        exist-1S.OBL money-my

    c. omik-inə   paiso
       exist-3S.OBL money
       'He has money.'

    d. omik-inə$_i$ [paiso-ni   taotao]$_j$
       exist-3S.OBL money-his Taotao
       'He has Taotao's money.'

---

13. While the first structure 'Y's X exists' is existential in nature, the second parallels the locative construction. Compare Mantauran Rukai (i) and (ii).

  (i)   a. omiki paiso   ana
         exist money   there
         'There is money there.'

       b. omiki paiso-li
         exist money-my
         'I have money.' (lit. 'My money exists.')

 (ii) a. omik-iaə   paiso
         exist-1S.OBL money
         'I have money.' (lit. 'The money exists to me.')

       b. omik-iaə   varaŋə-li ḷolai ʔapəcə
         exist-1S.OBL belly-my child sleep
         'The child is sleeping on my belly.'

e. omik-inə$_i$ ðipolo$_i$ [paiso-ni taotao]$_j$
   exist-3S.OBL Dipolo money-his Taotao
   'Dipolo has Taotao's money.'

At this stage, a word should be said about Truku Seediq, which behaves a bit differently from Paran Seediq. I have attempted to show in section 3.1. that in Paran Seediq, the verb *niqan* should be analyzed as *niq-an,* a verb marked as LF that takes two nominal arguments. It is quite clear that in Truku Seediq a process of grammaticalization is under way: (i) the verb is a one-place predicate; (ii) if the possessor is indicated by means of a noun, it is marked as nominative and corefers with the suffix attached on the theme NP. This grammaticalization process is thoroughly discussed in Tsukida (1999), and I follow her analysis in this study.

(26) Truku Seediq (adapted from Tsukida, 1999)
    a. niqan laqi-mu
       exist child-my
       'I have a child.'

    b. *?niqan-ku   laqi[14]
       [exist-1S.NOM child]

    c. niqan laqi-na ka   awi
       exist child-his NOM Awi
       'Awi has a child.'

Abstracting all the variation pointed out above, we can schematize the second structure as (27a, b).

(27) a. Amis, Atayal, Bunun, Kavalan, Paiwan, Puyuma, Rukai, Truku Seediq
        exists [X   Y]$_{NOM}$ (lit. Y's X exists)

    b. Mantauran Rukai
        exists [Y]$_{OBL}$ [X]$_{NOM}$ (lit. X exists to Y)

3.3 Y Has X vs. Y's X Exists

In this section, I attempt to determine the semantic variation associated with the two structures 'Y has X' and 'Y's X exists' in the three languages (Bunun, Kavalan, and Puyuma) where they are both found, and finally, to compare them with Mantauran Rukai.

In the former three languages, the same kind of semantic variation is found where both *have*-structures occur: the first structure 'Y has X' indicates inalienable or inherent possession; the second structure 'Y's X exists' refers to the existence of the possessed entity. Compare (28–29).

---

14. Tsukida (1999) indicates that this example is totally ungrammatical. One Seediq informant in his early sixties whom I consulted believes that it is grammatically correct when it yields the following interpretation(s):'I am pregnant' or 'I belong to a(nother) man' (implied: having born his child). Similar examples with inanimate themes are, on the other hand, rejected.

(28) 'Y has X'
    a. Isbukun Bunun
       ?aiða-an-ik   ?uvað
       exist-LF-IS.NOM child
       'I have a child.' (implied: of my own)

    b. Hsinshe Kavalan
       yau-iku   tu   sunis
       exist-IS.NOM ACC child
       'I have a child.' (implied: of my own)

    c. Nanwang Puyuma
       unian-ku   ɖa   paisu
       not exist-IS.NOM OBL money
       'I have no money.' (implied: of my own)

(29) 'Y's X exists'
    a. Isbukun Bunun
       ?aiða? ?inak   ?uvað
       exist   my   child
       'I have a child.' (lit. 'My child exists.')

    b. Hsinshe Kavalan
       yau   sunis-ku
       exist child-my
       'I have a child.' (lit. 'My children exist.' [may not be my own child])

    c. Nanwang Puyuma
       ulaya ku-paisu
       exist my-money
       'I have money.' (lit. 'My money exists.')

Interestingly enough, the syntactic variation found in Mantauran Rukai (cf. 'Y's X exists' vs. 'X exists at Y's place') yields the same semantic difference as that found in languages that display the two *have*-structures 'Y has X' and 'Y's X exists'. As in the other languages, 'Y's X exists' refers to the existence of the possessed entity and usually implies its presence at speech time. As for the second structure 'X exists at Y's place', it can be regarded as semantically equivalent to 'Y has X' because it indicates inalienable or inherent possession.

(30) Mantauran Rukai
    a. omiki paiso-li
       exist   money-my
       'I have money.' (implied: 'My money exists.')

    b. omik-iaə   paiso
       exist-IS.OBL money
       'I have money.' (implied: 'My money is here in my pocket.')

4. Conclusion

I have tried to show in this essay that the Formosan languages display either one or two *have*-structures. This has been achieved by a thorough investigation of the marking of the verb and of the nominal arguments, and in that respect, Table 1—based in part on Yeh et al. (1998) and Lin (1996)—may aid in visualizing the main syntactic variation that divides these languages.

Table 1. Verbal Marking and Case Realization of Nominal Arguments in the Two *have*-structures Found in Formosan Languages*

LANGUAGE	DIALECT	STRUCTURE	MARKING OF THE VERB	MARKING OF THE TWO NPS	
				THEME	LOCATIVE PHRASE
Seediq	Truku	Y's X exists	unmarked	GEN	NOM
Seediq	Paran	Y has X	marked by LF -*an*	NOM	ACC
Bunun	Isbukun	Y has X	marked by LF -*an*	OBL	NOM
		Y's X exists	unmarked	NOM	GEN
Saisiyat	Tungho	Y has X	unmarked	ACC	NOM
Kavalan	Hsinshe	Y has X	unmarked	ACC	NOM
		Y's X exists	unmarked	NOM	GEN or OBL
Puyuma	Nanwang	Y's X exists	unmarked	NOM	OBL
Amis	Central	Y's X exists	unmarked	NOM	GEN
Atayal	Wufeng	Y's X exists	unmarked	NOM	GEN
Paiwan	Northern	Y's X exists	unmarked	NOM	GEN
Rukai	Budai	Y's X exists	unmarked	NOM	OBL
Rukai	Mantauran	X exists to Y	marked by *om-*	NOM	OBL
		Y's X exists	marked by *om-*	NOM	GEN

*   Table 1 also appears in a slightly different form in Zeitoun et al. 1999:29.

I have said nothing about Tsou, because in many respects, it does not fit into the typological picture provided above. In the early sixties, Tung (1964) pointed out the difficulty in accounting for the existential/possessive construction, and this problem, to my knowledge, has not been resolved. To conclude this essay, I would like to repeat the two of the examples given in the introduction, to call attention now in retrospect to the ways in which Tsou departs from what is found in the Formosan languages as a whole.

(31)   Tfuya Tsou
     a. pan   to   peisu-si
       exist OBL money-his
       'I have money.'

b. pan to   peisu ta   moʔo
   exist OBL   money OBL Mo'o
   'Mo'o has money.'

Possessive sentences differ from other types of verbal/nominal sentences in this language in at least two respects. (i) Verbal sentences are (nearly) always introduced by an auxiliary verb, whereas possessive sentences are headed by *pan* 'exist', which does not belong to the class of auxiliary verbs[15] reported in this language. (ii) Verbal and nominal clauses usually include a predicate followed by at least one NP, marked as nominative, but in possessive sentences, *pan* is followed by an oblique NP. That is, the existential sentence is subjectless. Thus, we find quite a different pattern from that found in the two types of languages described above, where either the locative phrase/possessor or the theme/possessed entity is marked as nominative and thus behaves as the subject of the sentence.

Though I am leaving the question open as to how these examples should be integrated into the typological perspective presented here, an increasing number of linguists are now working on Tsou—Stanley Starosta was among the first to get interested in this language—and I am sure that no one will remain insensitive to this question.

---

15. Auxiliary verbs include *mo, mi-, moso, moh-, i-, o(h)-, te, nte, ta, tena, ntena, da,* and *nto(h)-,* and usually carry not only focus but also temporal/aspectual distinctions. For a detailed analysis, see Tung et al. (1964), Szakos (1994), and Zeitoun (1996). Joseph Szakos (pers. comm.) believes that *pan* should be treated as a verb, because it can, for instance, be causativized, e.g., *poa-pan* . . . 'make (it) exist (come into existence) . . .'.

REFERENCES

Bresnan, Joan, and J. Kanerva. 1989. Locative inversion in Chichewa: A case study of factorization in grammar. *Linguistic Inquiry* 20:1–50.

Chang, Yung-li. 1997a. An outline of Seediq grammar [In Chinese]. Ms.

————. 1997b. An outline of Kavalan grammar [In Chinese]. Ms.

————. 1997c. Voice, case, and agreement in Seediq and Kavalan. Ph.D. dissertation, National Tsing-hua University.

Clark, Eve. 1978. Locationals: Existential, locative, and possessive constructions. In *Universals of human language*, vol. 4, *Syntax*, ed. by Joseph Greenberg, 85–126. Stanford: Stanford University Press.

Freeze, Ray. 1991. Existentials in Austronesian. Paper read at the Sixth International Conference on Austronesian Linguistics. Honolulu, May 21–24.

———— →. 1992. Existentials and other locatives. *Language* 68:553–595.

Kuno, Susumo. 1971. The position of locatives in existential sentences. *Linguistic Inquiry* 2: 333–378.

Huang, Lillian M. 1993. *A study of Atayal syntax.* Taipei: Crane Publishing.

Huumo, T. 1996. Bound spaces and the semantic interpretation of existentials. *Linguistics* 34:295–328.

Lee, Pei-jung. 1997. The case marking and focus systems in Kavalan. M.A. thesis, National Tsing-hua University.

Li, Paul. 1973. *Rukai structure.* Special Publications No. 64. Taipei: Institute of History and Philology, Academia Sinica.

————. 1997. *The Formosan tribes and languages of I-lan county* (In Chinese). I-lan Historical Series, 1. I-lan: I-lan county government.

Lin, Hsiu-hsu. 1996. Isbukun syntax. Ms.

Lyons, John. 1967. A note on possessive, existential, and locative sentences. *Foundations of language* 3:390–396.

→ Nichols, Johanna. 1986. Head-marking and dependent-marking grammar. *Language* 62:56–119.

Szakos, Joseph. 1994. Die Sprache der Cou: Untersuchungen zur Synchronie einer austronesischen Sprache auf Taiwan. Ph.D. dissertation, University of Bonn.

Tan, Cindy R. 1997. Simple sentences in Puyuma. M.A. thesis, National Normal Taiwan University.

Tsukida, Naomi. 1999. Locative, existential, and possessive clauses in Seediq. In *Selected Papers from the Eighth International Conference on Austronesian Linguistics*, ed. by Elizabeth Zeitoun and Paul Jen-kuei Li, 599–636. Symposium Series of the Institute of Linguistics, Academia Sinica, No. 1. Taipei: Institute of Linguistics (Preparatory Office), Academia Sinica.

Tung, T'ung-ho, S. H. Wang, T. K. Kuan, T. F. Cheng, and Margaret M. Yan. 1964. *A descriptive study of the Tsou language, Formosa.* Institute of History and Philology, Special Publications No. 48. Taipei: Institute of History and Philology, Academia Sinica.

Yeh, Marie M. 1998. The relationship between existential and possessive sentences in Saisiyat. Paper read at the First International Conference on Languages in Taiwan and Language Acquisition, National Hsinchu University, Hsinchu, May 31–June 1.

Yeh, Marie M., Lillian M. Huang, Elizabeth Zeitoun, Anna H. Chang, and Joy J. Wu. 1998. A preliminary study on the negative constructions in some Formosan languages. In *Proceedings of the Second International Symposium on Languages in Taiwan*, ed. by Huang Shuan-fan, 79–110. Taipei: Crane Publishing.

Zeitoun, Elizabeth. 1996. The Tsou temporal, aspectual, and modal system revisited. *Bulletin of History and Philology* 67.3:503–532. Taipei: Academia Sinica.

→ Zeitoun, Elizabeth, Lillian M. Huang, Marie M. Yeh, and Anna H. Chang. 1999. Existential, possessive, and locative constructions in the Formosan languages. *Oceanic Linguistics* 38:1–42.

16: THE SYNTAX AND SEMANTICS OF SAISIYAT NEGATORS

Marie Meili Yeh
National Lien Ho Institute of Technology

1. Introduction

Saisiyat is a Formosan language spoken by the aboriginal people dwelling in the northern mountain areas in Hsinchu and Miaoli counties. Like many other Formosan languages, Saisiyat employs a variety of forms to express negation. Yeh (1991) identifies the eight negators in table 1 based on her fieldwork.[1] This study will be devoted to the discussion of the relation between *ʔokaʔ* and the four negators *ʔokay*, *ʔokik*, *ʔamkay*, and *ʔamkik*, because the five bear morphological and syntactic resemblance. Examining the morphological and syntactic properties of these negators, we find that *ʔamkik* and *ʔamkay* are contracted forms of *ʔam ʔokik* and *ʔam ʔokay*, respectively, and *ʔam* is a modal verb. In addition, *ʔokik* and *ʔokay* may come from the reduction of *ʔokaʔ* plus one of the two morphemes *ʔi* or *ʔik*, which serve to introduce the negated element. This analysis is in accord with Starosta's (1995:698) remark that "in Proto-Formosan, a LIGATURE preposition *ka* or *a* introduced clauses following a negative auxiliary verb, and probably other complement clauses as well." Data from other Formosan languages will be presented to show that this phenomenon is observed not only in Saisiyat.

Table 1. Negators in Saisiyat

NEGATOR	DISTRIBUTION AND FUNCTION
ʔokik	before nonverbal predicates or stative verbs in declarative constructions
ʔokay	before action verbs in declarative constructions
ʔamkik	before nonverbal predicates or stative verbs to express the negation of modality in declarative constructions
ʔamkay	before action verbs in declarative constructions, with modal implication
kayniʔ	before NP arguments or predicates to express the negation of volition
ʔokaʔ	before NP arguments in possessive/existential constructions; before predicates
ʔiziʔ	before verbs in imperative constructions
ʔinʔiniʔ	before predicates or adverbials, denoting 'not yet'

---

1. In this study, the term negator is used in a broad sense to cover all the negative forms. Abbreviations used include 1, 1st person; 2, 2nd person; 3, 3rd person; AF, Agent Focus; ACC, accusative; ASP, aspect; BEN, benefactive; CAUS, causative; EXCL, exclusive; GEN, genitive; INCL, inclusive; IF, Instrumental Focus; LF, Locative Focus; LNK, linker; LOC, locative; NEG, negator; NOM, nominative; PL, plural; PF, Patient Focus; POSS, possessive; Q, Question; REAL, realis; SG, singular; TNS, tense.

2. The Negators *ʔokik* and *ʔokay*

The two negators *ʔokik* and *ʔokay* are both used in declarative constructions. With respect to the syntactic distribution, the negator *ʔokay* is more restricted; it can only cooccur with action verbs. Note that the focus marking in the affirmative and negative constructions is different. Compare:

(1) a. ʔoyaʔ ʃ-om-əβət ka   korkoriŋ
       mother beat-AF   ACC child
       'Mother beat the child.'[2]

   b. ʔoyaʔ ʔokay ʃəβət   ka korkoriŋ
       mother NEG   AF-beat ACC child
       'Mother did not beat the child.'

(2) a. korkoriŋ ʃəβət-ən ni   ʔoyaʔ
       child   beat-PF GEN mother
       'The child was beaten by Mother.'

   b. korkoriŋ ʔokay ʃəβət-i ni   ʔoyaʔ
       child   NEG   beat-PF GEN mother
       'The child was not beaten by Mother.'

(3) a. kahœy si-ʃəβət ni   βakiʔ   ka korkoriŋ
       wood IF-beat GEN grandfather ACC child
       'The wood was taken by Grandfather to beat the child.'

   b. kahœy ʔokay ʃəβət-ani ni   βakiʔ   ka korkoriŋ
       wood NEG   beat-IF   GEN grandfather ACC child
       'The wood was not taken by Grandfather to beat the child.'

As illustrated above, there exist two sets of focus markers in Saisiyat, which are summarized in table 2.[3] While the focus markers in set I appear in affirmative declarative constructions, those in set II are found in imperative constructions[4] or negative constructions after *ʔokay* as well as *ʔinʔiniʔ* and *ʔiziʔ*, the two negators that will be introduced in section 5.

Table 2. Focus Markers in Saisiyat

	AF	PF	LF	IF
I	m-, -om-, ma-, Ø	-ən	-an	si-
II	Ø	-i	—	-ani

---

2. As is pointed out in Yeh (1991), verbs with focus markers only are neutral forms. They do not carry special tense distinction. In this study, translation is normally in the past tense unless a habitual reading is implied.

3. In fact, most Formosan languages possess two sets of focus markers. For details on the distribution of the two sets of focus markers, see Yeh (1997a).

4. However, PF imperative verbs are not suffixed with *-i*. Instead, they are bare root forms. Therefore, the distinction between the two sets of focus markers cannot be captured by the feature [imperative].

(4) Imperative Constructions

    a. ʃəβət   ka   korkoriŋ
       AF-beat ACC child
       'Beat the child!'

    b. tatpoʔ βəːay-ani ka   korkoriŋ
       hat   give-IF ACC child
       'Give the hat to the child!'

(5) Negative Constructions

    a. βakiʔ   ʔinʔiniʔ ʔi waːiʔ
       grandpa NEG   AF-come
       'Grandpa has not come yet.'

    b. ʔiziʔ ʃəβət ka   korkoriŋ
       NEG beat ACC child
       'Don't beat the child!'

Starosta, Pawley, and Reid (1982) call the focus markers that occur in imperative and certain negative constructions DEPENDENT FORMS because they occur only in imperative constructions or in embedded sentences. Following their work, the terms dependent and independent focus markers will be adopted in this study. To sum up briefly, *ʔokay* is followed by action verbs with dependent focus markers.

Compared with *ʔokay*, the negator *ʔokik* is more widely used. It is generally found to negate nonverbal predicates (6 and 7) or verbal predicates with stative verbs (8 and 9).[5]

(6) a. yako   ʃayʃiyat
       1SG.NOM Saisiyat
       'I am Saisiyat.'

    b. yako   ʔokik ʃayʃiyat
       1SG.NOM NEG   Saisiyat
       'I am not Saisiyat.'

(7) a. sia   ray tawʔan
       3SG.NOM LOC house
       'He is at home.'

    b. sia   ʔokik ray tawʔan
       3SG.NOM NEG   LOC house
       'He is not at home.'

(8) a. yako   raːam hisia
       1SG.NOM know 3SG.ACC
       'I know him.'

    b. yako   ʔokik raːam hisia
       1SG.NOM NEG   know 3SG.ACC
       'I do not know him.'

---

5. Sentence (9b) shows that *ʔokik* cooccurs with an adjective. Adjectives are classified as stative verbs, because they share the morphosyntactic properties of verbs (Yeh 1991:67–69).

(9) a. minkoriŋan kərpə:
   woman   fat
   'The woman is fat.'

 b. minkoriŋan ʔokik kərpə:
   woman   NEG fat
   'The woman is not fat.'

The sentences below show that, like ʔokay, the negator ʔokik also cooccurs with action verbs, though the sentences negated by the two negators are different syntactically and semantically.

(10) a. ʔoyaʔ   maŋ ʃ-om-əβət ka   korkoriŋ
    mother ASP beat-AF- ACC child
    'Mother is beating the child.'

 b. ʔoyaʔ   ʔokik maŋ ʃ-om-əβət ka   korkoriŋ
   mother NEG ASP beat-AF- ACC child
   'Mother is not beating the child.'

(11) a. sia   r-om-in-aʔœ: ka   pinoβæ:æh
    3SG.NOM drink-AF-ASP- ACC wine
    'He has drunk wine.'

 b. sia   ʔokik r-om-in-aʔœ: ka   pinoβæ:æh
   3SG.NOM NEG   drink-AF-ASP ACC wine
   'He has never drunk wine.'

Syntactically, the action verbs after ʔokik and ʔokay are different in their focus marking. Comparing (10) and (11) with (12) and (13), we find that while the verbs after ʔokay change their focus markers to the dependent forms, the focus marking of the verb after ʔokik remains unchanged.

(12) a. ʔoyaʔ   ʃ-om-əβət ka   korkoriŋ
    mother beat-AF- ACC child
    'Mother beat the child.'

 b. ʔoyaʔ   ʔokay ʃəβət   ka   korkoriŋ
   mother NEG AF-beat ACC child
   'Mother did not beat the child.'

(13) a. sia   r-om-in-aʔœ: ka   pinoβæ:æh
    3SG.NOM drink-AF-ASP- ACC wine
    'He has drunk wine.'

 b. sia   ʔokay raʔœ: ka   pinoβæ:æh
   3SG.NOM NEG AF-drink ACC wine
   'He did not drink wine.'

Moreover, they are different with respect to the presence of temporal or aspectual marking. Note that the action verbs negated by ʔokik are in the progressive or perfective aspect, while those following ʔokay do not carry any tense or aspectual marker.

261

As for the semantic difference, (10b) and (11b) express the negation of an ongoing event or an experience, while (12b) and (13b) deny the occurrence of an event. Even when the aspectual marker is absent (14a), the sentence with an action verb negated by ʔokik will still be interpreted as the negation of an ongoing event. However, (14b) is interpreted as the negation of the occurrence of the event *writing*. Compare:

(14) a. korkoriŋ ʔokik k-om-aːat ka ʔinsiaa raroːoʔ
      child   NEG write-AF- ACC 3SG.POSS name
      'The child is not writing his name.'

    b. korkoriŋ ʔokay kaːat ka ʔinsiaa raroːoʔ
      child   NEG write ACC 3SG.POSS name
      'The child did not write his name.'

A similar distinction is also observed in Mayrinax Atayal between *iniʔ* and *yakaat* and in Wulai Atayal between *iniʔ* and *yat* (Huang 1993:74-87, Huang 1995:160-170, and Huang and Davis 1989). Take the following Wulai sentences for example.

(15) Wulai Atayal (Huang 1993:75)

a. iniʔ-ku   qaniq   b. yat-ku   qaniq
   NEG-1SG.NOM eat   NEG-1SG.NOM eat
   'I didn't eat.'   'I won't eat.'

Huang (1993) as well as Huang and Davis (1989) remark that the difference between the two negators *iniʔ* and *yat* lies in how they relate to the event. As (15a) indicates, *iniʔ* directly negates the execution of the event while *yat* denies an opportunity or preparatory circumstance that leads to the occurrence of an event, as in (15b). Huang (1993:81) further remarks that this pattern suggests that *yat* will appear with predicates that are more stative or predicates that identify an object, while *iniʔ* will cooccur with action verbs. The Saisiyat negator *ʔokik* behaves like *yat* in Wulai Atayal, and *ʔokay* behaves like *iniʔ*.

Given that the negator *ʔokik* cooccurs with nonverbal predicates, stative verbs, and actions verbs in the progressive or perfective aspect, we might wonder if the three share certain properties. Nonverbal predicates, which serve to describe properties of some type, are doubtless stative. However, are action verbs in the progressive or perfective aspect stative as well? Tang (1988) remarks that English progressive and perfective verbs appear to be stative because they can appear after *seem to*, and only stative verbs can appear after *seem to*.[6] Compare:

(16) a. John seems to know the answer.
    b. * John seems to run.
    c. He seems to be running.
    d. He seems to have been running.

The following sentences show that in Saisiyat, only stative verbs, perfective verbs, or progressive verbs can appear after *nakhalaʔ* 'seem to'. Thus, the three seem to form a natural class.

---

6. The situation seems to be parallel in Mandarin Chinese (Yeh 1991:88).

(17)   ʔoyaʔ   nakhalaʔ raːam ka   hiniʔ ʔowaw
        mother seem   know ACC this   matter
        'Mother seems to know this matter.'

(18) a. ʔoyaʔ   nakhalaʔ ʔam βahiʔ ka   kayβaːən[7]
           mother seem   ASP wash ACC clothes
           'Mother seems to be washing the clothes.'

        b. *ʔoyaʔ nakhalaʔ βahiʔ ka kayβaːən

(19) a. βakiʔ   nakhalaʔ r-om-in-aʔœː ka pinoβæːæh
           grandpa seem   drink-AF-ASP- ACC wine
           'Grandpa seems to have drunk wine.'

        b. *βakiʔ nakhalaʔ r-om-aʔœː ka pinoβæːæh

For stative and nonstative predicates to make use of different negators is not idiosyncratic in Saisiyat. According to Chen (1996), the two negators *ini* and *uxe* in Seediq can also be differentiated by the semantic properties of the predicates. Thus, stative verbs must have their aspect changed when negated by *ini*, the negator that cooccurs with action verbs.

(20) Seediq (Chen 1996:63)
        a. msipus riyuŋ mudu   nii
           sweet   very   tangerine this
           'This tangerine is sweet.'

        b. uxe   msipus mudu   nii
           NEG   sweet   tangerine this
           'This tangerine is not sweet.'

        c. ini   ku-sipus mudu   nii
           NEG   -sweet tangerine this
           'This tangerine will not be sweet.' / 'This tangerine is not sweet.'

3. The Negator *ʔokaʔ*

The negator *ʔokaʔ* is mostly found before a nominal argument in existential or possessive constructions as the negative counterpart of the possessive/existential verb *hayzaʔ* 'have'.

(21) a. yako   hayzaʔ ka   rayhil
           ISG.NOM have   ACC money
           'I have money.'

        b. yako   ʔokaʔ ka   rayhil
           ISG.NOM NEG   ACC money
           'I have no money.'

---

7. Except in the future tense, *ʔam* can also express progressive aspect.

(22) a. kawaʃ βaβaw hayzaʔ ka βintœʔæn
     sky   above have   ACC star
     'There are stars in the sky.'

   b. kawaʃ βaβaw ʔokaʔ ka βintœʔæn
     sky   above NEG   ACC star
     'There are no stars in the sky.'

The case realization of the nominals in these sentences indicates that *hayzaʔ* and *ʔokaʔ* are transitive verbs taking the Theme argument as their grammatical object and the Possessor or the Location as subject. The fact that *ʔokaʔ* can have tense or aspect marking (23) and can be focus-marked[8] or causativized (24) further supports its being analyzed as a verb. Consider:

(23) a. hiniʔ ka   βa:laʔ mayhal ʔam ʔokaʔ ka ʔalaw[9]
     this NOM river after will NEG   ACC fish
     'Later, this river will not have any fish.'

   b. hiniʔ ka   ʔomæh ʔ-in-okaʔ ka tawʔan, ʔisahiniʔ hayzæʔ ila
     this NOM field NEG-ASP- ACC house   present have   ASP
     'This field used not to have any houses, but now it does.'

   c. yako   kin ʔokaʔ ka rayhil
     ISG.NOM ASP NEG   ACC money
     'I have less and less money.'

(24) a. hizaʔ korkoriŋ ʔampowaʔ si-ʔokaʔ ka yaβaʔ
     that child   why   IF-NEG   ACC father
     'Why does that child not have a father?'

   b. pa-ʔokaʔ ka raro:oʔ
     CAUS-NEG ACC name
     'Remove the name.'

In addition to appearing before NP arguments in an existential or a possessive construction, *ʔokaʔ* is also found to coocur with a predicate (Yeh 1991:93). It should also be noted that sentences whose predicate is negated by *ʔokaʔ* have the morpheme *ila* or *na* intervening between *ʔokaʔ* and the predicate. Besides, the morpheme *ʔik* (for stative verbs) or *ʔi* (for action verbs) occur before the negated predicate.[10]

(25) a. yako   ʔam ʔokaʔ ila ʔi ʃəβət ka korkoriŋ
     ISG.NOM will NEG   ASP AF-beat ACC child
     'I will not beat the child any longer.'

---

8. However, this usage is increasingly rare, according to the informant.

9. As illustrated in this sentence, *ka* can mark both Nominative and Accusative case in Saisiyat. For a detailed account of the Saisiyat case-marking system, please refer to Yeh (1995a).

10. This is true of most of the data. There are also sentences with morphemes other than *ila* and *na* intervening between *ʔokaʔ* and *ʔi*.

b. t-in-awʔan ʔæhæʔ ila tinalʔomæh, ʔokaʔ na ʔik sizæh
build.house-ASP- one ASP year NEG ASP finish
'It has been built for one year, but is still not finished.'

According to Yeh (1991), *ila* and *na* can also appear in sentence-final position. When this happens, the negator *ʔokay* or *ʔokik* is used instead, as in (26). The similarity between (25) and (26) indicates that *ʔokik* and *ʔokay* might be the reduced forms of *ʔokaʔ ʔik* and *ʔokaʔ ʔi*, respectively. In other words, the negator for declarative constructions is actually *ʔokaʔ*, and *ʔik* or *ʔi* serves to introduce the element being negated.

(26) a. yako ʔam ʔokay ʃəβət ka korkoriŋ ila
ISG.NOM will NEG AF-beat ACC child ASP
'I will not beat the child any longer.'

b. t-in-awʔan ʔæhæʔ ila tinalʔomæh, ʔokik sizæh na
build.house-ASP- one ASP year NEG finish ASP
'The house has been being built for one year [now], but is still not finished.'

This analysis is further supported by the fact that when giving a short answer to (27a), *ʔam ʔokaʔ* is used instead of *ʔamkay*. Compare (27c) with (27a–b).

(27) a. rimʔan ʔam ʔ-omo-ral ay
tomorrow will rain-AF- Q
'Will it rain tomorrow?'

b. ʔam ʔokaʔ
will NEG
'It won't.'

c. rimʔan ʔamkay ʔoral
tomorrow NEG AF-rain
'It will not rain tomorrow.'

Moreover, the existence of *ʔi* and *ʔik* is also attested with another negator *ʔinʔiniʔ* (see section 5). Consider:

(28) a. tabər ʔinʔiniʔ ʔi waːiʔ
Tabər NEG AF-come
'Tabər has not yet come.'

b. yako m-waːiʔ ʔinʔiniʔ ʔik ʔæhæʔ tiyamsoŋ
ISG.NOM AF-come NEG one hour
'I have arrived here not yet for an hour.'

Here we come to the following paradigm of the negator *ʔokaʔ*. According to this paradigm, *ʔokaʔ* can be followed by different types of complements: while the noun phrase functioning as its argument is introduced by the accusative case marker *ka*, the complement serving as the predicate of the sentence must be preceded either by *ʔik* (if it is stative) or by *ʔi* (if nonstative).

(29)　ʔokaʔ　+　ka　+　argument
　　　　ʔokaʔ　+　ʔik　+　predicate [+stative]
　　　　ʔokaʔ　+　ʔi　+　predicate [-stative]

This phenomenon is by no means idiosyncratic in Saisiyat. In many other Formosan languages, a morpheme also intervenes between the negator and the negated element. Take Paiwan for example. Yeh et al. (1998) claim that *inika* is the negator used in declarative constructions.

(30)　Paiwan (Yeh et al. 1998)
　　　a.　na-k-əm-ac　　a　　ʔatuvi tua vatu
　　　　　PAST-bite-AF- NOM snake ACC dog
　　　　　'The snake bit a dog.'

　　　b.　ini-ka na-k-əm-ac　　tua　ŋiaw aza vatu
　　　　　NEG PAST-bite-AF- ACC cat　　that dog
　　　　　'That dog did not bite the cat.'

Nevertheless, given that the aspectual marker *aŋa* can separate *ini* and *ka* as in (31), the negator should be *ini*,[11] and *ka* should be treated as the morpheme introducing the negated element.

(31)　Paiwan
　　　a.　ini-aŋa　ka t-əm-əkəl̪　ti　　kama tua vuvu
　　　　　NEG-ASP　　drink-AF- NOM father ACC wine
　　　　　'Father has not yet drunk the wine.'

　　　b.　ini-aŋa　ka-kən　　a　　　madudu
　　　　　NEG-ASP -1SG.NOM LNK　angry
　　　　　'I am not angry yet.'

The fact that yes-no questions are answered with *ini* alone further supports the position that it is a negator per se.

(32)　Paiwan
　　　Q: sə-payuan-sun　　　　　　A: ini
　　　　sə-Paiwan-2SG.NOM　　　　　NEG
　　　　'Are you Paiwan?'　　　　　'No.'

In addition, Yeh et al. (1998) also point out that in Mantauran Rukai, the morpheme *ka* appears before the negated element and claim the existence of double negation in Mantauran Rukai.

(33)　Mantauran Rukai
　　　a.　o-tipitipi taol̪o　　　　　b.　o-tipitipi-kai　　ka taol̪o
　　　　　REAL-beat dog　　　　　　　REAL-beat-NEG　　dog
　　　　　'He beat the dog.'　　　　　'He did not beat the dog.'

---

11. The data in the other languages are from the field notes of Anna Chang (Paiwan), Joy Wu (Amis), and Elizabeth Zeitoun (Mantauran Rukai).

Moreover, in Amis and Seediq, a morpheme (*ka* in Amis and *ku* in Seediq) is found to appear before the negated predicate, but only when the predicate is stative. Take for example the following sentences.[12]

(34) Amis (Yeh et al. 1998)

   a. ma-fanaʔ kaku ciŋraan
      AF-know 1SG.NOM 3SG.ACC
      'I know him.'

   b. caʔay ka-fanaʔ kaku ciŋraan
      NEG   -know 1SG.NOM 3SG.ACC
      'I don't know him.'

(35) Seediq (Chen 1996:47)

   a. pbaro riyuŋ qhuni nii
      tall   very tree this
      'This tree is very tall.

   b. ini ku-pbaro qhuni nii
      NEG -tall   tree this
      'This tree is not tall.'

What then is the function of the morphemes *ʔi* and *ʔik* in Saisiyat, *ka* in Amis, Paiwan, and Mantauran Rukai, as well as *ku* in Seediq? Starosta (1995:698) points out that Proto-Formosan employed a ligature preposition *ka* or *a* to introduce clauses following a negative auxiliary verb, and that the morphemes in these languages seem to be ligatures.[13] Nevertheless, differences can be observed. While Saisiyat, Amis, and Seediq make use of different ligatures for stative and nonstative predicates (if we take zero as one type of ligature), Paiwan and Mantauran Rukai resort to the same ligature regardless of the type of predicate.

4. Negation and Modality

This section discusses the relation between negation and modality. To negate future events or possibility in Saisiyat, the negators *ʔamkik* and *ʔamkay* are employed. Consider the following two pairs of sentences:

(36) a. sia   rimʔan   ʔam ray tawʔan
       3SG.NOM tomorrow will LOC house
       'He will be home tomorrow.'

---

12. Note that in Wu (1995) and Chen (1996), *ku* and *ka*, respectively, are treated as verbal prefixes.

13. Chen (1996:46-51) follows Chang's (1992) analysis of Paiwan *ka* in causative constructions to treat Seediq *ku* and Saisiyat *-k* as activizers. Though *ka* in Amis, *ku* in Seediq, and *-k* in Saisiyat do show properties of an activizer, this analysis does not seem to hold for the two *ka*'s in Paiwan and Mantauran Rukai because they are not restricted to occurring before stative verbs.

    b. sia   rimʔan   ʔamkik ray tawʔan
       3SG.NOM tomorrow NEG   LOC house
       'He will not be home tomorrow.'

(37) a. rimʔan   ʔam ʔ-om-oral
       tomorrow will rain-AF-
       'It will rain tomorrow.'

    b. rimʔan   ʔamkay ʔoral
       tomorrow NEG   AF-rain
       'It will not rain tomorrow.'

The fact that (36b) can also be expressed as (38a) and (37b) as (38b) indicates that *ʔamkik* and *ʔamkay* are the reduced forms of *ʔam ʔokik* and *ʔam ʔokay* respectively.

(38) a. sia   rimʔan   ʔam ʔokik ray tawʔan
       3SG.NOM tomorrow will NEG   LOC house
       'He will not be home tomorrow.'

    b. rimʔan   ʔam ʔokay ʔoral
       tomorrow will NEG   AF-rain
       'It will not rain tomorrow.'

Yeh (1991, 1995b) points out that *ʔam* can denote both epistemic and deontic modality. However, *ʔamkik* and *ʔamkay* can only express the negation of epistemic modality. Negation of deontic modality is conveyed by another negator, *kayniʔ*.

(39) a. maʔan   ʔam rayhil
       1SG.GEN want money.
       'I want money.'

    b. yako   kayniʔ ka   rayhil
       1SG.NOM NEG   ACC money
       'I do not want money.'

(40) a. yako   ʔam m-paʔrəm
       1SG.NOM want AF-sleep
       'I want to sleep.'

    b. yako   kayniʔ m-paʔrəm
       1SG.NOM NEG   AF-sleep
       'I don't want to sleep.'

Example (39) shows that *kayniʔ* can take a noun phrase as its complement.[14] Also note that the verbs after *kayniʔ* are affixed with independent focus markers, whereas the verbs after the negator *ʔamkay* carry the dependent focus markers. Compare:

---

14. Note that the case realization in the affirmative construction and in the negative construction is different. In the affirmative construction, it follows the pattern of NAF constructions to realize the Agent argument as Genitive, whereas in the negative construction, it follows the pattern of AF constructions to realize the Agent argument as Nominative (see also Yeh 1995:375).

(41) a. yako   kayniʔ s-om-iʔæl ka ʔalaw
        ISG.NOM NEG   eat-AF- ACC fish
        'I do not want to eat the fish.'

    b. yako   ʔamkay siʔæl ka ʔalaw
        ISG.NOM NEG   AF-eat ACC fish
        'I will not eat the fish.'

(42) a. korkoriŋ kayniʔ ʃəβət-ən noka minkoriŋan
        child   NEG   beat-PF GEN woman
        'The child did not want to be beaten by the woman.'

    b. korkoriŋ ʔamkay ʃəβət-i noka minkoriŋan
        child   NEG   beat-PF GEN woman
        'The child will not be beaten by the woman.'

Yeh (1991:98–99) points out that the difference in the focus marking might be ascribed to the difference in the grammatical status of *kayniʔ* and *ʔamkay*. The fact that *kayniʔ* can be used alone as a transitive verb in a sentence indicates that it is a lexical verb. On the other hand, *ʔamkay* cannot be used alone and thus is more like an auxiliary verb.

5. The Negator *ʔiziʔ* and *ʔinʔiniʔ*

The negator *ʔiziʔ* is generally found in imperative constructions expressing prohibition or a negative command as exemplified in (43).

(43) a. ʔiziʔ saːəŋ   b. ʔiziʔ siʔæl ka   pazay
      NEG AF-sit   NEG AF-eat ACC rice
      'Don't sit down.'   'Don't eat the rice.'

However, its distribution is not restricted to imperative constructions. It can appear in declarative constructions (44) and can also cooccur with the third person subject (44 and 45b), though it is not allowed to cooccur with a first person singular subject, as (45c) indicates.

(44) a. korkoriŋ ʔiziʔ raʔœː   ka pinoβæːæh,
        child   NEG AF-drink ACC wine

        tatiniʔ   kayzæh r-om-aʔœː ka pinoβæːæh
        old.man good   drink-AF- ACC wine
        'Children can not drink; only adults can drink.'

    b. βakiʔ   ʔiziʔ palraŋaʔ ka βaβoy
        grandfather NEG AF-kill ACC pig
        'Grandfather cannot kill a pig.'

(45) a. ʃoʔo   ʔiziʔ saːəŋ
        2SG.NOM NEG AF-sit
        'You cannot sit down.'

    b. sia   ʔiziʔ saːəŋ
       3SG.NOM NEG AF-sit
       'He cannot sit down.'

    c. *yako   ʔiziʔ saːəŋ
       1SG.NOM NEG AF-sit

    d. yamiʔ/ʔitaʔ   ʔiziʔ saːəŋ
       1PL.EXCL.NOM/1PL.INCL.NOM NEG sit
       'We cannot sit down.'

For the first person singular subject, *ʔamkay* is used instead. This restriction may arise because the speaker is always the one who gives commands or makes prohibitions.

    (46)   yako   ʔamkay raʔœː ka pinoβæːæh
          1SG.NOM NEG   drink ACC wine
          'I cannot drink wine.'

The negator *ʔinʔiniʔ*, with its meaning close to 'not yet', can be used to negate a predicate (47) or an adverbial (48). Consider:

    (47) a. tabər ʔinʔiniʔ ʔi waːiʔ
         Tabər NEG   AF-come
         'Tabər has not yet come.'

        b. ʔalaw ʔinʔiniʔ ʔi talək-i
         fish NEG   cook-PF
         'The fish has not been cooked yet.'

    (48) a. ʃoʔo   rimaʔ ʔinʔiniʔ ʔæhæʔ tinalʔœmæh
         2SG.NOM go NEG   one   year
         'You have not been gone for a year yet.'

        b. yako   ʔinʔiniʔ oːnæhŋəː m-waːiʔ
         1SG.NOM NEG   long   AF-come
         'I have not been here for a long time yet.'

As mentioned in section 3, verbs after *ʔinʔiniʔ* have the morpheme *ʔi* appearing before them. Nonetheless, as indicated in (49), sometimes *ʔi* or *ʔik* does not occur.

    (49)   sia   ʔinʔiniʔ putŋor ray tawʔan
          3SG.NOM NEG   arrive LOC home
          'He has not yet arrived home.'

The following sentences seem to indicate that *ʔinʔiniʔ* is the negation of *ila*, an alternative expression of *ʔokaʔ na*.

    (50) a. ʃoʔo   rimaʔ ʔæhæʔ tinalʔœmæh ila
         2SG.NOM go   one   year   ASP
         'You've been gone for a year (now).'

b. ʃoʔo   rimaʔ ʔokaʔ na ʔæhæʔ tinalʔœmæh
2SG.NOM go NEG ASP one   year
'You haven't been gone for a year yet.'

c. ʃoʔo   rimaʔ ʔinʔiniʔ ʔæhæʔ tinalʔœmæh
2SG.NOM go NEG   one   year
'You haven't been gone for a year yet.'

6. Concluding Remarks

Table 3 gives the four negators in Saisiyat and their corresponding distribution. It indicates that the Saisiyat negators can be divided into two categories based on the mood of the construction. The negator *ʔiziʔ* is mainly used in imperative constructions, while the other three are found in declarative constructions. In addition, existential/possessive and nonexistential/possessive declarative constructions are negated by the same negator, though the morphemes that introduce the complements are different. In this respect, Saisiyat is different from many other Formosan languages where existential/possessive and nonexistential/possessive declarative constructions are negated by different negators (see Yeh et al. 1998 for details).

Table 3. Negators in Saisiyat

NEGATOR	CONSTRUCTION TYPE	DISTRIBUTION
ʔokaʔ	Declarative [+existential/possessive]	*ka* + argument
	Declarative [-existentialʔpossessive]	*ʔi* + predicate [-stative] *ʔik* + predicate [+stative]
kayniʔ	Declarative [+volition]	*ka* + argument, Ø + predicate
ʔinʔiniʔ	Declarative [not yet]	*(ʔi)* + predicate [-stative] *(ʔik)* + predicate [+stative], *(ʔik)* +adverbial
ʔiziʔ	Imperative	Ø + predicate

There are two sets of focus markers in Saisiyat. In some negative constructions, the verbs are affixed with the independent forms, while in others, the focus markers of the verb belong to the dependent forms. Table 4 summarizes focus marking in negative constructions.

Table 4. Focus Marking of the Verb in Negative Constructions

NEGATOR	LIGATURE	VERB FOCUS MARKING
ʔokaʔ	ʔi	+ dependent
	ʔik	− dependent
ʔiziʔ	Ø	+ dependent
kayniʔ	Ø	− dependent
ʔinʔiniʔ	ʔi	+ dependent

What factor determines the focus marking of the verb in negative constructions? Does the difference result from the difference in subcategories of the negative verbs or the difference in the cooccurring ligature? The fact that the same negator *ʔokaʔ* can be followed by verbs focused with both dependent and independent forms seems to indicate that it is the ligature that determines the focus marking of the verb—the ligature *ʔi* is followed by verbs focused with dependent forms while the ligature *ʔik* is followed by verbs affixed with independent forms. However, the picture appears to be different with the two negators *kayniʔ* and *ʔiziʔ*. The two both use zero ligature, but they are followed by verbs with different sets of focus markers—while the negator *kayniʔ* is followed by verbs with independent focus markers, the negator *ʔiziʔ* appears before verbs with dependent focus markers. It turns out that for these two negators, it is the negator that determines the focus marking of the verb. The negator *kayniʔ* is followed by verbs with independent focus markers because it is a lexical verb. As *ʔiziʔ* is an auxiliary verb, the verbs following it are affixed with dependent focus markers.

If we compare the focus marking in negative constructions with that in the corresponding affirmative constructions, we may find that when the ligature is *ʔik* or Ø, the focus marking of the verb remains the same as that in the affirmative construction. However, if the ligature is *ʔi*, the focus marker will be changed from the independent form to the dependent form. In other words, the ligature *ʔik* or Ø and *ʔi* exert different effects on their complements. Then, what is the grammatical status of the ligature? According to Schachter and Otanes (1972), a ligature functions as an indicator of the relation between the elements it connects. Since the ligatures *ʔik* and *ʔi* introduce clausal complements (Starosta 1995), they are similar to complementizers functionally. Tang (1997) provides strong evidence to argue for the Paiwan *tu* clause as finite and the *a* clause as nonfinite. Are Saisiyat *ʔi* and *ʔik* complementizers? Does the difference between *ʔi* and *ʔik* complements also lie in finiteness? If they are complementizers, why are they different from the other complementizers in complex constructions? For example, the Saisiyat complementizers for complex constructions are Ø and *komoʃaʔ* (Yeh 1997b), and in Paiwan, they are *tua, tu, ta,* and *a* (Tang 1997). To answer the above questions, further study on the intrinsic properties of the two sets of focus markers and more thorough research on complex constructions are necessary.

REFERENCES

Chang, Hsiou-chuan. 1992. Causative constructions in Paiwan. M.A. thesis, National Tsing Hua University.

Chen, Jie-huei. 1996. A preliminary study on Tkdaya Seediq negators (in Chinese). M.A. thesis, National Tsing Hua University.

Huang, Lillian M. 1993. *A study of Atayal syntax*. Taipei: Crane Publishing.

———. 1995. *A study of Mayrinax syntax*. Taipei: Crane Publishing.

Huang, Lillian M., and Philip W. Davis. 1989. Negation in Mandarin and Atayal: A comparison. Paper presented to the 22nd International Conference on Sino-Tibetan Linguistics and Languages, Honolulu, Hawai'i.

Payne, John R. 1985. Negation. In *Language typology and syntactic description*, vol. 1, *Clause structure*, ed. by Timothy Shopen, 197–242. Cambridge: Cambridge University Press.

Schachter, Paul, and Fe T. Otanes. 1972. *Tagalog reference grammar*. Berkeley: University of California Press.

Starosta, Stanley. 1995. A grammatical subgrouping of Formosan languages. In *Papers for International Symposium on Austronesian Studies Relating to Taiwan*, ed. by Paul J.-K. Li, Cheng-hwa Tsang, Ying-kuei Huang, Dah-an Ho, and Chiu-yu Tseng, 683–726. Taipei: Institute of History and Philology, Academia Sinica.

Starosta, Stanley, A. K. Pawley, and L. A. Reid. 1982. The evolution of focus in Austronesian. In *Papers from the Third International Conference on Austronesian Linguistics*, vol. 2, *Tracking the Travelers*, ed. by Amran Halim, Lois Carrington, and S. A. Wurm, 145–170. Pacific Linguistics C-75. Canberra: Australian National University.

Tang, Jane C.-C. 1997. On clausal complements in Paiwan. Paper presented to the Eighth International Conference on Austronesian Linguistics, Nankang, Taipei, Dec. 28–30.

Tang, Ting-chi. 1988. *A cognitive grammar of English: Form, meaning, and function*, 1 (in Chinese). Taipei: Student Book Company.

Wu, Joy J. 1995. Complex sentences in Amis. M.A. thesis, Taiwan Normal University.

Yeh, Marie M. 1991. Saisiyat structure. M.A. thesis, National Tsing Hua University.

———. 1995a. Focus and case marking system in Saisiyat. In *Papers from the First International Symposium on Languages in Taiwan*, ed. by Feng-fu Tsao and Mei Hui Tsai, 29–58. Taipei: Crane Publishing.

———. 1995b. A preliminary study on Saisiyat tense and aspect (in Chinese). In *Papers in Formosan linguistics*, ed. by Paul J.-K. Li and Ying-jin Lin, 369–384. Taipei: Committee of Education Study, Ministry of Education.

———. 1997a. The mapping between focus and the semantic role of the grammatical subject. In *Proceedings of the Conference on the Development of Languages in Taiwan*, 103–120. [Preconference issue for participants; conference held at National Hsinchu Teachers College, Hsinchu, Taiwan , June 6–7.]

———. 1997b. Pivotal constructions in some Formosan languages. In *Papers for the Eighth International Conference on Austronesian Linguistics*, 85–108. Taipei, Taiwan: Institute of Linguistics (Preparatory Office), Academia Sinica.

Yeh, Marie M., Lillian M. Huang, Elizabeth Zeitoun, Anna H. Chang, and Joy J. Wu. 1998. A preliminary study on the negative constructions in some Formosan languages. In *Papers from the Second International Symposium on Languages in Taiwan*, ed. by S.-F. Huang, 81–111. Taipei: Crane Publishing.

17:   SUBORDINATE CLAUSES AND ERGATIVE PATTERNS IN SHOSHONI

Francis Lindsey, Jr.
Kansai Gaidai University

1. Introduction[1]

Shoshoni is generally considered to be an accusative language, yet it has some features that suggest an ergative syntax. The most obvious of these is the number agreement between some common verbs and their objects (if the verb is transitive) or their subjects (if the verb is intransitive). Another is a type of subordinate clause in which the apparent subject occurs in the objective form.[2] In this paper, I examine number agreement and these subordinate clauses, along with other phenomena that may be useful in determining the extent to which an ergative syntax must be proposed for Shoshoni.

I argue that the apparent ergative number agreement for verbs is not a part of the syntax. Then I provide a Lexicase derivational rule that accounts for the object forms in subject roles in subordinate clauses.

2. Some Preliminary Information

Shoshoni is a Central Numic language of the Uto-Aztecan family. It is most closely related to the Western Numic (Mono, Northern Paiute, Paviotso, Bannock) and Southern Numic (Kawaiisu, Chemehuevi, Southern Paiute, Ute) languages. It is more distantly related to Hopi, Papago, Yaqui, Tarahumara, Nahuatl (including Classical Aztec), and a number of other languages in what is now the western U.S. and Mexico.

2.1 Shoshoni Phonology and Orthography

Shoshoni is a mora-timed language. There are five or six vowels,[3] each occurring long or short. The spelling system developed by Miller (1972) uses these symbols for the

---

2.   I have used the terms subjective and objective for case forms because Miller (1972) used them. They correspond to nominative and accusative, for the most part. However, the subjective form does not always function as a subject. It may be used with postpositions, for example.

3.   There is a question as to whether [e] is /e/ or an alternative pronunciation of the diphthong /ai/. The sound spelled ai is pronounced as either [e] or [ai] in many words, with some speakers preferring [e], other speakers preferring [ai], and others using both in apparent free variation. Miller (1972) used <u>ai</u> (with the underline) when he heard [e] and *ai* when he heard [ai].

vowels: i a̲i̲ a o u e, with doubled letters indicating the long vowels. The underlined a̲i̲ represents [e]. The orthographic *e* represents a high back unrounded vowel [ï]. The other symbols have the traditional values. Vowels are often unvoiced at the end of an utterance. They are regularly unvoiced before /h/.

The stop consonants occur in single, geminate, preaspirated, and prenasalized forms (using traditional terms for Numic languages). Consonants are lenis, voiceless, and unaspirated in initial position. In intervocalic position, single stops are pronounced as voiced spirants. Preaspirated stops (analyzed as clusters of /h + C/) are pronounced as voiceless fricatives. Prenasalized consonants (analyzed as /n + C/) are pronounced as a nasal followed by a voiced stop. The stop consonants are written as *p ts t k kw* in their single forms. The symbols have their normal values except that in the Gosiute dialect *ts* is pronounced as an interdental [t̪]. The nasals *m n* have distributions similar to the stops, although geminate and prenasalized forms are indistinguishable.[4] The intervocalic form of *n* varies a great deal in normal speech from the expected nasalized [ỹ] to an [n] phonetically similar to the geminate form.

The consonants *s w y* are not affected by preceding consonants. That is, they do not undergo spirantization, gemination, preaspiration, or prenasalization.

Shoshoni words are always pronounced with a final vowel, although that vowel is often unvoiced. However, because spirantization, gemination, preaspiration, and prenasalization operate across morpheme and word boundaries in normal speech, some words are assumed to have underlying final consonants that cause phonetic variation in the initial consonants of following words and suffixes.[5] The underlying consonants that cause changes in following consonants are written as " for gemination, *h* for preaspiration, and *n* for prenasalization. These final consonants (or final features as Miller calls them) do not show up overtly before a following vowel or *s w y*.

The glottals *h* and ' (glottal stop) occur, although they are often lost in medial position. There are no consonant clusters other than those described above.

Table 1. Final Consonants and Their Phonetic Effects on a Following Stop

hai	[hai]	'uncle'
hai-pa̲i̲	[haiβe]	'have (an) uncle'
tua"	[tua]	'son'
tua"-pa̲i̲	[tuappe]	'have (a) son'
haih	[hai]	'crow'
haih-pa̲i̲	[haiɸe]	'have (a) crow'
tsoon	[t̪oo]	'bead'
tsoon-pa̲i̲	[t̪oombe]	'have bead(s)'

---

4. To me, that is. Miller (pers. comm.) was not so sure.

5. Readers may notice a similarity between this and the famous Samoan problem in Gleason's (1955:31) introductory linguistics workbook.

These phonetic details are often important in syntactic analysis: Final -*n* (never pronounced before vowels and *s w y*) is the possessive marker. The same suffix may have different phonetic realizations due to the effects of stem-final consonants.

2.2 Shoshoni Syntax

The most common word order in Shoshoni is subject, adverbs, object, verb.

Nouns have three distinct forms: a subjective (unmarked) case form, an objective case form, and a possessive form. The unmarked form is used with postpositions. Numbers are singular, dual, and plural; however, only nouns referring to humans commonly show number. Inanimate nouns very rarely have singular or plural forms.

Verbs generally take tense/aspect markers, although they also occur without them. Verbs may also take number markers, but these are rare in normal speech. They are somewhat more common in traditional tales when the subjects are humans or mythological figures.

Adjectives are verbs, taking the same tense/aspect markers. In NPs, adjectives precede the nouns that they modify.

2.3 Notes on the Data

In the early 1970s, I studied linguistic techniques under Wick R. Miller at the University of Utah. Most of the classroom sessions for the two-term course were devoted to the collection of Shoshoni data from Clell Pete, a native speaker of the Gosiute dialect of Shoshoni. The data I collected consisted of his responses to students' how-do-you-say questions.

At the time, Miller was finishing up a collection of texts, translations, and a dictionary. That publication (Miller 1972), along with my class notes, provides nearly all of the data I have examined in the preparation of this study.

These two sources provide quite distinct types of data. My class notes show Shoshoni to be a rather common SOV language with the expected postpositions and with modifiers preceding heads. The collection in Miller 1972 consists of historical and mythological texts in addition to a few extended responses to Miller's questions. These stories, not surprisingly, contain language with much more complex sentences than what I collected.

3. Number Agreement

Shoshoni has subject-verb number marking. With imperatives only, this marking is obligatory; when no overt marker is present in commands, a singular subject is indicated. In all other cases, this type of marking is optional. It is generally used only when number is being emphasized. The number of the subject can be marked by a suffix *peweh* (dual) or *pemme* (plural) that directly follows the verb.[6] Because only

---

6. Miller (1972) regards these markers as enclitics rather than suffixes because of stress considerations.

some nouns (primarily animate nouns) can be overtly marked for number, and the use of the number markers with the verbs is optional, it is not surprising that the data contain very few examples of this type of agreement.

On the other hand, a second type of number agreement is fairly common in texts. This is an apparent ergative agreement. Miller (1972:18) describes it by saying that "a number of common verbs have special dual and plural forms, indicating number of the subject for intransitve verbs, object for transitive verbs."

Such ergative agreement is extremely rare in accusative languages; in fact Croft (1990: 90) specifically claims that languages with nominative-accusative case marking do not have ergative agreement.

The verbs that have these special forms are listed in tables 2 and 3. They have been placed in two tables because they fit neatly into two semantic groups. The verbs in table 2 relate to human actions. These verbs agree with their subjects. The subjects generally (but not always) take singular, dual, or plural forms. The verbs in table 3, on the other hand, are concerned with the manipulation of objects. These verbs seem to agree in number with their objects.

All of the verbs in table 2 have a meaning associated with human (or, at least, animate) movement or behavior. A surprising fact is that almost all of these verbs are intransitive. The transitive verbs on this list (*paikkaH*, *waseG* 'kill,' and *yekwiG*, *niweneG* 'say') do not consistently agree with their objects. Look at (1–3), which precede table 2. Example (1) shows *paikkaH* used with a third person object pronoun unmarked for number, but from context clearly referring to one person. Example (2) shows *paikkaH* used with a singular objective form that might refer to either a singular or plural antelope. Example (3) shows a clear use of *paikkaH* with a plural objective form.

(1) U paikkahkwa.
    u '3rd. obj. pron.'; *paikka-hkwa* 'kill-past'
    'He killed him.'

(2) Kwahatena paikkatoi, . . .
    kwahaten-a 'antelope-obj.'; *paikka-toi* 'kill-fut.'
    'He was going to kill antelope, . . . '

(3) Mawaihti, iwaihti tuineppewaihti u paikkahkwattsi u tekkayu.
    mawaih-ti 'ones like this-pl. obj.'; *iwaih-ti* 'ones like this-pl. obj.';
    tuineppe-waih-ti 'boys-like-pl. obj.'; *u* '3rd. obj. pron.'; *paikka-hkwa-ttsi*
    'kill-continuous-subord. first, . . .'; *u* '3rd. obj. pron.'; *tekka-yu* 'eat-past
    continuous'
    'Ones like this, boys like this, they would kill and eat.'

Given these facts, it seems unnecessary to provide any special rules for number agreement other than to indicate that these verbs are sensitive to the number of their subjects. This is not particularly surprising, because these verbs refer to human actions, and nouns referring to humans are almost the only ones regularly marked for number. These verbs differ from other verbs, transitive or intransitive, only in that number is indicated by stem changes rather than the addition of the optional suffixes, *-peweh* and *-pemme*. The two transitive verbs appear to follow the same pattern as the verbs in table 3.

Table 2. Movement and Action Verbs
with Apparent Ergative Number Agreement

SINGULAR	DUAL	PLURAL	GLOSS
anniH*	—	anniiteH, hamiaH	'fall over'
eppeiH	ekkoiH	ekkoiH	'sleep'
hapiG	kopiG, kwapiG	kopiG, kwapiG	'lie (down)'
kateG	yekwikka, yekwi	yekwiG	'sit'
kuhnaiG	kukuhnaiG	nutaa	'(start to) run'
mia	mimia	mia	'walk, go, move'
nahnaG	nanahna	nahnaG	'grow up'
nemi	nenemi, yeyenkaH	yenkaH	'move; live'
nukki	nunukki	nutaa	'run'
pahaiG	papahaiG	yokkaH	'fall off'
paikkaH	waseG	waseG	'kill'
piteH	pippite	piteH	'arrive, come'
taikwaG	niweneG	niweneG	'talk'
tepuiH	tetepuiH	tepuiH	'wake up (intr.)'
tiaiH	kokoiH	koiG	'die'
to'aiH, to'iH	toto'aiH	keyaG, keaG	'emerge'
waiG	wawaiH	waiG	'climb down'
wene	tsatsakkiH	topoiH	'stand'
yahnaiG	nanikkoiH	nanikkoiH	'laugh'
yaiH	yayaiH	waikuG	'enter'
yakai	nawoiH	nawoiH	'cry'
yekwiG	niweneG	niweneG	'say'
yekwiG	niweneH	niweneH	'talk'
yetseG	yoyotiH	yotiG	'get up; fly'

* The final G and H are dictionary markers indicating which forms of tense/aspect suffixes are used. G-verbs take geminate suffixes; H-verbs take preaspirated suffixes.

One possible explanation of the data in table 3 is that the singular and plural forms are, in fact, unrelated except for sharing meaning—much like *kill* and *massacre* in English. The words *wiiG* 'throw something light aside' and *paitiG* 'scatter, sow seeds' are different in the same way that *toss* and *scatter* are.

There is some evidence that the distinction between the forms is not as rigid as claimed. In counting the uses of *yaaG* and *himaG* 'get, carry' in the first four tales from Miller 1972, I found *yaaG* used 14 times with clearly singular objects, 3 times with plural objects, and 4 times with objects for which number is unclear. The verb *himaG* was used only 4 times, but in all of those instances with the same objects as *yaaG*. This is precisely the type of distribution we would expect if the agreement were simply a semantic implication rather than a syntactic phenomenon.

Table 3. Instrumental Verbs
with Apparent Ergative Number Agreement*

SINGULAR	PLURAL	GLOSS
-ka'aH	-ketaH	'cut (flexible object)'
-ka'iH	-ponkaH	'push (small object)'
-kepaH	-kepiiH	'break (hard object)'
-kumpaH	-tokaiH	'kill (using an object)'
-kwai'aiH, -kwaiyaH	-kwaiyuiH	'take out from a hole'
-kwitunaH	-kwituniiH	'wring out'
-kwitupiH	-kwitupii	'wrap with rope'
-metekkiH	-metekiiH	'turn (something) over'
-nuaG	-nuii	'move (tr.)'
-pahkaG	-paakuiG	'split (tr.)'
-paikoaH	-paikoaiH	'split down the middle'
-pataH	-patiiH	'spread (cloth) out'
-siwaH	-siwiiH	'tear'
-sommiH	-somiiH	'bend (stiff object)'
tawiG	paitiG	'throw (big object)'
-tehpaG	-tepaiH	'open'
tekiG	tappuiH	'put'
-tekwaH	-paiH	'hit; knock down'
-temaH	-temiiH	'lock in'
-tempukkaH	-tempukaiH	'pin (object)'
-tepekkiH	-tepekkiiH	'dig (hole)'
tomaaH	tottomaH	'make (arrow)'
-tsi'aH	-soiH	'chip'
wiiG	paitiG	'toss aside'
yaaG	himaG	'get, carry'

* Verbs preceded by a hyphen are not full verbs but rather stems that take a regular set of instrumental prefixes such as *tsa"* 'with the hand', *to"* 'with the fist', *ku* 'with the head'. The verb *yaaG* 'carry' may optionally take these prefixes.

Shoshoni inflects for singular, dual, and plural. However, the verbs in table 3 have only singular and plural forms. If their agreement were a syntactic phenomenon, we would expect some evidence of dual forms. There is none.

The noun morphology is clearly accusative. That is, objects are objective forms that differ from subjective forms, in spite of the ergative pattern of number agreement. Note that (4) contains the objective form *newi* with both *paikkaH* from table 2 and *tekkaH* 'eat', a normal transitive verb. Examples (5) and (6) show the expected objective forms for *tekkaH* and *yaaG*—the latter from table 3.

(4) Newi paikkahkwatti, newi tekkannu.
newi 'person/Indian (obj. form < *newe* + -*i*)'; *paikka-hkwa-tti* 'kill-go-sub-ord. suffix first . . .'; *tekka-nnu* 'eat-used to'
'He killed and ate people/Indians.'[7]

(5) Usuten taipo newi tekkayu.
usuten 'this/that one (subj. form)'; *taipo* 'white man (subj. form)'; *newi* 'person/Indian (obj. form)'; *tekka-yu* 'eat-past continuous'
'That white man ate the Indian.'

(6) Aikka taka un tohopaia yaakki.
aikka 'that (obj. form)'; *taka* 'only'; *un tohopai-a* 'his leg-obj.'; *yaa-kki* 'carry-toward'
'He just brought his leg.'

In summary, there are two types of verbs that might be claimed to have ergative number agreement. The first type includes intransitive verbs that take human (or at least, animate) subjects. As the subjects of intransitive verbs are identical in accusative and ergative languages (+Nom in Lexicase terms), this type of agreement does not require an ergative analysis. The second type includes transitive verbs indicating an action that affects a person or thing. The objects of these verbs take the same objective forms as all the other verbs of the language.

With the exception of those referring to humans, few nouns in Shoshoni are marked as singular or plural. Such nouns show syntactic agreement with verbs only when they are subjects—never when they are objects. The noun objects claimed to agree with their verbs in number show no overt inflectional or morphological evidence of that number agreement. Thus there is no reason to assume an ergative pattern of syntactic number agreement. In short, the data indicate that the apparent ergative agreement is actually nothing more than a meaning implication of single or multiple objects for certain transitive verbs.[8]

4. Is Shoshoni Really Accusative?

When I brought the number agreement in section 3 to the attention of William O'Grady and Stanley Starosta (both faculty members of the University of Hawai'i), they immediately asked me if I was sure that Shoshoni was accusative. I was and still am. However, since the question is an obvious one, I will respond to it.

In the data I collected personally, there is not a single sentence that would make me question an accusative analysis. The sentences below are typical. Object forms

---

7. The third person subject pronoun is zero.

8. As I make this claim, I am ignoring an important theoretical consideration for Lexicase: Are semantic features different from syntactic features? As the framework is currently set up, the notation for syntactic features is identical to the notation for semantic features. Starosta (pers. comm.) has mentioned that he uses the ? implication marker used by Lee (1989) to mark case frames, even for semantic implications. I would prefer to use the ⊃ for implications that are semantic, but I have seen no clear definition of what constitutes syntax as opposed to semantics in the Lexicase theory.

are consistently marked by one of the object suffixes -*a*, -*i*, or -*tta*. The suffix is pre-
dictable, although Miller (1972:19) states that it is only partially so, and that in some
instances two forms, apparently in free or dialectic variation, exist. In spite of this,
there is no reason to believe that the different forms mark different cases.

(7) Punku tosakaite.
   punku 'horse (subj. form)'; *tosa-kai-te* 'white-be-characteristically'
   'The horse is white.'

(8) Newe punkui puikka.
   newe 'person/Indian (subj. form)'; *punku-i* 'horse-obj.'; *puikka* 'see'
   'The Indian saw the horse.'

(9) Ne newi puikka.
   ne 'I (subj. form)'; *newi* 'person/Indian (obj. form < *newe* + -*i*)'; *puikka*
   'see'
   'I saw the Indian.'

(10) Ne aiha puikka.
   ne 'I (subj. form)'; *aih-a* 'crow-obj.'; *puikka* 'see'
   'I saw a crow.'

(11) Sukka wokaipitta teasen tekkamminna pemme mai, ne tsoonee.
   sukka 'that (obj. form)'; *wokai-pin-tta* 'prickly pear cactus-plant/animal-
   obj'; *teasen* 'all'; *tekka-mmi-nna-pemme* 'eat-regularly-past-plural'; *mai*
   'quotative particle'; *ne* 'I (subj. form)'; *tsoo-nee* 'great grandparent-subj. pl.'
   'They all ate cactus, my great grandparents.'

There are occasional sentences that do not use the normal objective; these are gener-
ally considered the result of noun incorporation in the verb (by lexical derivation):

(12) Ne aih-puikka.
   ne 'I (subj. form)'; *aih* 'crow (subj. form)'; *puikka* 'see'
   'I look in the way a crow does.'

(13) Nemme tapu(n)nukkutui mai niiwenete.
   nemme '1st. pl. excl. subj.'; *tapun* 'cottontail (subj. form)'; *nukku-tui*
   'roast-fut.'; *mai* 'quotative particle'; *niiwene-te* 'say (with dl./pl. subj.)-
   ongoing'
   'We're going cottontail roasting, they said.'

(14) Taputtsi uteemai tapun-yekkwi.
   Taputtsi 'Cottontail (the mythological figure)'; *utee-mai* 'them-with';
   tapun 'cottontail (subj. form)'; *yekkwi* 'kill'
   'Cottontail went cottontail hunting with them.'

The stories in Miller 1972 for the most part confirm an accusative syntax,
although occasionally the sentences are not quite as transparent as I would like. There
are even a few sentences in which the expected form is not used:

(15) Newe nanihai teahwai mai ne niikwinte.
newe 'Indian' (Miller's spelling is phonetic; it may be possessive *newen*);
na-niha-i 'reflex.-name-obj.'; *teahwai* 'tell about'; *mai* 'quotative particle';
ne 'I (subj. form)'; *niikwin-te* 'say-ongoing'
'Tell about Indian names, he says to me.'
(The expected form is *nia* 'me (obj. form)' rather than *ne*.)

(16) Aikka sonippeha, huuppitta, pohopitta, aisen memme ma puikka aikka ma seakkanna.
aikka 'this(obj. form)'; *sonippeh-a* 'hay-obj.'; *huuppi-tta* 'wood-obj.';
pohopi-tta 'sage brush-obj.'; *aisen* 'now'; *memme* 'you (pl. subj. or poss. form)'; *ma* '3rd. obj.'; *puikka* 'see'; *aikka* 'this(obj. form)'; *ma* '3rd. obj.';
seakka-nna 'grow(intr.)-past'
'Hay, trees, sage brush, right now you can see it growing.'
(The verb *seakkanna* should take subj. forms, not *aikka, ma*.)

(17) Mea u meenu.
mea moon (subj. form); *u* '3rd. obj.'; *mee-nu* 'make-past'
'He made a moon of it (the sun's gall bladder).'
(*Mea* is an unmarked subj. form rather than the expected obj. form.)

The total number of such sentences is small, and there is usually an explanation.[9] The most common problem involves the use of *ne* 'I' as an object and possessive form, instead of *nia* and *nian,* as in (15) and (11), respectively. I have no explanation, but this substitution is not uncommon. I can determine no conditioning factors. In fact, the sentence preceding (15) in the text includes both *ne* and *nia* in parallel object roles.

The least transparent sentences are from mythological and historical tales. As pointed out by Shimkin (1964), the Shoshoni story teller makes extensive use of sequential devices: parallel development to a climax, simple reiteration or repetition, sharp contrast, and parenthetical incorporation. These artistic performance devices, especially parallel forms and repetition, occasionally hinder a clear analysis of an utterance:

(18) Suten witsa aittun pa'anai kimmakinte(n) ukka newe un tohopai wekka'appeha ta'uhta kimma.
suten 'that (subj. form)'; *witsa* 'should'; *aittun* 'from there'; *pa'a-nai* 'high-from'; *kimma-kin-te* 'come-toward-ongoing'; *ukka* 'that (obj. form)'; *newe* 'Indian (subj. form)', possibly *newen* (poss. form) *un* '3rd. poss.'; *tohopai* 'leg (subj. form)'; *wekka'a-ppeh-a* 'cut off-nominalizer-obj.'; *ta'uhta* 'find'; *kimma* 'come'
'If he had come down from this direction, that Indian would have come and found his cut-off leg.'

In this example *ukka* and *wekka'appeha* are coreferential objective forms functioning as the object of the verb *ta'uhta.* If the translation (from Miller 1972) represents

---

9. One of the reviewers of this study asked about passives. I suspect that most of the examples of subjective forms in object meanings are in passive sentences, but I have not yet successfully determined how passives work in Shoshoni, so I cannot make this suggestion with much certainty.

the Shoshoni structure, *newe*, which is coreferential with the subject *suten*, sits in among a group of objects.[10] Example (19) is similar.

(19) Suten tuku sukka suten taipo tainnappe u paikkattsi, un takkui tekkanu.
      suten 'that (subj.)'; *tuku* 'just'; *sukka* 'that (obj. form)'; *taipo* 'white-man (subj.)'; *tainnappe* 'mature man (subj.)'; *u* '3rd. obj.'; *paikka-ttsi* 'kill-first, . . .'; *un* '3rd. poss.'; *takku-i* 'meat-obj.'; *tekka-nu* 'eat-past'
      'That white man had simply killed him and eaten his meat.'

Once again, the phrases referring to the subject, *suten* and *suten taipo tainnappe*, are mixed with the objects *sukka* and *u*.

In spite of such unclear examples, Shoshoni is clearly an accusative language. Armed with just a few sentences like those listed above, I would be quite hesitant to suggest that it was not.

5. Subordinate Clauses

In Shoshoni, subordinate clauses can be divided into those that may have subjects coreferential with the subject of the main clause and those whose subjects may not be coreferential with those of the main clause.

In adverbial clauses with the subject coreferential with that of the main clause, the arguments take the expected (accusative) subjective and objective forms. The most common verb suffix in clauses of this type is *-ttsihi* 'first . . . , then' as in these sentences.

(20) Pemman kahnika pitettsi[11] suten nuhnuhnuh mai yekwiti nankanu, mai suten.
      pemman '3rd reflexive poss.'; *kahni-ka* 'house-at'; *pite-ttsi* 'arrive-TTSIHI'; *suten* 'that (subj.)'; *nuhnuhnuh* 'nhnh sound'; *mai* 'quotative particle'; *yekwi-ti* 'say' probably < *yekwi* + *-ten* 'nominalizer' + *-i* 'obj.'; *nanka-nu* 'hear-past continuous'; *mai* 'quotative particle'; *suten* 'that (subj.)'
      'When they arrived at their house, they heard that nhnh sound, they say.'

(21) Sute newe u yaatsi wihyu, kimmanu.
      suten 'that (subj.)'; *newe* 'person/Indian (subj.)'; *u* '3rd. obj.'; *yaa-tsi* 'carry-TTSIHI'; *wihyu* 'then'; *kimma-nu* 'come-past continuous'
      'After that person took it, he came (back).'

Other subordinate verb suffixes listed by Miller (1972:17) include *-ppeh* 'perfective', *-ten* 'that which Vs', and *-kanten* 'that which characteristically Vs'. These verbs generally form relative clauses (or noun phrases), but they are occasionally in adverbial position and might be translated as adverbial clauses. The first of these forms, *-ppeh*, has perfective meaning in main clauses as well as in subordinate clauses. The subjects and objects in clauses with *-ppeh* verbs have normal subject and object

---

10. An alternative analysis would be that *newe* is an underlying *newen* (poss.) coreferential with *un* and that *tohopai* (subj. form) is incorporated in the verb *wekka'a* before noun derivation. However, noun incorporation is normally used when the object is generic, not specific as in this example.

11. The form of the suffix *-ttsihi* varies a great deal in the texts.

forms. There does not seem to be any significant difference between *-ppeh* verbs in subordinate and main clauses.

(22) Sateweh ko'ippeh un tutuaneweh.
sateweh 'those (dual subj.)'; *ko'i-ppeh* 'die-PPEH'; *un* '3rd. poss.'; *tutu-aneweh* 'sons (dual subj.)'
'Those two who died were her children.'

(23) Newi tekkappeh.
newi 'person/Indian (obj.)'; *tekka-ppeh* 'eat-PPEH'
'They ate people.'

The exceptions to this general rule involve the use of possessives when *-ppeh* derives nouns from verbs. Compare the use of the subjective form in (22) to the use of possessive forms for arguments of *-ppeh* verbs (24) and (25). These derived nouns, when functioning as objects, take the regular objective suffix *-a* as in (26).

(24) Kai wihyu hakkan nihappeh, newe wihyu sesemme ma puikkante Otsi ma niikwinnimma.
kai 'not'; *wihyu* 'then'; *hakkan* 'whose? or anyone's (poss.)'; *niha-ppeh* 'name-PPEH'; *newe* 'person/Indian (subj.)'; *wihyu* 'then'; *sesemme* 'one (redup.)'; *ma* '3rd. obj.'; *puikkan-te* 'see-ongoing'; *Otsi* 'Fox'; *ma* '3rd. obj.'; *niikwin-ni-mma* 'call-to and fro-past'
'He got that name from nobody in particular, but just some people saw it here and there, and called him Fox.'

The phrase *kai wihyu hakkan nihappeh* in (24) appears to be a negative copular sentence. Such NP = NP sentences do not have overt verbs in Shoshoni. The predicate has the expected subjective case form. In (25) *iten newen tekkappeh* is the subject of the main verb *kammanna* 'taste (intr.)' while in (26) *teihtettsia tekkappeha* shows the objective case suffix signaling its use as object of the verb *tekkayu*.

(25) Iten newen tekkappeh kaitsaan kammanna.
iten 'this (poss.)'; *newen* 'Indian (poss.)'; *tekka-ppeh* 'ate-PPEH'; *kai-tsaan* 'not-good'; *kammanna* 'taste'
'Indian food doesn't taste good anymore.'

(26) Teihtettsia tekkappeha tekkayu ukkwihyunte.
teihtettsia 'a little (obj.)'; *tekka-ppeh-a* 'eat-PPEH-obj.'; *tekka-yu* 'eat-continuous'; *ukkwihyunte* 'there was, it happened'
'They ate a little bit of food there.'

My data contain few examples of *-ten* and *-kanten* in subordinate clauses, but from these it appears that the structural pattern is similar to that for *-ppeh*. They take the normal subjective and objective forms, as in (27).

(27) Huuppin aiten tammen puikkaneen huuppin aittsu seakkanten newe nii-wenehkante.
huuppin 'tree (subj. form)'; *aiten* 'that (subj. form)'; *tammen* '1st. pl. incl.'; *puikkaneen* 'seen things (pl. subj.)'; *huuppin* 'tree (subj. form)'; *aittsu*

'there'; *seakkan-ten* 'grow-TEN'; *newe* 'Indian'; *niiwene-hkante* 'speak-characteristically'
'All those trees we saw among us, all growing things there spoke the Indian language.'

Clauses whose "subjects" are never coreferential with the subject of the main clause take different subordiate suffixes (Miller 1972). These are *-ku* 'durative' and *-kka* 'momentary, completed'. Verbs with these suffixes are unusual in that their apparent subjects have objective case forms

(28) Sukka miaku nia puikka.
sukka 'that (obj.)'; *mia-ku* 'walk-subord. suffix'; *nia* 'me'; *puikka* 'see'
'When he$_i$ was walking, he$_j$ saw me.'

(29) Nemmi teteihku pettuyumminna.
nemmi '1st. pl. excl.'; *teteih-ku* 'small-subord. suffix'; *pettuyummin-na* 'heavy-past'
'When we were small, they were heavy.'

(30) Sapai u pitenuhka newe kia suten umaka piteheppeh.
sapai 'there'; *u* '3rd. obj.'; *pite-nu-hka* 'arrive-slowly-subord. suffix'; *newe* 'person/Indian (subj.)'; *kia* 'perhaps'; *suten* 'that (subj.); *umaka* '3rd.-at'; *pite-heppeh* 'arrive-perfective'
'When he$_i$ arrived there, an Indian approached him$_i$.'[12]

The two suffixes *-kka* and *-ku* are those that are relevant to this paper, since they do not take the expected case forms for an accusative language. In the following discussion, I concentrate on the suffix *-ku* to avoid confusion, because the form *-kka* occurs both on subordinate verbs and on main verbs (with different meanings) according to Miller (1972:17). I use examples with *-kka* only when the same pattern is evident for *-ku*.

6. The Suffix *-ku*

There are three phonetically identical suffixes *-ku* in the data. Only the third is important to this discussion. In the next two subsections I explain why the first two are not relevant, before turning my attention to the third.

6.1 *-ku* on Main Verbs

There are a few sentences in which *-ku* occurs on the main verb:

---

12. One reviewer suggested that the *u* 'him' referring to the subject of *pitenuhka* is actually the object of the main verb *piteheppeh*. While possible, it seems unlikely in light of the many other examples like (28) and (29). Also, the simple objective pronoun *u* almost always directly precedes its verb.

(31) Suten newe u punittsi, kuhnaikku.
suten 'that (subj.)'; *newe* 'person/Indian (subj.)'; *u* '3rd. obj.'; *puni-ttsi*
'see-first ... , then'; *kuhnai-kku* 'run-away'
'That Indian saw him and ran off.'

Fortunately all of these occurrences of *-ku* on the main verb correspond to a meaning of 'away'. Shoshoni has an adverbial suffix *-kwa* 'away'. Final vowels are generally unvoiced. An unvoiced /a/ with lip rounding from the /kʷ/ can easily be confused with a /u/. For this reason, I will assume that these are examples of the adverbial suffix *-kwa* in final position.

Arguments of verbs with this suffix take the normal case forms.

6.2 *-ku* on Locatives and Postpositions

Somewhat more problematic is the *-ku* that occurs on locatives and postpositions.

(32) Tepanmaihku ma tekkayu.
tepan-maihku 'pine nuts-with'; *ma* '3rd. obj.'; *tekka-yu* 'eat-used to';
'They ate it with pine nuts.'

(33) Haiyattsa aipaikku u kwetti.
haiyattsa 'exclamatory adverb = oh yes!'; *aipaikku* 'there or then'; *u* '3rd.
obj.'; *kwetti* 'shoot'
'He shot him right on the spot.'

I am not at all certain what is happening here, but pospositions and locatives clearly change form. Neither the function nor the distribution of the different forms is completely clear. Miller (1972:22) states that the five forms vary with the verb in the clause. Table 4 contains a few examples, along with the types of verbs they occur with.

The last category is not so well defined as the others. That is, the verbs that occur with these forms do not form as clear a semantic group. Postpositional phrases are usually in the same position in the sentence as subordinate clauses. Given these facts, one might wish to claim that the forms ending in *-ku* are actually subordinate verbs. However, there are three good reasons for not doing this. First, in spite of the large number of such postpositions in the texts, there is not a single case of a postposition ending in *-ku* taking an objective-marked subject. Postpositions always follow unmarked (subjective) nouns. Second, the meaning associated with the subordinate verb suffix *-ku* is 'momentary, completed' while the meaning normally associated

Table 4. Postposition Forms

	'ON TOP OF'	'WITH (INSTR.)'	'AT'
INTR. V. OF REST	-pan	-man	-ka"
MOTION (TO) V.	-pai	-mannai	-kai
MOTION (FROM) V.	-panti	-manti	-katti
TR. V. OF REST	-panten	-manten	-katten
(TR.?) V. OF MOTION	-panku	-manku	-kakku

with postpositions is more like that of *-kka* 'durative'. Third, if *-ku* and *-kka* are members of the same class of subordinate suffixes, there should also be posposition forms ending in *-kka*. There are none.

6.3 *-ku* on Subordinate Verbs in Adverbial Clauses

These are the examples that demonstrate the unusual objective forms in subject position, as in (28–30) above, repeated here as (34–36).

(34) Sukka   miaku   nia   puikka.
     that(obj.)   walk-KU   me   see
     'When he_i was walking, he_j saw me.'

(35) Nemmi   teteihku   pettuyumminna
     us   small-KU   heavy-past
     'When we were small, they were heavy.'

(36) With *-kka* in a similar construction:
     Sapai̱ u   pitenuhka newe   kia   suten   umaka piteheppeh
     there him arrive-KKA Indian(subj.) perhaps that(subj.)him   arrived
     'When he_i arrived there, an Indian approached him_i.'

Given the normal order of constituents in a Shoshoni sentence, the reason for this use of object forms for subjects in adverbial clauses is quite clear:

(37) $subject_{MAIN}$ $subject_{ADVERBIAL}$ $verb_{ADVERBIAL}$ $object_{MAIN}$ $verb_{MAIN}$

Shoshoni does not require overt subjects or objects, so the subject of an adverbial clause might be the first word or phrase in the sentence, easily mistaken for the subject of the main clause.[13] An objective form clearly indicates that the word is not the subject of the main clause.[14]

The trade off is that objects in these adverbial clauses would be indistinguishable from the subjects (in objective form). Not surprisingly, the data contain very few examples of transitive verbs in adverbial clauses with both subjects and objects. Transitive verbs taking *-ku* or *-kka* normally have only one argument. The overt noun seems to be the logical subject of the subordinate verb. However, most of the examples with transitive verbs have pronominal arguments, making it impossible to be certain whether the subject or object is the argument overtly appearing. In (38), does *u* refer to the man or the leg he was preparing to eat? My assumption that *u* refers to the man (the logical subject) is based on similar patterning for full noun phrases; there is nothing in the structure of (38) to indicate that this is so.

---

13. The historical and mythological texts in Miller 1972 are misleading in that the story tellers regularly use demonstratives as subjects and objects. In other speech, overt subject and object pronouns are somewhat less common.

14. If this explanation is correct, one might expect more languages to use this or a similar strategy. My experience with verb-final languages is limited to Shoshoni (and some elementary Japanese), so I am not in a good position to determine whether it is found elsewhere. Is this type of construction limited to Shoshoni?

(38) Tokain u tekkahtoihka, suten newe umaka pite.
tokain 'exactly'; *u* '3rd. obj.'; *tekkah-toi-hka* 'eat-fut.-KKA'; *suten* 'that
(subj.)'; *newe* 'Indian(subj.)'; *uma-ka* '3rd-at'; *pite* 'arrive'
'Just as he (the white man) was about to eat (the leg), that Indian
approached him.'

My data do include about ten sentences in which both the subject and object of the
subordinate -*ku* (or -*kka*) verb are present. Examples (39), (41), and (43) are typical.

(39) Ikka taka ne kai keppuihka, ne huttsia makkah hannimminku hukkana.
ikka 'this (obj. form)'; *taka* 'only'; *ne* 'I'; *kai* 'not'; *keppuihka* 'taste (tr.)';
ne 'I'; *huttsi-a* 'grandmother (obj.)'; *makkah* 'there'; *hannimmin-ku* 'pre-
pare-KU'; *hukkan-a* 'pickleweed (obj.)'
'I haven't tried it, but my grandmother used to gather pickleweed.'

Example (39) is an unusual structure. It does not have the expected SOV word
order. The pause indicated by the comma makes it look similar to (40):

(40) Tokain tekkatoi, suten taipo.
tokain 'exactly'; *tekka-toi* 'will-fut.'; *suten* 'that (subj.)'; *taipo* 'white man
(subj.)'
'He was about to eat, that white man.'

In these sentences, a noun phrase corresponding to an argument in the main sen-
tence follows the main verb. In (39), *ikka*, the object in the main clause, is clearly
coreferential with *hukkana*, the logical object of the -*ku* verb *hannimminku* as well as
the main verb *keppuihka*. Given the case markings, the phrase *ne huttsia makkah han-
nimminku* can be considered a clause modifying *hukkana*. Following this reasoning, a
better translation of (39) might be 'I haven't tried it, the pickleweed that my grand-
mother used to gather there'.

This is not a completely satisfactory analysis of sentence (39) because -*ku* verbs
generally form adverbial clauses rather than adjectival clauses. However, a similar
analysis of (41) and (43) seems justifiable. The nouns *taiponee* in (41) and *newenee* in
(43) function as subjects of the main clause. Translations more clearly reflecting the
Shoshoni structure might be those shown below in (42) and (44).

(41) Wihyu ne puikkanku taiponee wihyu wopimpa'an taka nayekwinte
naappehkante.
wihyu 'then'; *ne* 'I'; *puikkan-ku* 'see-KU'; *taiponee* 'white men(pl. subj.)';
wihyu 'then'; *wopim-pa'an* 'wagon-on; *taka* 'only'; *nayekwin-te* 'travel';
naappehkante 'aux. used to'
'I saw the white men who traveled by wagon.'

(42) same as (41)
Wihyu <u>ne puikkanku taiponee</u> wihyu wopimpa'an taka nayekwinte
naappehkante.
'The <u>white men that I saw</u> used to travel on wagons.'

The noun *taiponee* is clearly the subject of the main verb *nayekwinte naappehkante*. It
has subjective form and logically fills the subject role in the gloss. A similar analysis

is possible for (43), where *newenee*, a noun in the subjective form, functions as the subject of the main verb *punkapai*. The phrase ending in *-ku* functions as an adjective clause, as shown in the translation (44).

(43) Ekittsi wihyu ne puikkanku newenee ne nuha punkapai.
ekittsi 'now'; *wihyu* 'then'; *ne* 'I'; *puikkan-ku* 'see'; *newenee* 'people/Indians (pl. subj.)'; *nuha* 'used to (adv.)'; *punkapai* 'horse-have'
'I can see that the Indians and I used to have horses.'

(44) same as (43)
Ekittsi wihyu ne puikkanku newenee ne nuha punkapai.
'The Indians that I see now and I used to have horses.'

The fact that phrases headed by verbs with *-ku* and *-kka* are never accompanied by more than one overt argument might lead to the claim that such phrases are not actually subordinate clauses. When compared to other subordinate clauses, specifically those with verbs carrying the *-ttsihi* suffix or those with normal main clause tense/ aspect, they are clearly deviant. All other types of subordinate clause allow both subjects and objects. However, the adverbial *-ku* and *-kka* phrases pattern exactly as other subordinate clauses, except for the limitations on number of arguments and the use of objective case form for logical subjects:

(45) Itsappe pahotiaihkante, newi miakkinku, u miankettsi, u paikkahkwattsi, u tekkanu, newi itsappe.
Itsappe 'Coyote (subj.)'; *pahotiai-hkante* 'hungry'; *newi* 'person (obj.)'; *mia-kkin-ku* 'move-to and fro-KU'; *u* '3rd. obj.'; *mianke-ttsi* 'chase-first . . . , then'; *u* '3rd. obj.'; *paikkahkwattsi* 'kill-first . . . , then'; *u* '3rd. obj.'; *tekka-nu* 'eat-continuous'; *newi* 'person (obj.)'; *itsappe* 'Coyote (subj.)'
'When Coyote got hungry, when people were walking, he chased them, killed them, and ate them—Coyote (did that to) people.'

Notice the parallel usage of the *-ku* clause *newi miakkinku* and the *-ttsihi* clauses *u miankettsi* and *u paikkahkwattsi* in (45). Notice also the position between the adverbial clause *pahotiahkante* 'when he was hungry' and the *-ttsihi* clauses.

7. A Lexicase Description

This paper has examined two problems for an accusative description of Shoshoni. The first, in section 3, relates to apparent ergative number agreement. A description of this using Lexicase formalism is straightforward. Subcategorization rules mark nouns as +human or –human, then further subcategorize +human nouns for number. Verbs used with human nouns (those in table 2) are marked for number in the lexicon. An inflectional rule prohibits different number on the verb and its subject. The apparent number agreement of verbs (in table 3) is simply handled by a semantic implication in the feature matrix of the verb. The second, in section 6.3, is the use of a single noun phrase with objective case form in phrases headed by a subordinate verb ending in *-ku* or *-kka*. This is not so easily handled.

289

One of the strengths of lexicase is that it requires description without reference to multiple layers of structure or transformations. When relating different word classes, we are required to use derivation. While phrasal derivation is not prohibited (see the discussion of NP::Det derivations for English possessives of Starosta 1988:100), the assumed change of a subject form to an object form in phrases is. The derivational rule needed would relate the head of -*ku* and -*kka* clauses to other verbs:

$$
\text{DR-1 } [+V] \quad :: \quad \begin{bmatrix} +V \\ \\ ? \begin{bmatrix} +\text{actr} \\ -\text{Nom} \end{bmatrix} \end{bmatrix}
$$

$$] \quad :: \quad ku]$$

In other words, a verb may correspond to a verb (ending in -*ku*) that expects an actor in object form as a dependent.

This is a far-from-satisfactory rule. Consider the case frame of the original (underived) verb. An intransitive verb would expect a PATIENT that would be indexed by linking rules as the actor and subject. (For information about linking rules, see Lee 1989, Springer 1992, Starosta 1998.) The universal rule assigning a nominative case form to the expected PATIENT (of intransitive verbs) in an accusative language must be disabled, presumably by an ordering of linking rules. The question of rule ordering for linking and chaining rules is one that has for the most part been avoided in the Lexicase literature because it has not seemed to have much importance (Lindsey 1992:39–41). Clear exceptions to general rules may demand that we pay more attention to the ordering question.

A transitive verb would also expect a PATIENT, with the accusative case-marking rule assigning an accusative case form. Since a PATIENT is assumed in the case frame of every verb (Starosta 1988:128), the derived verb must also expect a PATIENT. Which original argument assumes the PATIENT role in the derived case frame? Since there is only one overt argument, the actor, it would seem logical to assign PATIENT status to that argument. However, such a solution would not allow for the use of chaining rules (Lee 1989) in the interpretation of modified nouns like those in (39), (41), and (43) as the implied PATIENT of the subordinate verb.

This problem of case frames is not fatal, however. The derivation rules relating passive verbs and transitive verbs in English involve the deletion of an AGENT from the case frame of the derived (passive) verb. This is justified, because Lexicase requires a syntactic definition of case relations (PATIENT, AGENT, MEANS, etc.) rather than one based on semantic criteria like the cases of Gruber (1965) and Fillmore (1968). Following this logic, we could claim that our derived forms do not have the original PATIENT in the case frame and that the original ?[AGT] (expected or implied AGENT) corresponds to the ?[+PAT] of the derived form. This may even be reasonable to some extent, due to the distribution of adjectival clauses in Shoshoni. Some

clauses, usually short in length, precede the noun. Other clauses, almost always with an overt relative pronoun, follow the noun. We might presume that the adjective clauses that precede have no missing case relation that must be interpreted, while those that follow do have such missing case relation. On the other hand, there does not seem to be much motivation for this approach.

In spite of the problems connected with DR-1, it does seem to account for the facts of the syntax. What this rule does not do is account for the native speaker's recognition of the fact that the "subject" of the verb with -*ku* is not coreferential with the subject of the main clause.

I believe this to be the same type of choice a speaker makes in deciding whether to use a pronoun, a reflexive, or a semantically fuller NP. I have argued elsewhere (Lindsey 1985) that that choice is a performance decision directly related to the speaker's evaluation of the amount of information needed to make the meaning clear, following logically from Grice's (1975) maxims of quantity. In that paper, I argued that the speaker of English uses special forms (stressed pronouns, reflexives, demonstratives) even when they are not required—and sometimes when they are prohibited by the syntax, as in the the case of coreferential full NPs within the same sentence—in order to clarify reference.

I believe the Shoshoni speaker uses the syntactic structure available (actor in the objective form along with -*ku* or -*kka*) in much the same way an English speaker uses marked forms to signal meaning.

8. Conclusion

It might be claimed that Shoshoni has some ergative patterns because (a) some transitive verbs show number agreement with their objects, and (b) some subordinate clauses have subjects with unexpected case forms. A close examination of the data shows that (a) can be attributed to the meanings of the verbs and that it has no morphological[15] or syntactic consequences. The (b) clauses are more difficult to account for. Object forms clearly perform the function of distinguishing the subjects of subordinate clauses from the subject of the main clause. The derivational rule required to account for the syntax has unpleasant consequences for Lexicase theory. I am dissatisfied with this solution in that such a complex set of theoretical adjustments is necessary to describe a construction whose functional load is so clear.

---

15. Sharp-eyed readers will have observed that verbs from table 3 often end in -*ii*. While there may be a historical relationship between such a suffix and plurality of the object, the ending is not productive in the modern language.

REFERENCES

Croft, William. 1990. *Typology and universals*. Cambridge University Press.

Fillmore, Charles J. 1968. The case for case. In *Universals in linguistic theory*, ed. by E. Bach and R. T. Harms. New York: Holt, Rinehart, and Winston.

Gleason, Henry Allan, Jr. 1955. *Workbook in descriptive linguistics*. New York: Holt, Rinehart, and Winston.

Grice, H. P. 1975. Logic and conversation. In *Syntax and semantics*, vol. 3: *Speech acts*, ed. by Peter Cole and Jerry L. Morgan, 41–58. New York: Academic Press.

Gruber, J. S. 1965. Studies in lexical relations. Ph.D. dissertation, Massachusetts Institute of Technology.

Lee, Nagiko I. 1989. Complementation in Japanese: A lexicase analysis. Ph.D. dissertation, University of Hawai'i.

Lindsey, Francis Jr. 1985. Pronoun interpretation: An interpretive component for the lexicase theory. University of Hawai'i *Working Papers in Linguistics* 17:71–94.

———. 1992. A computational model for the testing of linguistic hypotheses concerning language change. Ph.D. dissertation, University of Hawai'i.

Miller, Wick R. 1972. *Newe natekwinappeh: Shoshoni stories and dictionary*. Salt Lake City: University of Utah Press.

Shimkin, Dmitri B. 1964. On Wind River Shoshone literary forms: An introduction. In *Language in culture and society*, ed. by Dell Hymes, 344–351. This is an abridged version of the article from *Journal of the Washington Academy of Sciences*, 1947, 37:329–352.

Springer, Hisami Konishi. 1992. Japanese topicalization: A lexicalist dependency analysis. In *Pan-Asiatic Linguistics: Proceedings of the Third International Symposium on Language and Linguistics*, 1191–1206. Bangkok: Chulalongkorn University.

Starosta, Stanley. 1988. *The case for lexicase: An outline of lexicase grammatical theory*. London: Pinter Publishers.

———. 1998. Lexicase revisited. In *Studies in linguistics in honour of Béla Korponay: A special issue*, ed. by József Andor, Béla Hollósy, Tibor Laczkó, and Péter Pelyvás, 279–334. Debrecen: Kiadta a Kossuth Lajos Tudományegyetem, Angol-Amerikai Intézete.

INDEX