

PACIFIC LINGUISTICS

Series C - No. 31

SOUTH-EAST ASIAN LINGUISTIC STUDIES

by

Nguyen Dang Liem, Ed.



Department of Linguistics  
Research School of Pacific Studies  
THE AUSTRALIAN NATIONAL UNIVERSITY

*PACIFIC LINGUISTICS* is published through the *Linguistic Circle of Canberra* and consists of four series:

- SERIES A - OCCASIONAL PAPERS
- SERIES B - MONOGRAPHS
- SERIES C - BOOKS
- SERIES D - SPECIAL PUBLICATIONS

EDITOR: S.A. Wurm.

ASSOCIATE EDITORS: D.C. Laycock, C.L. Voorhoeve, D.T. Tryon, T.E. Dutton.

EDITORIAL ADVISERS:

- |  |  |
|--|--|
| B. Bender, University of Hawaii              | A. Healey, Summer Institute of Linguistics                       |
| A. Capell, University of Sydney              | N.D. Liem, University of Hawaii                                  |
| S. Elbert, University of Hawaii              | H. McKaughan, University of Hawaii                               |
| K. Franklin, Summer Institute of Linguistics | G.N. O'Grady, University of Victoria, B.C.                       |
| W.W. Glover, Summer Institute of Linguistics | K. Pike, University of Michigan; Summer Institute of Linguistics |
| G. Grace, University of Hawaii               | E. Uhlenbeck, University of Leiden                               |

ALL CORRESPONDENCE concerning *PACIFIC LINGUISTICS*, including orders and subscriptions, should be addressed to:

The Secretary,  
*PACIFIC LINGUISTICS*,  
Department of Linguistics,  
School of Pacific Studies,  
The Australian National University,  
Canberra, A.C.T. 2600.  
Australia.

Copyright © The Authors.

First published 1974.

Reprinted 1978.

The editors are indebted to the Australian National University for help in the production of this series.

This publication was made possible by an initial grant from the Hunter Douglas Fund.

National Library of Australia Card Number and ISBN for the complete set: 0 85883 144 9

National Library of Australia Card Number and ISBN for this volume: 0 85883 099 X

## TABLE OF CONTENTS

	<i>Page</i>
SEMANTIC ANALYSIS OF DATANG IN INDONESIAN, by Soenjono Dardjowidjojo	1
1. INTRODUCTION	1
2. CENTRALITY OF THE VERB	4
2.1. The problems	6
2.2. The analysis	6
3. CONCLUSION	19
<i>Notes</i>	21
<i>References</i>	22
ON MORPHOLOGICAL AND SYNTACTIC RELATIONS IN A SOUTHEAST ASIAN LANGUAGE, by Cesar A. Hidalgo	25
0. INTRODUCTION	25
1. THE PREDICATIVE STEM	26
1.1 The simple stem	27
1.2 The derived stem	27
2. THE PREDICATIVE AFFIXES	29
2.1 The stem and focus affixes	30
2.2 The stem and voice affixes	32
3. PROPOSAL	35
<i>Notes</i>	36
<i>References</i>	38

	<i>Page</i>
LAO FINAL PARTICLES, by Arthur G. Crisfield	41
THE DEVELOPMENT OF THE REGISTERS IN STANDARD KHMER, by Philip N. Jenner	47
<i>Notes</i>	58
OBSERVATIONS ON THE SURIN DIALECT OF KHMER, by Philip N. Jenner	61
0. INTRODUCTION	61
1. SEGMENTAL PHONEMES	63
1.1. Vowels	63
1.2. Consonants	63
2. SUPRASEGMENTALS	66
2.1. Stress	66
2.2. Intonation	66
3. STRUCTURE OF THE WORD	67
4. SAMPLE TEXT	68
<i>Notes</i>	73
PASSIVE AND ERGATIVE IN VIETNAMESE, by Marybeth Clark	75
0.0. INTRODUCTION	75
0.1. The problem	75
1.0. THE GRAMMAR	76
1.1. Rules for the structure of sentences	76
1.2. Justification of the rules	76
1.3. A sampling of lexicon	78
1.4. Conventions and redundancy rules	79
2.0. SUBMISSIVE VERBS	80
3.0. ERGATIVE CONSTRUCTIONS	84
<i>Notes</i>	86
<i>References</i>	87
SUBMISSIVE VERBS AS ADVERSATIVES IN SOME ASIAN LANGUAGES, by Marybeth Clark	89

	<i>Page</i>
0. INTRODUCTION	89
1. PASSIVE AND ADVERSATIVE	89
2. THE SITUATION IN VARIOUS LANGUAGES	90
2.1. Japanese	90
2.2. Vietnamese	94
2.3. Thai and Lao	98
2.4. Mandarin Chinese	99
2.5. Cambodian	102
2.6. English	104
3. A SUMMARY OF SUBMISSIVES	105
4. A COUPLE OF SUBMISSIVE QUESTIONS	106
<i>Note</i>	108
<i>References</i>	108
MODERN LINGUISTIC THEORIES AND CONTRASTIVE ANALYSIS OF ENGLISH AND VIETNAMESE, by Nguyen Dang Liem	111
0. INTRODUCTION	111
1. TAGMEMICS AND PRE-ASPECTS TRANSFORMATIONAL PROCEDURE	114
1.1. Clause types in contrast	115
1.2. Clause classes in contrast	119
2. CASE GRAMMAR	122
3. GENERATIVE TRANSFORMATIONAL INSIGHTS	124
4. STATISTICAL STUDIES	124
5. THEORETICAL CONSIDERATIONS	125
<i>References</i>	126
CLAUSES AND CASES IN ENGLISH AND SOUTHEAST ASIAN LANGUAGES (BURMESE, CAMBODIAN, CANTONESE, LAO, THAI, AND VIETNAMESE) IN CONTRAST, by Nguyen Dang Liem	129
0. INTRODUCTION	129
1. CLAUSE UNITS IN THE LANGUAGES	130
1.1. English	131
1.2. Burmese	133
1.3. Cambodian	138

	<i>Page</i>
1.4. Cantonese	141
1.5. Lao and Thai	143
1.6. Vietnamese	146
2. CLAUSE UNITS IN CONTRAST	149
2.1. Clause types in contrast	149
2.2. Clause classes in contrast	150
3. DEEP VERSUS SURFACE STRUCTURE	150
4. PHILOSOPHY OF APPROACH	153
References	154
THE VALUE OF AU AND AI IN MIDDLE KHMER, by Philip N. Jenner	157
Notes	172
THE WORD ẮC IN KHMER: A SEMANTIC OVERVIEW, by Saveros Pou	175
Notes	189
Abbreviations	191
A CLASSIFICATION OF VERBS IN VIETNAMESE AND ITS PEDAGOGICAL IMPLICATIONS, by Nguyen Dang Liem	193
INTRODUCTION	193
CLASSIFICATION OF VERBS	195
1. là 'be'	195
2. chậm 'slow'	196
3. lạnh 'cold'	198
4. chết 'to be dead'	199
5. bị 'be adversely affected'	199
6. mua 'buy'	201
7. bán 'sell'	202
8. chọn 'choose'	202
9. ăn 'eat'	203
10. nói 'speak'	203
11. biết 'know'	204
12. ghét 'hate'	204
13. ở 'stay at, be at'	205
14. ra 'go out'	206

	<i>Page</i>
15. đi 'go'	207
16. tới 'arrive'	207
17. có 'have'	208
References	209

---

EDITORIAL NOTE

Some of the articles included in this volume were written around 1972, and theoretical views expressed in them may be at variance with the authors' present views on the subject. This note has been inserted at the specific request of the Editor of this volume on behalf of the authors.

---





# SEMANTIC ANALYSIS OF DATANG IN INDONESIAN<sup>1</sup>

SOENJONO DARDJOWIDJOJO

## 1. INTRODUCTION

When Chomsky's *Syntactic Structures* was published, a grammar was considered to have a tripartite arrangement: the phrase structure, the transformational structure, and the morphophonemics (p. 46). Semantics at that time was placed outside the grammar and was considered harmful to linguistics. Two years earlier he said that "if it can be shown that meaning and related notions do play a central role in linguistic analysis, then its results and conclusions become subject to all of the doubts and obscurities that plague the study of meaning, and a serious blow is struck at the foundations of linguistic theory" (Chomsky, 1955, p. 141). As the time passed, however, generative grammarians began to feel that the elimination of meaning from linguistic analysis became more and more impossible. Katz and Fodor (1963) presented a detailed study of semantic theory which has a significant role in the development of the present generative grammar. This theory was later integrated into a more comprehensive work (Katz and Postal, 1964) where a grammar was then viewed as consisting of three components: syntactic, phonological, and semantic. It was later "standardised" with the appearance of Chomsky's *Aspects of the Theory of Syntax*. One characteristic that the above works share is that they hold syntax to be central and the other two components peripheral.

At the time when *Aspects* was published, George Lakoff finished his dissertation which later became known as *Irregularity in Syntax*. Although it was originally intended as a minor revision to *Aspects*, it soon became apparent that it dealt with fundamental issues that at the end led to the rejection of syntax and semantics as two separate entities. One logical conclusion from this indifferentiation is the merging

of deep structures with semantic representations. Lakoff's approach simplifies the base, as many of the problems are now handled transformationally.

Another revision to the standard theory, which again was originally minor at the beginning, was developed by Fillmore (1968). In his pursuit of a deeper deep structure, he introduced a concept known as case within the base component. In this model a sentence is viewed as consisting of a Modality (M) and a Proposition (P), the former dealing with sentence modalities such as negation, mood, and aspects, and the latter "a tenseless set of relationships involving verbs and nouns" (p. 23). Central to this "case grammar" are the roles of the cases which determine the selection of the verb within a sentence.

As this model gained a great number of followers, it also showed an inadequacy to handle certain problems (Huddleston, 1970). This prompted Fillmore to revise his theory which resulted in a model substantially different from his previous one (Fillmore, 1971). The new model did not have a modality and a proposition, but, instead, "a predicator in construction with one or more entities each of these related to the predicator in one of the semantic functions known as (deep structure) cases" (p. 4). Since a predicator is mostly, although not always, a verb, his recent model can, therefore, be viewed as consisting of a verb plus a set of nouns each with its own case marker.<sup>2</sup> The fact that he does not mention the verb-case relationship compels me to assume that it is still the array of cases which determines the choice of the verb.

Although it is clear from the foregoing analysis that the role of semantics in linguistic theory is being given more serious thought than before, the real breakthrough in the use of semantics as a theoretical foundation comes from a man who was reared in the structuralist tradition - Wallace Chafe. He believes that since the function of language is to relate meaning to sound, he assumes that "at the heart of an adequate theory of language must be an adequate theory of semantic structure" (Chafe, 1971, p. 11). He further assumes that "the total human conceptual universe is dichotomized initially into two major areas" (p. 96), that is, the area of the verb, which embraces states and events, and the area of the noun, which embraces things. Contrary to the practice followed by some grammarians, such as Fillmore, Chafe considers the verb as central, and the noun peripheral. This necessarily means that it is the verb which determines what noun or nouns may or must co-occur, and not the reverse.

The way this semanticist theory operates can be summarised as follows: At the bottom of the whole matter is a semantic structure in which configurations of meanings are to be found. Through post-semantic processes,

which are similar to transformations, the configurations are transformed into a series of post-semantic representations which eventually lead to a surface structure. In order to reach the phonetic structure, the surface structure has yet to be subjected to several processes. This includes the symbolisation processes which convert the still semantically oriented surface structure into its underlying phonological structure. After undergoing a series of phonological processes, the surface structure will then appear in its phonetic form. See Figure 1 (simplified).

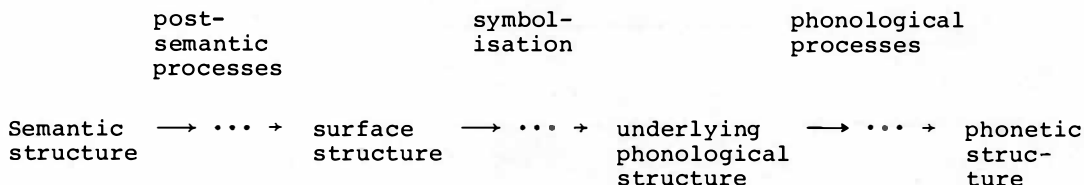


FIGURE 1

Each verb and noun under consideration is specified in terms of three semantic units: selectional, lexical, and inflectional. A selectional unit is intended to narrow down the areas of conceptual space. For a verb this includes state, process, action, and ambient. For a noun the selectional unit consists of count, potent, animate, human, unique, and feminine. After a noun or a verb has been selectionally specified, it reaches a point where no further specification is possible. For units of this kind the term lexical unit is used which appears ordinarily in the form of a root. The only specification a lexical unit may undergo is through derivation which includes among other things relative, inchoative, causative, for a verb, and countiser, anthropomorphiser, defeminiser for a noun.<sup>3</sup> The last semantic units are inflectional, which for a verb includes generic, perfective, progressive, past, obligational, etc. For a noun it may consist of units such as definite, generic, plural, unique, etc. In addition to these, Chafe postulates another inflectional unit which he calls "new". He assumes that when two persons are communicating, some of the information being communicated is new, and that there is at least one item which is old. The new information is being introduced to the hearer's mind for the first time, while the old information is shared, or assumed to be shared, by the speaker and the hearer.

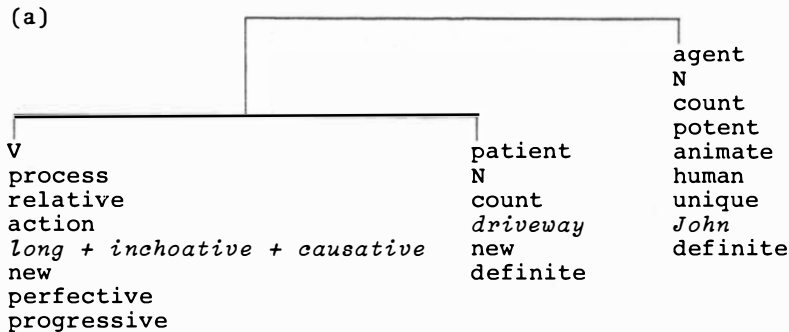
The centrality of the verb within a sentence brings about the verb-noun relations of the following types: patient, agent, experiencer, beneficiary, instrument, complement, and location. Unless a verb is specified as ambient, which indicates that the state is an all-encompassing

state with no reference to a particular thing within the environment, every sentence must have either a patient, an agent, or both. The occurrence of an instrument noun depends on the presence of an action-process verb. The other six are determined by the presence of a certain set of selectional units within the verb.

Applying these semantic units to the sentence

(1) *John has been lengthening the driveway.*

Chafe comes up with the following semantic structure (p. 240):



where the italicised elements are lexical units and those above and below them represent the selectional and inflectional units respectively.

In order to bring (a) to surface structure, a series of post-semantic processes must be applied. This involves subject and object formations, agreement, literalisation, linearisation, and deletions of various kind. After phonological processes have been utilised, (a) appears in the phonetic form of (1).

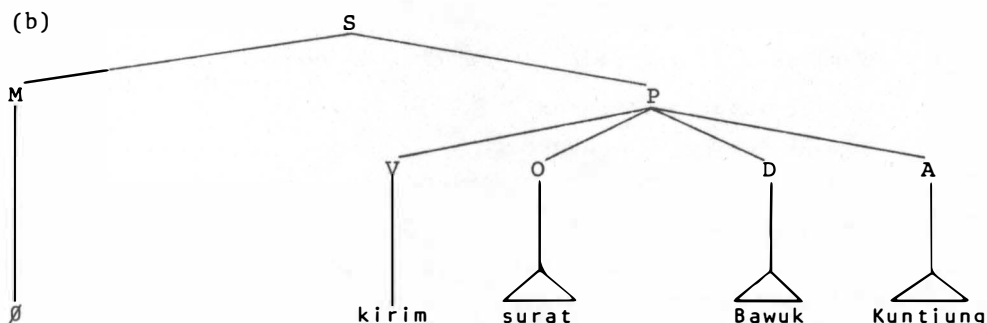
## 2. CENTRALITY OF THE VERB

This paper is an attempt to apply Chafe's theory of language. There are two main reasons why this particular theory is adopted. First, as a native speaker of Indonesian, I feel that using case array the way Fillmore does to determine what verb can or must occur in a sentence is counter-intuitive.<sup>4</sup> If a sentence such as

(2) \**Dia memarahi patung itu*  
*he angry statue the*

is to have a meaning at all, the possible interpretation would be to consider *patung itu* as animate, and not to assign a unique meaning to the verb *memarahi*. In this way, (2) means something like '*He is angry with the statue.*' This particularly becomes more obvious if we consider the whole verb set, because what determine the meaning of a sentence in

Indonesian are, in most cases, the verb affixes such as the -i of *memarahi* above. Secondly, even if these affixes can be introduced transformationally as Fillmore has suggested (1968, p. 30, footnote 39), there is still a problem in Indonesian in cases where the case frame has /\_\_\_ O + D + A/, since the choice of any one of the three to become the surface subject has not given us a sufficient indication as to what particular affix must be chosen for the verb. Given the concept of 'a man by the name of *Kuntjung*, sending a letter to a girl named *Bawuk*' we cannot produce a well-formed sentence unless we also know, in addition to the A becoming the subject, what exactly comes after the verb - *Bawuk* or the letter - because this very choice determines the affixes that occur with the verb root. If this sentence is diagrammed as in (b)<sup>5</sup>,



this means that in order to form a sentence with the meaning intended, we must select not only A as the subject but also either D or O for the object that comes immediately after the verb. If A and O are chosen, we have

- (3) *Kuntjung mengirim surat kepada Bawuk.*  
*send letter to Bawuk*

If, on the other hand, A and D are chosen, we have

- (4) *Kuntjung mengirimi Bawuk surat.*  
*send letter*

The corollary of this constraint makes (5) and (6) unacceptable:

- (5) \**Kuntjung mengirimi surat kepada Bawuk.*  
 (6) \**Kuntjung mengirimkan Bawuk surat.*

We see, therefore, that the well-formedness of the sentences resulting from (b) is determined not by one, but by two cases. It is thus clear that using the noun to determine the verb brings about many problems for Indonesian.

## 2.1. THE PROBLEMS

The problems under investigation are only a sample from a general phenomenon which involves a subtle distinction of meanings on one hand, and their phonological representations on the other. The verbs chosen are *datang* with all its derived forms. Consider now the sentences

- (7) *Penari itu akan datang.*  
*dancer the will come*  
*'The dancer will come.'*
- (8) *Gombloh akan mendatangkan penari.*  
*will come dancer*  
*'Gombloh will bring in a dancer.'*
- (9) *Gombloh akan mendatangi penari.*  
*will come dancer*  
*'Gombloh will approach (physically) a dancer.'*
- (10) *Gombloh kedatangan penari.*  
*come dancer*  
*'Gombloh got visited by a dancer.'*<sup>6</sup>

While (7) does not pose a serious semantic problem, (8) through (10) require a careful attention in the following way: in (8) the subject<sup>7</sup> acts as an instigator which then results in the coming of the object; the object, therefore, is actually the actor of the coming; the method of instigating the action by the subject is not relevant; if he chooses to send a ticket, for instance, rather than picking her/him up, he virtually remains at his place. In (9) the subject is the actor that performs the act of coming, while the object becomes the locus toward which the coming of the subject is directed. He is, therefore, stationary. The method of approaching is also insignificant. In (10) the subject undergoes the coming of the object. The nature of coming itself is, from the point of view of the subject, unexpected and adversative. This is why, perhaps, the future aspect *akan* does not occur.<sup>8</sup>

The semantic differences among (7-10) are indicated by the presence or absence of the prefixes *men-* and *ke-*, and the suffixes *-kan* and *-i* in the verb root *datang*.

## 2.2. THE ANALYSIS

What I am trying to do here is to present the semantic analysis of these verbs and how they are brought up into the surface structure. I will, therefore, concern myself only with the semantic structure and the post-semantic processes, leaving out the symbolisation and phonological processes that convert the surface structure into its phonetic

form.

Since in this paper the verb is considered central, I will begin with the verb *datang*. As *datang* expresses an activity rather than a state of being, and also the fact that (7) answers to questions such as '*what will N do?*' rather than '*what will happen to N?*', it seems appropriate to consider this verb as an action verb. Given this initial element, the first rule in the generation of (7) is

S-1: V —  $\rightarrow$  action

where S refers to semantic rules, the broken shaft to "optional", and the double head to "is further specified as". The structure resulting from S-1 is (c):

(c) V  
action

Since no further selectional specification is possible, a lexical unit is now reached in the form of a root through S-2 with the structure given in (d):

S-2: action  $\rightarrow$  datang

(d) V  
action  
datang

The unbroken shaft indicates that the rule is obligatory.

At this point we can see that of the two elements, noun and verb, in (7), the information that the speaker and the hearer share is the concept of the agent, whereas the activity itself is new from the hearer's point of view. The inflectional unit "new" can, therefore, be attached only to the verb as in S-3, giving us the structure (e):

S-3: V  $\rightarrow$  V  
root root  
new

(e) V  
action  
datang  
new

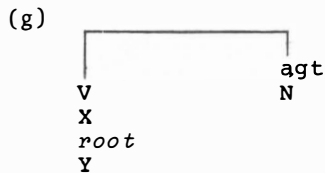
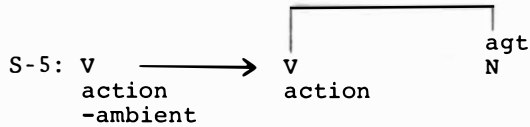
Finally, the inflectional unit which indicates aspectual future must be indicated. Using S-4 below,

S-4: V —  $\rightarrow$  future  
action

we have the complete structure of the verb in (7):

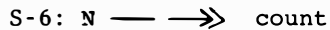
(f) V  
action  
datang  
new  
future

The semantic units found in (f) call forth the presence of a certain noun. Since the verb is a non-ambient action, it must be accompanied by an agent, introduced by rule S-5, which results in the structure (g). The single arrow head means "becomes".

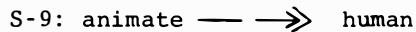
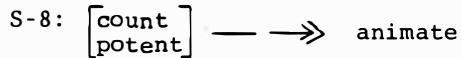
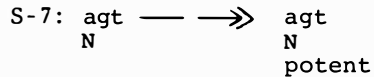


X and Y indicate the selectional and inflectional units respectively.

Looking at the agent itself, we see that it has units of its own, the first of which is "count".



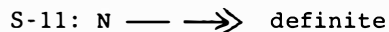
Additional selectional units are added by the rules -



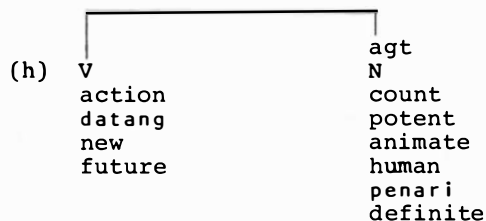
The lexical unit is introduced by rule S-10:



The only inflectional unit required is the unit definite.



Incorporating the last six rules and the verb structure of (f), we have the following semantic structure for (7):





As mentioned in 2.1., the verb *mendatangkan* in sentence (8) requires two nouns. While all the units of the verb given in (f) are applicable, two more important units must be added. Since *mendatangkan* expresses not only what someone, the agent, does, but also involves a change in condition of someone or something else, the patient, it must therefore be specified not only as action but also as process. A revision to rule S-1 can be presented as S-1a:

S-1a: V — —> (process)  
           -state            (action)

where ( ) indicates an inclusive disjunction.

The second rule that we must add is a consequence coming out from S-1a, that is, in the event that both process and action are selected, the root must be derivationally specified as causative. This rule must be introduced after a lexical unit has been inserted, that is, after S-2.

S-2a: V — —> V  
       process        process  
       action        action  
       root            root + *causative*

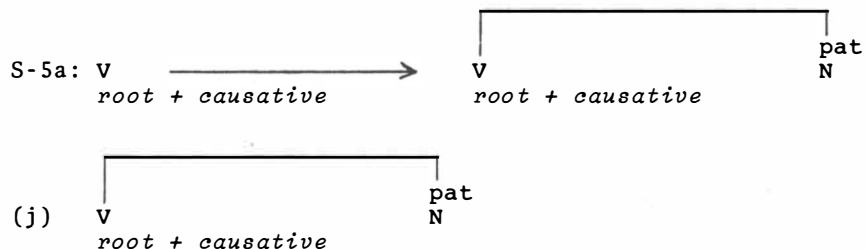
After modifying S-4 to become S-4a, needed to account for the fact that not only action, but also action-process verbs can be specified as future,

S-4a: V — —>> future  
       process  
       action

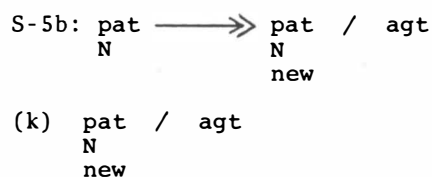
and applying S-1a, S-2, S-2a, S-3, and S-4a, we have the following structure for *mendatangkan*:

(i) V  
       process  
       action  
       *datang + causative*  
       new  
       future

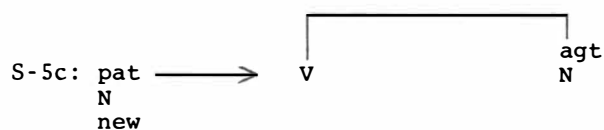
Turning now to the relation between the verb and its nouns, we see that while it is true that the noun preceding the verb, which I will call the first noun, is the one that instigates the action which then brings about the coming of the noun following the verb, the second noun, the latter is itself the actual performer of the coming. Our rules must account for this fact. My suggestion is as follows. The causative verb in (i) must be rewritten as (j) through rule S-5a (only relevant elements are given):



The patient must be further specified to indicate that it is he who does the coming. Before we can do this, however, we must specify pat as new due to the presence of an agent. Rule S-5b introduces new, resulting in (k). The slant line is to be read "in the environment of".



Now to indicate that pat in (k) is an actor, we need rule S-5c.

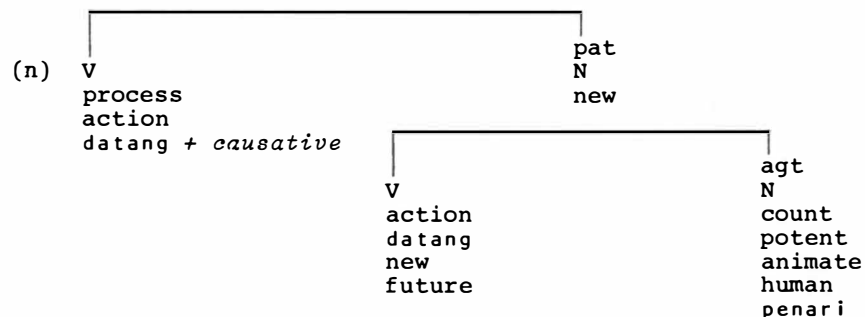


The resulting structure is (m):

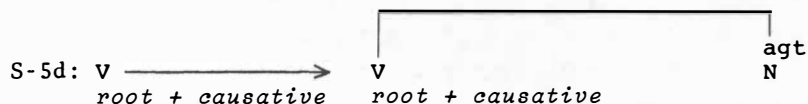


The structure of the verb in (m) is the same as that of (f). Except for the absence of "definite", the structure of the agent in (m) is the same as that given in (h).

Putting together S-5a through S-5c, we have the following structure:



The agent that instigates the action is introduced by rule S-5d.

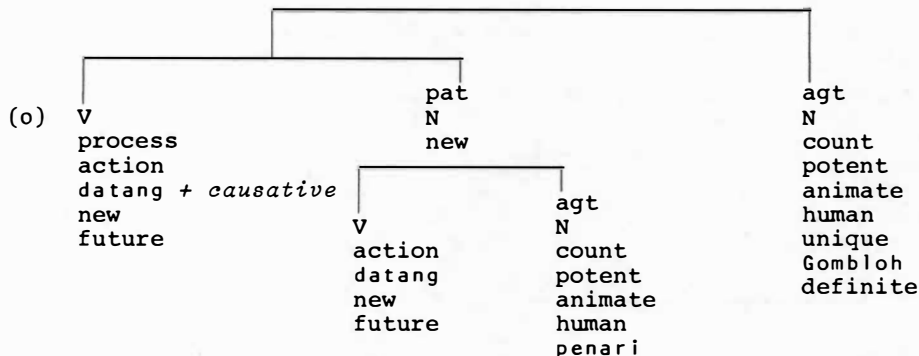


The agent differs from that in (h) only in that the former has a selectional unit "unique" which must be introduced before S-10, called S-9a. This is then followed by lexicalisation rule S-9b.

S-9a: human  $\longrightarrow$  unique

S-9b: [unique  
-feminine]  $\longrightarrow$  Gombloh

The result of applying S-6 through S-9b, combined with (n) produces the complete structure of (8).



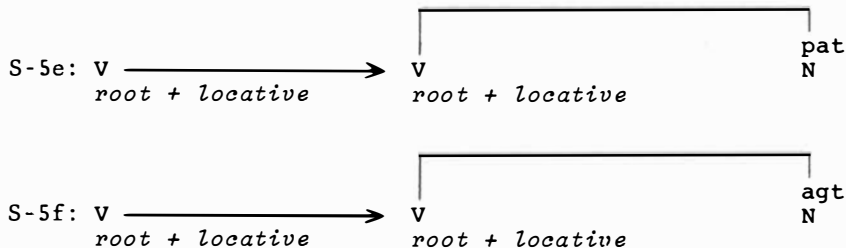
If we now compare (8) and (9), we see that the difference in meaning must be attributed to the verb structures, since other elements are constant. As it turns out, while *mendatangkan* has a causative derivational unit, *mendatangi* is locative, indicating the locus toward whom the action is directed. The difference between the rule to cover this fact and that given in S-2a lies only in the manifestation of the derivational unit. We can, therefore, propose to add to S-2a the locative unit, which then changes S-2a to -

S-2a: V  $\longrightarrow$  V  
           process      process  
           action      action  
           *root*      *root +* { *causative* }  
   { *locative* }

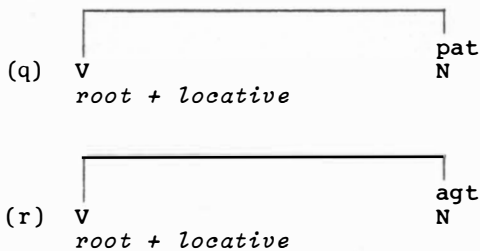
where the parentheses state an exclusive disjunction. If locative is chosen, the structure looks like (p):

(p) V  
       process  
       action  
       datang + *locative*  
       new  
       future

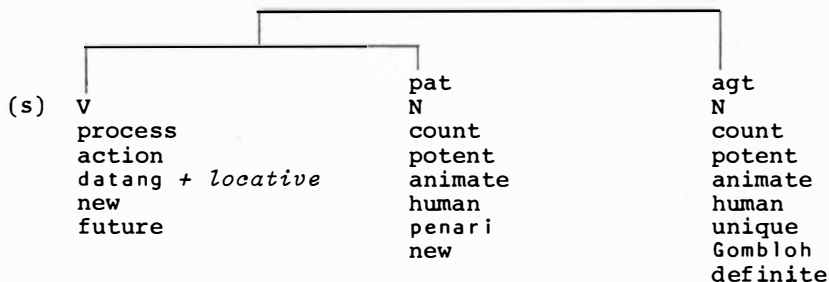
The semantic structure of the agent in (9) is the same as that of the right-most agent in (o). The structure of the patient is much less complicated than that in (o), since this patient does not have a double function. Two more additional rules are required that would generate the patient and the agent. Although these rules are parallel to S-5a and S-5d, they are different in that they produce nouns which are functionally different from the previous ones. We will number them S-5e for the patient, and S-5f for the agent.



The structures produced by these rules are -



If to the patient (q) we now apply rules S-6 through S-9, S-10, and S-5b, and to the agent of (r) rules S-6 through S-9a, plus S-11, we have a complete structure of (9):



Let us now look at sentence (10). As is the case with *mendatangkan* of (8), *kedatangan* in (10) also has a double function, although in a different manner. In relation to its first noun, the verb is a state verb. The noun is in a state of being, rather than a process or action.

In relation to its second noun, however, it does indicate an action, since this noun performs the action of coming. Since a state verb permits only a patient, and an action verb only an agent, the first and the second nouns in (10) must be considered patient and agent respectively. The situation where a patient occurs before and an agent after the verb is known as passive. I would call this kind of passive in (10) adversative. Contrary to what Chafe would do, I would consider adversative as derivational, rather than inflectional. Before a rule that would cover this fact can be introduced, S-5g, we need another rule that would generate an action-state verb. I will order this rule after S-1a, and call it S-1b.

S-1b: V  $\longrightarrow$   $\Rightarrow$  [action  
state]

Rule S-5g below states that the root of an action-state verb must be specified as adversative:

S-5g: V  $\longrightarrow$  V  
       action                    action  
       state                     state  
       root                      root + *adversative*

The unit action in the verb requires an agent. Rule S-5h would generate structure (t):

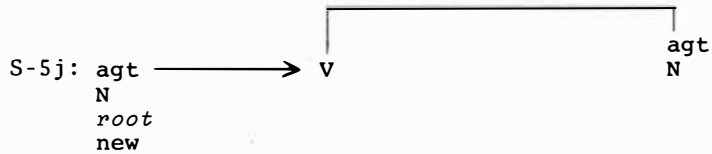
S-5h: V  $\longrightarrow$  V  
X  
root + *adversative*  
Y V  
X  
root + *adversative*  
Y agt  
N

(t) V  
X  
root + *adversative*  
Y agt  
N

Since the concept of the agent in (t) bears an old information only to the speaker, but not to the hearer, it must be specified as new:

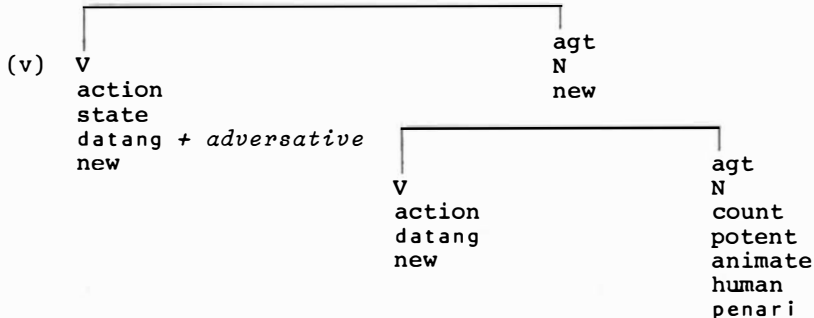
S-5i: agt  $\longrightarrow$   $\Rightarrow$  agt / V  
       N                    N     /    V  
       root                 root    /    root + *adversative*  
                               new

To indicate that agt in (t) is an agent only with respect to the root datang, we need rule S-5j.

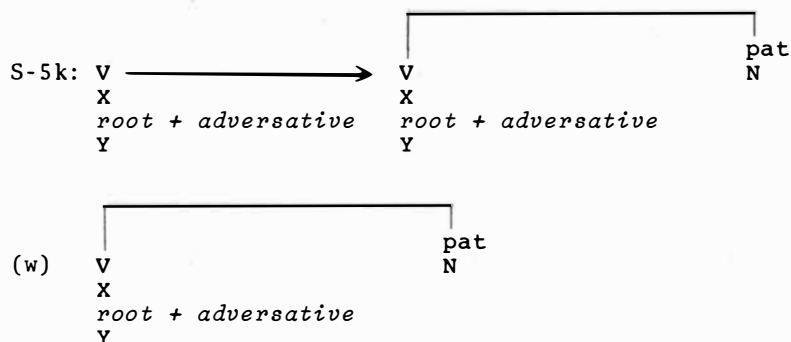


Except for the absence of "definite" in the above agt, and the unit future in the verb, the structure of (u) is the same as that of (h).

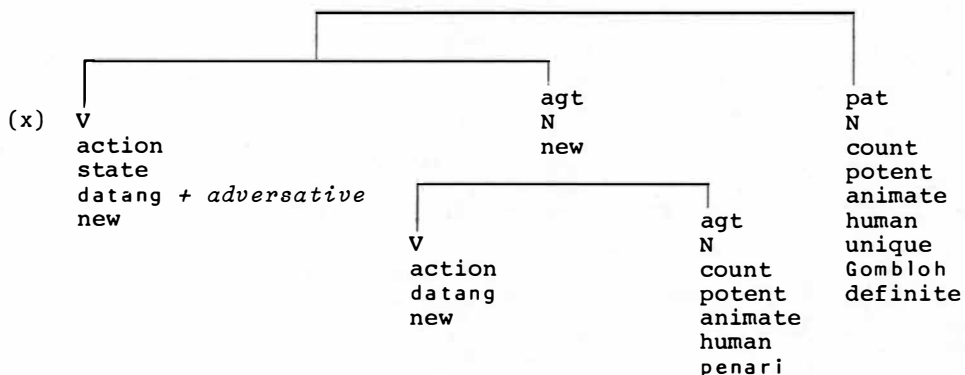
Putting together S-1b, S-2, S-3, and S-5g for the adversative verb, S-5h and S-5i for the agent marked "new", and then reproducing structure (h) minus the "definite" and "future", we have the following structure:



The adversative verb also requires a patient. This is introduced by rule S-5k, resulting in structure (w).



The semantic units of pat in (w) include those specified for the right-most agt of (o). Adding this patient with its semantic units to (v) produces the structure of (10):



What we see from (x) is that in the event that the verb bears an adversative unit, the patient is not specified as new and becomes the subject on the surface structure. The agent, on the other hand, must be specified as new and becomes the surface object. There has been a long standing argument among Indonesian scholars whether sentences such as (10) are passive or not. The structure in (x) seems to justify our affirmative position.

If we now compare the semantic structures of (h), (o), (s), and (x), we see several interesting phenomena. First, while the verb in (h) is an action verb, those in (o) and (s) are process-action, and that in (x) is action-state. The last three are related to the first in that they all share one and the same root. The difference among them lies only in the derivational units that each takes. Second, the different semantic units in each verb call forth a different set of nouns which may accompany them. These nouns have different semantic functions. Third, while the structures of (h) and (s) are simple, those of (o) and (x) are complex in that they involve "sub-branching". And fourth, the only nodes that bear no unit "new" are the right-most nodes.

With the semantic structures of (7) through (10) now complete, we can proceed to the post-semantic processes which eventually lead to the surface structures. For the purpose of illustration, I will take only sentence (8) with its structure given in (o).

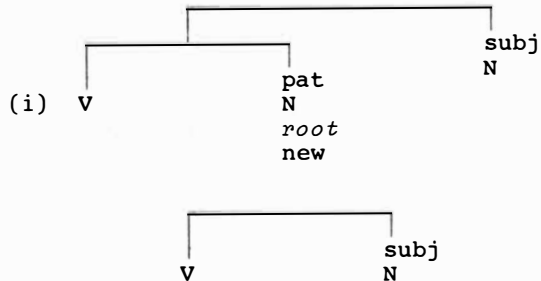
The first step is to form the subject and object. Two criteria are used, namely, the semantic interrelation between the agent and the patient, and the distribution of "new" to the nouns.

The following rules will cover the subject formation for (8) as well as for the other three sentences.

T-1a: agt → subj  
 N  
 root  
 -new

T-1b: pat → subj / V  
 adversative

The T indicates that it is a post-semantic process. Applying T-1a to (o), we have the following:

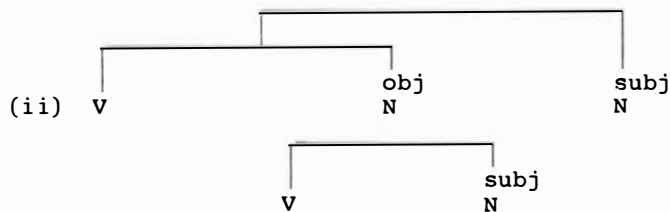


The object is formed by the following rules:

T-2a: agt → obj  
 N  
 root  
 new

T-2b: pat → obj  
 N  
 root  
 new

Applying T-2b to (i), we come up with the following structure:



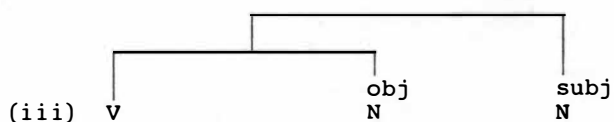
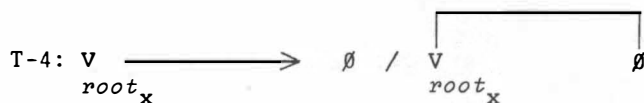
Since the subject of the "embedded" sentence is the object of the main sentence, the former must be raised to coalesce with the object. This is required also because of the fact that the embedded verb and the main verb share the same lexical unit.

T-3: subj → obj / V subj  
 root<sub>x</sub>

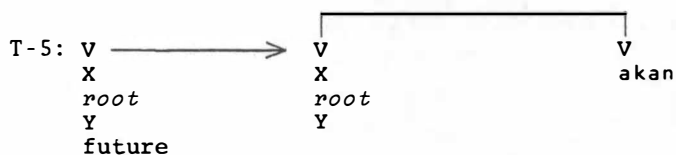
The subscript X indicates a certain co-reference. This process leaves



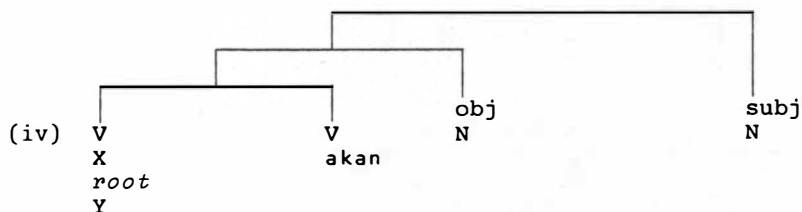
the verb without a noun, which can therefore be pruned. The result is (iii):



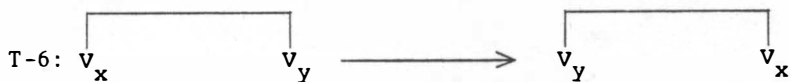
A set of processes called literalisation then follows. These are used to literalise the units under each noun and verb. Since in our case this is only applicable to the verb with its aspectual future, we need only the following rule:



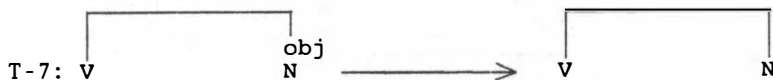
The resulting structure is (iv).



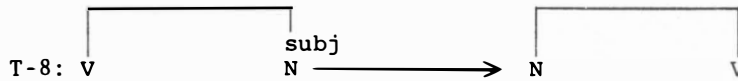
We need now to linearise these elements. There are two kinds: primary linearisation, which involves the ordering of the nouns and the verbs with relation to each other, and secondary linearisation, which deals with the ordering of the elements within each noun and verb. Starting from the left-most element with rule



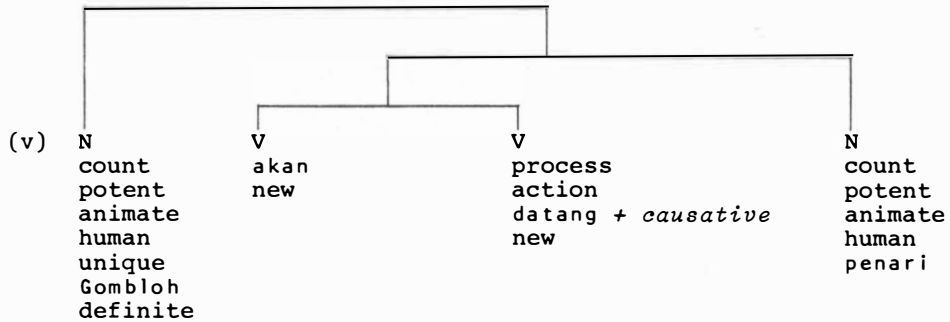
proceeding to state that an object noun follows its verb,



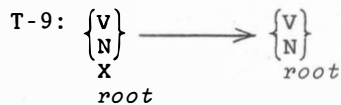
and finally stating that a subject precedes its verb,



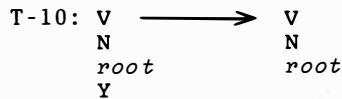
we arrive at the diagram below with all its semantic units added:



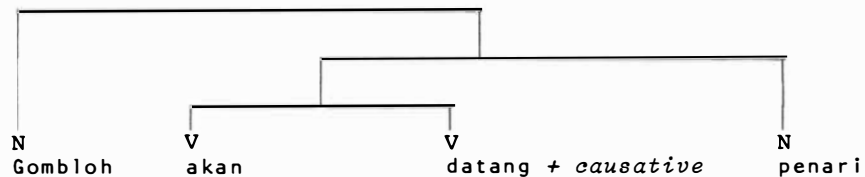
As mentioned earlier the necessity of listing the selectional units is motivated by the fact that it is through this process that we can arrive at a lexical unit. Therefore, once the lexical unit has been obtained, these selectional units become redundant. Through a process of deletion they can then be erased:



The inflectional units must also be deleted on the same ground.



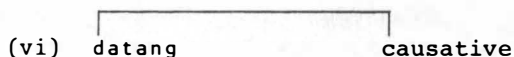
The application of T-9 and T-10 leaves us with the following structure:



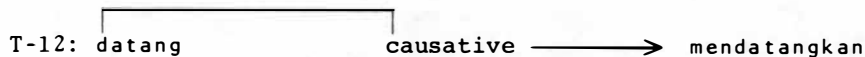
At this point, secondary linearisation can be applied. The only one we have is that which involves the linearisation of the derivational unit causative. Using X as a derivational unit, we have



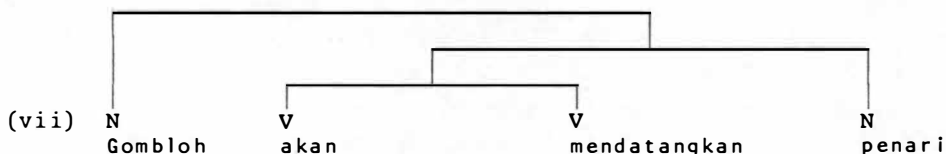
which, in our case, gives us



Specifying the causative further, which in Indonesian is indicated by the prefix *meN-* and the suffix *-kan* for an active sentence, we have



After specifying these last two rules, we finally have the following surface structure:



Before concluding, I must mention that the verb *mendatang* has been purposely left out from our discussion. The reason, quite frankly, is that I am not sure what the structure of this verb is like. It seems that *mendatang*, which means something like '*to perform a coming*', is the only example in the language where the prefix *meN-* can be added to a root which itself is an action verb. This is quite unusual for Indonesian, because what usually happens is that, if the *meN-* is to be added to a root, which itself is a full verb, this root is always a state root. The addition of the prefix changes it into, in most cases, a process. Secondly, the noun that occurs with *mendatang* is restricted to a very few words, the most common of which is the noun *ombak* '*wave*' as in *Ombak mendatang* '*The wave comes.*' Even here, the usage is restricted to poetic or other literary works. Thirdly, while in any process or action verb, the activity is usually motivated by a volition of some sort, this phenomenon seems lacking in the verb *mendatang*. To put it rather loosely, the noun, be it a patient or an agent, seems to lack the unit "potent" the way this term is ordinarily used. The sentence *Ombak mendatang*, therefore, means, more accurately, something like '*By the course of nature, the wave then slowly comes (to the shore).*' Finally, the other common usage of *mendatang* is in connection with temporal nouns, such as year, month, week, etc. In this usage, however, the relation between the verb and its noun is not of the "subject-predicate" but of the "modified-modifier" type. Thus *bulan mendatang* means '*the coming month*'.

### 3. CONCLUSION

From the foregoing analysis we can see that there are two problems

involved: meaning and form. From the meaning point of view, we have discovered that the problem lies in the units which make up the semantic structures of the verbs and the nouns, and their functional relations with each other. From the point of view of form, the problem involves a correct choice of affixes as determined by the semantic structures.<sup>9</sup>

The theoretical implication of the present analysis is that from the semantic point of view the concepts of coming, bringing in, and approaching are not only interrelated, but of such a nature that one is derivable from the other. It, therefore, may throw light upon the question whether the English verbs *bring* is the causative form of *come* (Binnick, 1971), *kill* the causative form of *die* (Lakoff, 1970), etc.

The pedagogical implication is that students learning Indonesian must be "made to see" the interrelatedness of the concepts and the derivability of the concepts and their phonological forms. Otherwise, an embarrassing situation could very easily arise, where a student, wanting to say that he will put someone to bed, comes up with the sentence

(11) Saja may meniduri gadis itu.

*I want sleep girl the*

where the verb *meniduri* is locatively derived from the root *tidur* 'sleep', which, therefore, gives the meaning '*I want to sleep on the girl = I want to make love to the girl.*'

## NOTES

1. The verbs under consideration are *datang* and its derived forms *mendatangkan*, *mendatangi*, *kedatangan*, and later *mendatang*. Although these are the only verbs analysed, it seems that many of the rules given are applicable to other verbs of similar nature.
2. In his 1971 model, Fillmore does not include the diagram to replace his old one.
3. Only the units which are of some relevance are given here. For further units, see Chafe's *Meaning and the Structure of Language*.
4. I have indicated this view in my earlier paper. See Dardjowidjojo, "The Men-, meN-kan, and meN-i Verbs in Indonesian" in *Philippine Journal of Linguistics*, 1971.
5. The use of this 1968 diagram is due to the fact that Fillmore does not have a diagram for his new model.
6. As it stands, a meaning cannot be really assigned to sentence (10) as it lacks temporal markers or any other time indicators. The translation given assumes a previous context.
7. The terms subject and object are surface structure terms.
8. I must point out that under a very remote circumstance, I would not be very surprised to find an adversative verb having aspectual markers such as *akan* given here.
9. This problem arises only in those who are learning the language as a foreign language.

## REFERENCES

- BINNICK, Robert I.  
1971 "Bring and Come", *Linguistic Inquiry*, vol. II, No. 2, pp. 260-265.
- CHAFE, Wallace L.  
1970 *Meaning and the Structure of Language*. University of Chicago Press, Chicago.
- CHOMSKY, Noam  
1955 "Semantic Considerations in Grammar", *Georgetown University Monograph Series on Languages and Linguistics*. Georgetown University Press, Washington, D.C.  
1957 *Syntactic Structures*. Mouton & Co., The Hague.  
1965 *Aspects of the Theory of Syntax*. The M.I.T. Press, Massachusetts.
- DARDJOWIDJOJO, Soenjono  
1971 "The meN-, meN-kan, and meN-i Verbs in Indonesian", *Philippine Journal of Linguistics*, vol. I, No. 1.
- FILLMORE, Charles J.  
1968 "The Case for Case", *Universals in Linguistic Theory*, Editors Emmon Bach and Robert T. Harms. Holt, Rinehart and Winston, Inc., New York.  
1971 "Some Problems for Case Grammar", *Working Paper in Linguistics*, the Ohio State University.
- HUDDLESTON, Rodney  
1970 "Some Remarks on Case-Grammar", *Linguistic Inquiry*, vol. I, No. 4.

KATZ, Jerrold J. and Jerry A. Fodor

1963 "The Structure of a Semantic Theory", *Language*, vol. 39.

KATZ, Jerrold J. and Paul M. Postal

1964 *An Integrated Theory of Linguistic Descriptions*. The M.I.T. Press, Massachusetts.

LAKOFF, George

1970 *Irregularity in Syntax*. Holt, Rinehart and Winston, Inc., New York.





## ON MORPHOLOGICAL AND SYNTACTIC RELATIONS IN A SOUTHEAST ASIAN LANGUAGE\*

CESAR A. HIDALGO

### 0. INTRODUCTION

In Charles Fillmore's paper on the features of a universal base, he argues for a universal system of deep-structure cases, the explanatory value of which is of a syntactic nature. He says:

The various permitted arrays of distinct cases occurring in simple sentences express a notion of "sentence type" that may be expected to have universal validity independently of such superficial differences as subject selection. The array of cases defining the sentence types of a language have the effect of imposing a classification of the verbs in the language (according to the sentence type into which they may be inserted), and it is very likely that many aspects of this classification will be universally valid.<sup>1</sup>

In our study of Ivatan,<sup>2</sup> using a tagmemic model, simple sentences (clauses) were typed according to clause expansion features or in Fillmorean terms, according to arrays of cases occurring in simple sentences. It was found out that the different sentence types correlate with various predicative types. Specifically, the different cases in the sentences are related to affixes in the predicative. It was also found out that not all affixes occur with all types of predicatives. The potential affixes in a predicative are restricted by stem type and other occurring affixes. This close relation between syntax (i.e. the array of cases occurring in a simple sentence) and the morphological structure of the predicative underscores the need for a closer scrutiny

---

\*Research on Ivatan (one of the Philippine languages) has been supported in part by grants from the Social Science Research Council, University of the Philippines, the Linguistic Society of the Philippines, and the Rockefeller Foundation. Paper read at the 28 International Congress of Orientalists, Canberra, January, 1971.

of predicative morphology, for a better understanding of Ivatan syntax and, presumably, the syntax of the other Philippine languages. This study of predicative morphology (in consonance with syntax) is especially relevant when we take into consideration the fact that a number of cases or in tagmemic terms clause nuclear tagmemes in Ivatan do not obligatorily occur. Their implicit presence is signalled by a predicative affix. Consider:

1. Naparutung si ina su manuk.  
     P                  S          O  
     *caused-cook fm mother fm chicken*  
     '*Mother caused (someone) to cook chicken.*'
  
2. Ipanutung ya.  
     P                  T/A  
     *cook-with dem*  
     '*(Someone) cooks with this.*'
  
3. Umuyug u ranum.  
     P                  T/S  
     *flow fm water*  
     '*Water flows.*'

In 1, the indirect object (Y) does not occur overtly. It is known to occur implicitly between speaker and hearer and may occur explicitly because of the causative affix {pa-<sub>1</sub>} in the predicative *naparutung* '*caused-cook*'. In the case of 2 and 3 the surface structures appear to be similar (i.e. both sentences consist of two elements, a predicate and a topic), but the underlying structures of the two are quite different. In 2, the structure of the predicative signals that the topic manifested by *ya* '*this*' is associative instrument ({ipa-<sub>1</sub>}). The predicative also signals the potential occurrence in the sentence of the subject and object functions. In 2, the speaker-hearer knows that these functions implicitly occur. In 3, the focus affix which is also an intransitiviser {um-} indicates plainly that no object function can potentially occur.

This paper attempts to present the relation between syntax and the morphological structure of the Ivatan predicative.

## 1. THE PREDICATIVE STEM

The predicative is a class of words that manifest features such as focus, tense, voice, or aspect (which are absent in the other form classes) and that fill the predicate slot of clauses other than the equational clause. It may, however, manifest non-predicate tagmemes on the

clause level generally manifested by nominals. Except for a few adjectival predicatives, it consists of at least two morphemes - a stem and an affix and may potentially consist of a stem and layers of affixes, including the suprasegmental morpheme stress or length.

The predicative stem may be simple or derived. Ivatan has nine types of simple stems (basically the verb stem, numeral stem, adjective stem, and common noun stem) which are formally distinctive on the basis of affixational potential. Affixes considered in categorising the simple stems are those that have bearing on syntax. These include both focus and voice affixes. Classifying, then, the simple stem types by their syntactically relevant affixational potential, they are thus categorised by the potential syntactic constructions into which they may participate.

### 1.1 THE SIMPLE STEM

The nine simple stem types may be labelled as follows: 1) verb stem 1 (vs1) which is generally an inherently transitive verb stem, i.e. it may take the transitive voice affix {N-}<sup>3</sup> without having to be affixed with a derivational affix that will then allow the prefixation of {N-}; 2) verb stem 2 (vs2) which is generally an inherently intransitive verb stem; 3) noun stem 1 (nsl) which is a concrete noun such as *amung* 'fish'; 4) noun stem 2 (ns2) which is an abstract noun pertaining to an emotion such as *adaw* 'love', *amu* 'fear'; 5) noun stem 3 (ns3) which is a concrete noun pertaining specifically to meteorological conditions like *chimuy* 'rain', *chidat* 'lightning'; 6) noun stem 4 (ns4) which is an abstract noun pertaining to a quality such as *avid* 'beauty', *pya* 'goodness', *karang* 'tallness'; 7) noun stem 5 (ns5) which indicates the period of the day and which may either be a bound form or a free form, e.g. *-kuyab* 'afternoon' and *ahep* 'night'; 8) numeral stem (num) such as *pitu* 'seven', *wahu* 'eight'; and 9) adjectival stem (adjs)<sup>4</sup> such as *dekey* 'small', *aru* 'many', *rakuh* 'big', *vayu* 'new'.

### 1.2 THE DERIVED STEM

The derived stem consists of a core manifested by either a nsl, ns3, or ns4 root and a derivational affix or consists of a nsl plus a pluraliser. The derivational affix is limited to a morphologically functional affix which increases or decreases the focus or voice affixational potential of the simple stem. Unlike the voice affixes which also increase or decrease the focus affixational potential of the simple stem, the predicative derivational affix is not syntactically relevant in signalling potential slots in syntax as the voice affixes do. Consider:

4. adaw 'love'  $\left\{ \begin{array}{l} \text{madaw (Sf)} \\ \text{ichadaw (Af)} \end{array} \right.$  versus -adadaw  $\left\{ \begin{array}{l} \text{mangadadaw (Sf)} \\ \text{adadawen (Of)} \\ \text{ipangadadaw (Af)} \\ \text{pangadadawan (Rf)} \end{array} \right.$

This simple noun stem 2, adaw, is inflectable for only two non-predicate focuses: subject and associative and thus the syntactic construction into which adaw as a predicative stem enters are unexpandable beyond the subject and associative slots in terms of nuclear clause tagmemes. The derived stem<sup>5</sup> -adadaw, however, may be inflected for subject, object, associative, and referent focuses so that the syntactic constructions into which -adadaw enters into are more complex in terms of potential clause expansion. For a case of a decrease, consider:

5. avid 'beauty'  $\left\{ \begin{array}{l} \text{makavid (Sf)} \\ \text{kaviren (Of)} \\ \text{ipakavid (Af)} \\ \text{pakaviran (Rf)} \end{array} \right.$  versus -avyavid  $\left\{ \begin{array}{l} \text{mayavyavid (Sf)} \\ \text{ipayavyavid (Af)} \\ \text{payavyaviran (Rf)} \end{array} \right.$

In the case of avid, ns4, the derivational affix decreases the affixational potential of the stem. The simple stem avid may be affixed with the augmentative {ipa-<sub>2</sub>} or the possessive affix {ka-<sub>3</sub>} which makes it possible for the stem to be inflected for object focus and thus allows the predicative to enter into a transitive construction. If avid were affixed with a derivational affix, the resulting stem -avyavid cannot be affixed with {ipa-<sub>2</sub>} or {ka-<sub>3</sub>} and the predicative with this stem participates only in an intransitive construction. The augmentative affix {ipa-<sub>2</sub>} is an inner affix whereas the associative focus affix {ipa-<sub>1</sub>} is an outermost affix.

The predicative derivational affix is a reduplicative affix (usually the first or first two syllables of the root are reduplicated). Affixation of the derivational affix to the noun root generally results in a bound form, e.g. -sasalawsaw from salawsaw 'wind'. A case when the non-simple noun stem does not result in a bound form is when a semantic change attends the stem change as in the non-simple stem conveying the diminutive, e.g. vavahay 'playhouse'.

In predicativising a stem, the affixation of the derivational affix to the noun root results in a meaning different from the predicative the stem of which is simple, e.g. maybaka 'raise cattle' versus maybabaka 'crawl on all fours'. This change in meaning is also evident when a plural noun is predicativised. When kusikusina 'kitchens' is predicativised forming maykusikusina 'do the things related to the kitchen', kusikusina 'kitchens' becomes singular and the notion of plurality is attached to the activities connected with the kitchen. Or

consider *mayvahay* 'play housekeeping' with the stem *vavahay* 'playhouse' and *mayvahavahay* 'do the activities connected with the house' with the stem *vahavahay* 'houses'. The simple noun stems that may be affixed with a derivational affix are *ns1*, *ns2*, *ns3*, and *ns4*. The derived noun stems (*dns*), then, are *dns1*, *dns2*, *dns3*, and *dns4*. *dns1* is further subclassified into *dlns1* and *d2ns1* (the former exhibiting a reduplication of only the first syllable of the simple stem and the latter exhibiting a reduplication of the first two syllables of the simple stem accompanied by meaning difference, e.g. *maylalamit* 'play with cloth remnant' versus *maylamilamit* 'work with clothes'). A difference in the location of the primary stress plus meaning difference occur in *dlns1*, hence *dlans1* and *dlbns1*, e.g. *-lalamit* from *lamit* 'cloth' and *lalamit* 'cloth remnant'.

## 2. THE PREDICATIVE AFFIXES

The predicative affixes are categorised in terms of form into: non-reduplicative affixes, reduplicative affixes, and suprafixes. In terms of function, the predicative affixes may be generally classified into two: the morphologically functional and the syntactically and/or semantically functional affixes.

The morphologically functional affixes do not give any semantic modification to the predicative stem nor are they directly relevant to syntactic structures. They function either to allow or restrict focus and/or voice affixes that may be affixed to the stem or to make distinct morphological distributions thereby making distinct semantic features of homophonous affixes.

The syntactically and/or semantically functional affixes may be subdivided into aspect and non-aspect affixes. The aspect affixes may be subcategorised into action aspect affixes (e.g. augmentative, distributive, abilitive or accidental, frequentative, repetitive) and time aspect affixes (e.g. habitual, punctiliar, durative, inceptive, continuative). Aspect affixes semantically modify the predicative stem but are not syntactically relevant in terms of clause expansion. The distributive aspect, however, indicates a plural subject or object, depending on the predicative focus. The action aspect affixes are inner affixes while the time aspect affixes occur in outermost position. The durative aspect suffix {-an<sub>3</sub>}, however, occurs in an inner position if the focus affix is a suffix as in *paychamahpanan* 'is done all night'. The non-aspect affixes are syntactically relevant in terms of clause expansion. They are subcategorised into focus affixes and voice affixes. The focus affixes occur in outermost position and the voice affixes in inner position. A detailed discussion of the various affixes is not possible in

this paper. Consider, then, the general scheme of Ivatan predicative affixes in terms of function as shown in the configuration as a summary of the Ivatan affixes (on next page).<sup>6</sup>

## 2.1 THE STEM AND FOCUS AFFIXES

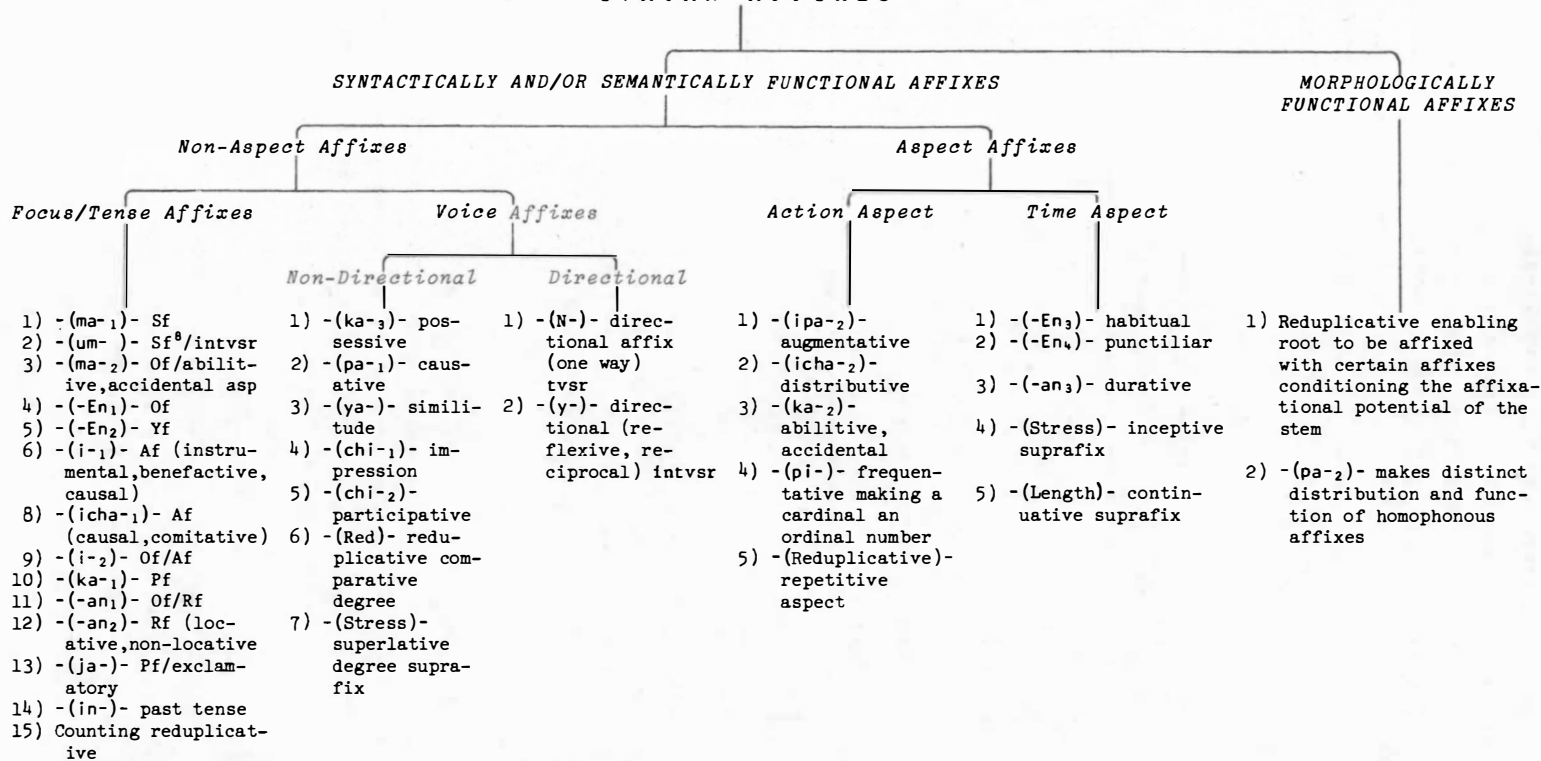
Focus refers to a relationship between the predicate and a non-predicate clause level tagmeme, the topic, where the basic function (case use) of the topic is signalled formally by an affix (the focus affix) in the predicative manifesting the predicate. The *topic*<sup>7</sup> can be any of the clause level tagmemes: subject, object, associative, referent, indirect object, object/referent, or object/associative. Each of these tagmemes or functions (considered here as deep structure cases) are marked by certain particles, the function markers. When any of them functions at the same time as topic, the function marker signalling its base function is replaced by a topic marker (e.g. *u* before common nouns) and the basic function of the topic is signalled by the predicate. Such tagmeme functioning as topic is then said to be focussed. Inflecting, then, the predicative for focus is a device for identifying which of the clause constituents has been singled out as topic. The focus affix indicates the logical function (the deep structure function) of the noun phrase serving as topic. For instance, the focus affix *-en* in *ahapen* 'get' indicates that the logical function of the topic *u vahayang* 'the knife' in the sentence

6. *Ahapen mu u vahayang.*  
*get you fm knife*  
 'Get the knife.'

is *object of*.

Different stems may inflect differently in terms of focus. The number and types of focus affixes to which a stem may be inflected signals the number and types of potential nuclear clause tagmemes which may constitute constructions into which the stem may enter. For instance, if a stem is inflectable for subject, object, referent, and associative focuses, the stem may enter into syntactic constructions where subject, object, referent, and associative tagmemes are among the constituents and if a stem is not inflectable for a particular focus type, e.g. object focus, it does not enter into a syntactic construction where an object tagmeme occur unless a non-focus affix that allows such focus inflection occurs with the stem (e.g. the derivational reduplicative affix in *-adadaw*) or unless a non-focus affix signals the occurrence of the object (e.g. the causative affix {*pa-1*} in *mapawyug* 'cause *x* to flow').

TABLE 1  
SUMMARY OF AFFIXES  
IVATAN AFFIXES



<sup>8</sup>The subject focus may also convey the verbal notion 'becoming' which -(ma<sub>-1</sub>)- does not, e.g. umhutab 'become bubbly', mahutab 'bubbly', umidak 'become white', maydak 'white', umasin 'become salt', masin 'salty'.

This does not mean, however, that all *n* tagmemes can always occur in one single syntactic construction when the stem is inflectable for all *n* tagmemes. The type of focus of the construction may impose constraints on the type of tagmemes that may occur in a particular construction. That focus is a function of the predicative stem, consider the comparison between verb stem 1, verb stem 2 and noun stem 2 in terms of focus inflections when they are unaffixed with optional voice affixes (Table 2 where *x* = presence).

TABLE 2

<i>Non-Predicate Focus Affixes</i>	Vs1	Vs2	Ns2
1. Subject	x	x	x
2. a) Object b) Obj/Ref c) Obj/As	x		
3. Associative	x	x	x
4. Referent	x	x	
Sample stem	-rutung 'cook'	-uyug 'flow'	adaw 'love'

If the predicative stem is unaffixed with optional voice affixes, -uyug 'flow' does not enter into a syntactic construction where an object is present and adaw 'love' does not enter into a syntactic construction where an object or a referent occurs.

## 2.2 THE STEM AND VOICE AFFIXES

Voice is a feature of the predicative realised by affixes which signal the relationship of the participants to the action such that the potential occurrence or non-occurrence of a clause level tagmeme is indicated.<sup>9</sup> The voice affix is syntactically relevant in terms of clause expansion. For instance, the causative {pa-<sub>1</sub>} occurring with a particular type of verb stem signals the potential occurrence of the indirect object (Y) tagmeme making the construction ditransitive and indicating the sub- as agent rather than actor in the situation, e.g.

7. Naparutung si ina ji Marya su manuk.  
 P                                  S (agent) Y (actor) O  
 'Mother caused Mary to cook chicken.'

or occurring with another type of verb stem, it signals the potential



occurrence of an object tagmeme, thereby transitivity an otherwise intransitive predicative, e.g.

8. Napawyug si Marya su ranum.

P                    S                    O

'Mary caused the water to flow.'

Other voice affixes are directional in nature. They indicate whether an action is of one direction, which, when occurring with a particular type of predicative stem, carries over to a recipient indicating the potential occurrence of the object slot (a transitive action) as in:

9. mangarek 'kiss someone'

10. manweswes 'turn something or someone'

or whether an action is of a reflexive direction or reciprocal direction indicating the absence of an object in syntax, e.g.

11. mayarek 'kiss each other'

12. mayweswes 'turn oneself'

Voice affixes are either obligatory or optional to certain stem types. Different simple stems vary also according to the voice affixes they may take. For instance, the causative affix {pa-<sub>1</sub>} may be affixed to the simple stem -rutung 'cook' which belong to the category verb stem 1 as in maparutung 'cause or make x cook' but the causative affix cannot be affixed to the simple stem adaw 'love' which belongs to the category noun stem 2, unless some other affix occurs to allow the affixation of the causative affix, i.e. a derivational affix, e.g. mapadadaw 'allow x to show affection' with the derivational affix in -adadaw. The implication of this in syntax is that the nuclear grammatical slot, the indirect object slot, which is signalled by the causative affix may occur with a vsl predicative but not with a ns2 predicative. Consider the comparison between vsl, vs2, ns2 and dns2 in term of voice affixes and focus affixes in Table 3 where + = obligatory, - = absence, x = presence, and ± = optional.

(Table 3 overleaf.)

TABLE 3<sup>10</sup>

Voice and Focus Affixes	Vs1	Vs2	Ns2	Dns2
1. Directional voice affix (signals O slot)	+	-	-	+
a) Sf	x	x	x	x
b) O/Rf O/Af Of	x	-	-	x
c) Af	x	x	x	x
d) Rf	x	x	-	x
2. Causative voice affix (signals Y for vs1 and O for vs2)	±	±	-	±
a) Sf	x	x	-	x
b) Of O/Af	-	x	-	-
c) Af	-	-	-	-
d) Rf	-	x	-	-
e) Yf	x	-	-	x

The presence of a directional voice affix N- in the predicative signals the object slot in syntax. Consider:

13. Nanutung aku su manuk.

P S O  
cooked I fm chicken  
'I cooked chicken.'

14. Nadaw aku nu metdeh.

P S A  
was-in-love I fm child  
'I was in love (because of) a child.'

15. Nangadadaw aku su metdeh.

P S O  
(showed) love I fm child  
'I showed love to the child.'

In 14, nu metdeh 'function marker child' is not the object of nadaw 'was in love' but the cause of an emotional state of an affectant, thus the function marker nu (the associative function marker) instead of su (the function marker indicating the function *object of*).

The occurrence of the causative voice affix {pa-<sub>1</sub>} signals the indirect object when the predicative stem is vs1 and the direct object when the predicative stem is vs2.

16. Naparutung aku ji Marya su manuk.

P            S            Y            O  
*caused-cook I fm Mary fm chicken*  
*'I caused Mary to cook chicken.'*

17. Napawyug aku su ranum.

P            S            O  
*caused-flow I fm water*  
*'I caused the water to flow.'*

Table 3 further shows that the occurrence of an optional voice affix in the predicative not only signals the occurrence of tagmemes which cannot occur when this voice affix is absent but also limits the occurrence of focus affixes, thereby restricting the presence of certain tagmemes (the array of cases) in syntax. The associative does not occur in syntax when the causative voice affix is present in the predicative. The referent may occur, however, but the sentence is strained when the indirect object also occurs.

### 3. PROPOSAL

It is proposed, therefore, that syntactic studies should analyse morphology in consonance with syntax and that morphology and syntax should not be studied independently of each other. Our Ivatan studies support the relevance of this point of view in the case of the close correlation between sentence (clause) types and simple stem types. It lends empirical evidence to Fillmore's theorising quoted above, in so far as the array of cases defining the sentence types of a language have the effect of imposing a classification of verbs in the language. As to whether the aspects of the verb classification in this study are universally valid, we can only speculate.

The predicative classification in this study considered structural features such as stem type and affixation and semantic features such as transitive sensitivity, concreteness, abstractness (emotion and quality), meteorological condition, time, and quantity. While it is postulated that the simple stems exhibit inherent semantic features, this aspect of this paper needs further study.

## NOTES

1. Charles J. Fillmore, "The Case for Case", *Universals in Linguistic Theory*, ed. Emmon Bach and Robert T. Harms (New York: Holt, Rinehart and Winston, Inc., 1968), p. 21.
2. See Cesar A. Hidalgo and Araceli C. Hidalgo, *The Structure of Ivatan: Phonological, Lexical, and Grammatical Components* (1970), mimeographed; "Ivatan Morphology: the Predicatives", *The Philippine Journal of Linguistics* 1, No. 2 (December, 1970); and Araceli C. Hidalgo, "Focus in Philippine Languages", *The Philippine Journal of Linguistics* 1, No. 1 (1970).
3. This stem class is the largest of the stem classes. Further investigation indicates that a subclassification of this class is desirable, but this is not done in this paper.
4. The Postal-Lakoff doctrine considers adjectives as a subset of verbs. See Fillmore, op. cit., p. 27.
5. A dichotomy of inflectional affixes and derivational affixes is not presented in this study as a clear-cut distinction between the two is not possible: a derivational affix may also be an inflectional affix simultaneously. For instance, the noun stem avid 'beauty' may become a predicative by prefixing {ma-<sub>1</sub>} to form mavid 'beautiful'. The prefix {ma-<sub>1</sub>} may then be classified as a derivational affix in that it governs the form class of the word, but at the same time it may be classified as an inflectional affix, i.e. as a focus/tense affix.
6. For a detailed discussion of the Ivatan affixes, see Hidalgo and Hidalgo, *A Tagmemic Grammar of Ivatan* (Manila: Linguistic Society of the Philippines, 1971), pp. 51-136.

7. Howard McKaughan and other linguists use the terms subject and actor, respectively for the terms topic and subject in this paper. To use the term actor could be misleading for the grammatical unit to which this term refers could be actor, agent, etc. all marked formally in the same manner in Ivatan. The term subject is thus deemed preferable leaving the term topic for the tagmeme focused by the predicate, i.e. the surface structure function.

The process of marking as topic a clause tagmeme is labelled focusing, not topicalisation. Focusing in this study is called by Fillmore (1968) as primary topicalization and by McKaughan as subjectivalization. McKaughan calls topicalisation Fillmore's secondary topicalization which in Ivatan is the process of permuting a non-predicate clause tagmeme (E) to clause initial position and connecting E to the rest of the clause by a particle which results in giving the notion of underscoring or emphasising the identification of E. This process is called in our study identification-emphasis and topicalisation is used to refer to the process of permuting a non-predicate tagmeme to clause initial position and connecting this tagmeme to the rest of the clause with a conjunctive particle which results in a topic-comment character for the clause, a stylistic transformation having no overtone of emphasising an identification.

8. See note at bottom of Table I (page 31).

9. Not all clause level nuclear tagmemes are correlated with a voice affix, but clause level nuclear tagmemes, whether they are correlated with a voice affix or not, are associated with a focus affix.

10. See Hidalgo and Hidalgo, "Ivatan Morphology: the Predicatives", for other voice affixes, op. cit.

## REFERENCES

- BACH, Emmon and Robert T. Harms (eds)  
1968 *Universals in Linguistic Theory*. Holt, Rinehart and Winston, Inc., New York.
- CHOMSKY, Noam  
1965 *Aspects of the Theory of Syntax*. The M.I.T. Press, Cambridge, Massachusetts.  
1968 *Language and Mind*. Harcourt, Brace & World, Inc., New York.
- COOK, Walter A.  
1969 *An Introduction to Tagmemic Analysis*. Holt, Rinehart and Winston, Inc., New York.
- ELSON, Benjamin and Velma Pickett  
1965 *An Introduction to Morphology and Syntax*. Summer Institute of Linguistics, Santa Ana, California.
- HIDALGO, Araceli C.  
1970 "Focus in Philippine Languages", *The Philippine Journal of Linguistics* 1:25-32.
- HIDALGO, Cesar A.  
1967 "English Verb Inflection", *Lingua* 18:113-130.  
1969 *Ivatan Grammar: a Tagmemic-transformational Analysis*. Ph.D. dissertation, Georgetown University.
- HIDALGO, Cesar A. and Jurgen Heye  
1967 "An Outline of Southern Ivatan Phonology", *General Linguistics* 7:105-120.
- HIDALGO, Cesar A. and Araceli C. Hidalgo  
1970 *The Structure of Ivatan: Phonological, Lexical, and Grammatical*

*Components*. Social Science Research Council, University of the Philippines and the Linguistic Society of the Philippines, Quezon City. Mimeographed.

1970 "Ivatan Morphology: the Predicatives", *The Philippine Journal of Linguistics* 1:9-51.

1971 *A Tagmemic Grammar of Ivatan*. Linguistic Society of the Philippines, Manila.

LONGACRE, Robert E.

1964 *Grammar Discovery Procedures: a Field Manual*. Mouton & Co., The Hague.

LOPEZ, Cecilio

1965 "Contributions to a Comparative Philippine Syntax", *Lingua* 15: 3-16.

MACDONALD, Ross R. and Soenjono Dardjowidjojo

1967 *A Student Reference Grammar of Modern Indonesian*. Georgetown University Press, Washington, D.C.

REID, Lawrence A.

1966 *An Ivatan Syntax*. Oceanic Linguistic Special Publication No. 2. Honolulu, Hawaii.





## LAO FINAL PARTICLES

ARTHUR G. CRISFIELD

Lao has a wealth of final particles. They are one-syllable words which usually occur utterance finally but some may occur alone and a few combinations occur quite often. They seem to function in three ways: as question words, imperative words, or words of special emphasis. They may indicate something about the speaker's feelings or attitude or something about the relationship between two speakers.

Lao is comprised of many dialects. The discussion here is based on the speech of Vientiane, although there is variation among individuals as well as variation from one part of the city to another.

Phonologically the final particles differ somewhat from words in other grammatical categories. Although some follow expected phonological rules of Lao, most of them have variants of different tone and vowel length depending on emphasis or nuance of meaning. Usually there is a short, high variant contrasting with a long, high, rising-falling variant. In general, the long variant is sweeter, more beseeching, softer, or less abrupt. Since some final particles have tones on a pitch or contour not usually found in other Lao words, it is difficult to write them in Lao and know exactly what tone is intended. Although their meanings are sometimes elusive and they are subject to individual variation, the final particles are essential for competence in spoken Lao. In general, these particles are acceptable on all but the most formal speech occasions. Although the number of final particles is certainly not endless, I doubt if the following list is complete.

The transcription makes use of the following inventory. Vowels are doubled when long.

p	t	c	k	ʔ	l	y	u
ph	th		kh		e	ə	o
b	d				ɛ	a	ɔ
m	n	ñ	ŋ		ia	ya	ua
f	s			h			
w		j					
	l						

┘	23	low rising	sǎj 'where'
└	22	low	caj 'heart, mind'
┘	45	high rising	sáj 'fishtrap', phét 'peppery'
┘	44	mid level	sāj 'to put', phēt 'gem'
┘	21	low falling	sâj 'intestine', saât 'mat'
┘	53	high falling	sàj 'to use', saât 'nation'

The two falling tones are glottalised. Words with low tone are said with the low rising tone by most Vientiane speakers giving only five tones. When necessary here, the approximate phonetic value of the tone of a final particle will be indicated by the number scale alone and not by a diacritic.

1 bɔɔ, bɔ̃ɔ This is usually considered the "yes-no" question word rather than a final particle. It may be considered neutral as to attitude or emphasis.

lǎaw si paj bɔɔ 'Is he going?'

2 bɔɔ ʘ 454 When said with this tone, the question word signals doubt. The speaker can hardly believe it.

lǎaw si paj bɔɔ ʘ 454 'Is he going? (I doubt it.)'

3 bɔ̃? When said alone, this means something like 'Oh, I see, is that so?'. The speaker shows he has heard what has been said.

lǎaw si paj bɔ̃? '(You said) he is going?'

4 bú? This is probably a contraction of 'bɔ̃ huù' 'don't know'.

lǎaw paj sǎj bú? 'Where he went, I don't know.'

5 dáj This serves to contradict a statement or action.

lǎaw paj dáj 'He did go too!'

kin dáj '(That's for you!) Eat it!'

6 dé-dee ʘ 454 As a question particle, this means 'and...?, what

about...?, how about...?'

láaw si paj, càw dé 'He's going; how about you?'

It is also used with the meaning 'let me remind you that..., understand that...'.  
 láaw bō paj dé 'He didn't go, (you know).'

7 dē~dēē, nē~nēē As a question particle, this means 'who all, what all, etc.'

mí phǎj dēē 'Who all is there?'

sāj ñǎj dēē 'What all do you put in?'

It is also used in polite requests.

khōō nàmtaan dēē 'May I have some sugar?'

át patuu hāj dēē 'Please close the door.'

8 dǎ~dǎē ㄤ 454 This is used with expressions of encouragement, pardon, warning and others to emphasise the sincerity of the statement, or to be sure one is heard.

soók dii dǎē ㄤ 454 'Good luck, now!'

paj dǎ '(I'm) going, you hear?'

9 dōók This means 'absolutely, indeed, on the contrary'.

bōō mfi dōók 'There isn't any, absolutely.'

paj juū dōók 'Oh, indeed I am going!'

10 hý~hyy ㄤ 454 Alone this indicates puzzlement or wonder. As a question particle, it means 'I wonder, can you tell me'.

láaw paj sǎj hý '(I wonder, do you know) where he went?'

11 kō?, kó?~kō? This is a question particle used to ask that a previous statement be repeated.

an daj kō? 'What did you say (again)?'

láaw paj sǎj kō? 'Where is it (you said) he went?'

12 kō~kōō ㄤ 454 This is used to remind, turn the attention to something or bring up a previously mentioned subject.

kin khâw juū kō 'We were eating (you see)...'

lyàng pỳm kō 'Oh, about the book...'

13 lā This usually follows a "yes-no" question and demands a definite answer.

láaw paj mēēn bōō lā 'He went. Am I right or not?'

càw si paj bōō lā 'Well, are you going or not?'

14 mē, mē~mēē ㄤ 454 This is a request particle meaning 'go ahead with it, don't hesitate'. It is used when the speaker urges someone to do something he wants to do but is afraid to or dares not to.

paj mē 'Go ahead go (I think you should)!'

15 nā This introduces a new subject or draws someone's attention.

sàn nā láaw si paj bɔɔ? *'In that case then, is he going?'*

níi nā cǎw si paj bɔɔ? *'Say, are you going?'*

16 naa ㄟ 454 This can be used to plead or beseech. It is also used when making excuses or giving explanations.

paj naa ㄟ 454, bɔɔ don naa ㄟ 454 *'Oh please come along; it won't take that long!'*

fǒn tók naa ㄟ 454, cȳŋ máa saǎ *'It was raining so of course I was late.'*

It may indicate that the speaker has not been immediately understood.

phūu nǎn naa ㄟ 454, phūu súŋ súŋ naa ㄟ 454 *'That one, the tall one, (don't you see!).'*

17 nɔ̄?~nɔ̄, nɔ̄?~nɔ̄ This is sometimes used alone to indicate agreement with what has been said, such as when something is said and one says *'Isn't it the truth!'*. As a question particle, it requests agreement.

paj nɔ̄? *'Lets' go all right?'*

mēēn nɔ̄? *'That's right, huh?'*

18 nɔ̄~nɔ̄ ㄟ 454 This means *'I wonder'* and is used when asking oneself especially after having forgotten.

láaw paj sǎj nɔ̄ ㄟ 454 *'I wonder where (he said) he was going?'*

19 sǎ~saa ㄟ 454 This is a request particle meaning *'let it be done or decided'*. The longer variant is often used to plead. The short variant may sound curt or impatient depending on the situation.

paj sǎ *'Leave!'*

*'Go ahead and leave. (You don't have to stay).'*

juū saa ㄟ 454 *'Oh, please stay!'*

20 thɔ̄? This is a request particle meaning *'let's...'*.

paj thɔ̄? *'Let's go!'*

21 tī This is used like No. 3 'bɔɔ?' to show one has heard what has been said.

láaw si paj tī *'He's going, (you say).'*

22 tǎ~tii ㄟ 454 This means *'surely, quite sure'* but always indicates some reservation or slight doubt or uncertainty.

paj tǎ *'Sure, let's go, why not. (In response to the question: "Shall we go?")'*

laaw si paj tǎ *'He's going I'm quite sure.'*

laaw si paj tii ㄟ 454 *'Of course he's going.'*

23 tuā~tua ㄟ 454 This is a question particle meaning *'of course?'*.

láaw si paj tuā *'He's going of course?'*

24 wá~waa ㄎ 454 This is a question particle. It usually expresses surprise or even puzzlement. Used alone in response to a previous statement it means 'oh?, is that so?'.  
 bɔɔ paj wá 'Oh, aren't you going?'  
 lǎaw paj sǎj wá 'Now where did he get off to?'  
 khýy hɔ̀n thè waa ㄎ 454 'How could it be so hot?'

25 wə̀j~wə̀j ㄎ 53 This may indicate exuberance or impatience or annoyance. It is not used in formal speech.

cā̀ŋdaj wə̀j 'How's it buddy?'  
 paj sǎj máa wə̀j 'Where the heck have you been?'  
 bɔ̀ɔ paj wə̀j 'Heck with it, I'm not going.'



## THE DEVELOPMENT OF THE REGISTERS IN STANDARD KHMER

PHILIP N. JENNER

As far as I know, the concept of register as a linguistic feature was first applied to the vowel system of modern Mon by Harry L. Shorto, Professor of Mon-Khmer Studies at the School of Oriental and African Studies, University of London. Shorto recognises a *head register* consisting of a vowel subset "characterised by clear voice quality" and a *chest register* consisting of a parallel vowel subset "characterised by breathy voice quality in association with a general laxness of the speech organs and a somewhat centralised articulation of vowels". The quasi-tonal distinction between the two subsets is "inherent in all Mon words" and "similar to that of Cambodian."<sup>1</sup>

The first to apply the term to modern standard Khmer was Miss Eugénie J.A. Henderson, also of the School of Oriental and African Studies.<sup>2</sup> In her now classic description of Khmer phonology,<sup>3</sup> Henderson speaks of a *First Register*, corresponding to Shorto's head register, and a *Second Register*, corresponding to Shorto's chest register. Apart from the descriptive refinements she introduces, these are coextensive with the "a-series" and "o-series" vowels recognised by François Martini<sup>4</sup> as well as his predecessors and followers.<sup>5</sup> However, Henderson, a product of the British school of acoustic phonetics established by Daniel Jones and an exponent of the Firthian school of linguistics, defines her First Register as marked primarily by "a 'normal' or 'head' voice quality", and secondarily by "relatively high pitch". In contrast, her Second Register is marked primarily by "a deep rather breathy or 'sepulchral' voice, pronounced with lowering of the larynx, and frequently accompanied by a certain dilation of the nostrils", and secondarily by lower pitch.<sup>6</sup> While in Mon "the exponents of register are distributed throughout the articulatory complex but exclude pitch features,"<sup>7</sup> in Khmer "the

register of a syllable is closely bound up with the vowel nucleus of that syllable, the two being mutually interdependent..." but includes pitch.<sup>8</sup> For Henderson, then, the primary factor in register is contrastive (oral versus pharyngeal) resonance while contrastive (normal versus lower) pitch is a secondary factor. She allows that "in relation to the VN of the second register, those of the first are in general more open in quality,"<sup>9</sup> but this single reference to the common lowering of her First Register is patently not part of her definition of register. She insists in fact that "the different vowel 'colour' inherent in the registers ... ensures that no vowel nucleus of the first register can ever have exactly the same quality as a vowel nucleus of the second register, no matter how alike their general description may be apart from the question of register."<sup>10</sup>

The very excellence of Henderson's interpretation of the registers, reflecting her broad knowledge of Southeast Asian linguistics, posed a curious problem for others in the field. On the one hand, the question arose of reconciling her conclusions with those of Martini, who had had nothing whatever to say of resonance and pitch contrasts.<sup>11</sup> On the other hand, when it came to applying Henderson's findings to the development of improved pedagogical methods, it was found that their concern with phonetic phenomena called for modifications. A valuable study of the first question was made by Heinz-Jürgen Pinnow, who advisedly concluded that Henderson's phonological description and Martini's Saussurian phonemic description both have undeniable merits.<sup>12</sup> The second and equally serious matter has been ably resolved by Mrs Judith M. Jacob, Lecturer in Cambodian at the School of Oriental and African Studies. Dispensing with Henderson's contrastive pitch, Jacob accepts contrastive resonance and introduces the feature of contrastive tenseness, which is only implicit in Henderson's description but which Shorto attributes to Mon. Jacob stipulates that syllable nuclei of the First Register are "pronounced with a clear, 'head' voice and a certain degree of tension" while those of the Second are pronounced "with a breathy, 'chest' voice and a comparatively relaxed utterance."<sup>13</sup>

More importantly, however, Jacob specifies that this "distinction of voice quality in the utterance of the vowels and diphthongs of the two registers" is *potential* - that is to say, facultative. "This difference of voice quality," she says, "will ... not be heard in the speech of all speakers. It may be heard occasionally [*sic*] in the speech of some speakers and is then most easily noted in syllables uttered in isolation." She rules out, moreover, any registral contrast in the three potential pairs of long high falling diphthongs which she transcribes *iə/ɨə*, *uə/ùə*, *u:ə/ù:ə*, noting that the Cambodians themselves are "sometimes



confused about these diphthongs in some words, not knowing to which register they belong and therefore not knowing which spelling to use."<sup>14</sup>

Jacob's modification of Henderson's original terms bears out Pinnow's conclusion as noted above and makes it easier not only to adjust Martini's observation of functional facts to Henderson's observation of phonetic facts but also to translate Henderson's findings into other phonemic analyses. Thus the American linguist Franklin E. Huffman, a former student of Mrs Jacob, has developed a phonemic interpretation of standard Khmer which improves upon Martini's earlier system and makes no mention whatever of contrastive resonance, pitch, or tension.<sup>15</sup> Rejecting the term "register", he returns to the older conception of "series", which however he restricts to his exposition of the writing system. Such restriction is justified in the sense that the register of any given nucleus is a lexical fact which cannot be accounted for from the data of pure description but can be accounted for in terms of the quasi-historical environments reflected in the writing system. Elsewhere I have commented on some of the strong and weak points of Huffman's analysis.<sup>16</sup>

In this place I wish only to suggest that on pedagogical as well as analytical grounds it may be preferable (a) to return to a position more midway between Shorto, Henderson and Jacob on the one hand and Martini on the other and (b) to pay closer attention to the development of the modern vowel system.

While the actual term used is probably unimportant, my own view is that the label "register" should be retained in phonemic interpretations of standard Khmer but should be redefined (despite the confusion thereby created with Firthian usage) in terms of function. After all, Henderson's and Jacob's phonetic observations have by no means been invalidated or otherwise shelved by conversion into American phonemic terms and, as Jacob points out, contrast *may* be effected by means of systematic alternations of resonance and tension, if not also of Henderson's pitch. The potentiality of more than minimal functional distinctions is hence a feature of the language with which the instructor and student of Khmer must reckon. It is this potentiality that accounts for the presence in standard Khmer of what Fries and Pike called "coexistent phonemic systems".<sup>17</sup> On the one hand we have a *Hochsprache* or Received Pronunciation manifesting 31 contrasting nuclei and reflecting an ideal of delivery appropriate for situations in which speech is formal and largely premeditated, if not actually read or recited. On the other hand we have a parallel style of utterance, "normal" or neutral without being substandard, which manifests only 24 contrasting nuclei. Specifically, on the high, higher-mid and low levels of openness both systems have

6, 3 and 7 nuclei respectively. Above the mean-mid level, however, the Received Pronunciation has 6 nuclei (/ýγ, ìiə, ýγə, ùuə, èe, òo/) not found in the normal style, while below the mean-mid level it has 1 nucleus (/əə/) which the normal style lacks. One of the functions of a sound instructional method should be to explain the rationale of these coexisting systems.

For this and other reasons I maintain that the concept of register, by whatever name it is known, should not be confined to expositions of the writing system. The latter, after all, reflects a stage of the language before the development of the two vowel subsets in question, and whether it is fortunate for the Khmer (and us) that their writing is conservative enough to signal the interrelationships of these subsets is a moot point having little to do with linguistic analysis. The circumstance that it does show registral relationships should not be ignored, but it seems much more useful to my way of thinking to show how the registers are manifested on the phonemic level.

In the first place, the register of most of the 31 syllable nuclei is recognisable on the basis of (a) the nuclear shape itself, (b) the nature of the syllable initial, whether simple or clustered, or (c) a combination of these indices. For example, the 9 nuclei which I write /ii, εε, əə, uu, u, uə, ɔɔ, oə, oə/ fall exclusively within the High Register (Henderson's Second), while the 8 nuclei which I write /aə, aa, a, aaə, o, aao, aa, a/ fall, again exclusively, within the Low Register. This means that ambivalence is limited to the 3 pairs of "broken" nuclei mentioned above by Jacob plus the 4 pairs of simple phonemes which I write /ée : èe, ýγ : ýγ, éə : èə, óo : òo/. Among the consonants initial /m, n, ŋ, w, j, r/ presuppose High Register nuclei, while initial /b, d, q, s, h/ presuppose Low Register nuclei; ambivalence is thus confined to the four stops plus /n, l/. Hence while Henderson and Jacob mark all nuclei of the High Register with the grave accent, I prefer to add diacritics only in the 14 ambivalent cases just cited - with the deplorable innovation of reserving the grave for the Low (First) Register while using the acute for the High. Even a certain percentage<sup>18</sup> of these diacritics is redundant, thanks to the consonant environments specified.

In the second place, the two registers are interdependent in the sense that the phonemic inventory is incomplete without both while members of each registral pair show morphophonemic alternation in the presence of certain affixes: /ruh ~ ruəh/ 'to exist, be (alive)' > /rbəh/ 'object, thing'; /róom/ 'to surround' > /sraəom/ 'to envelop, encase'; /wíiəj/ 'to beat, whip' > /cwəaj/ 'to wind, wrap, plait'; /praəə/ 'to change' > /bəmreə/ 'change'; /roəp/ 'to count' > /prəp/ 'to

*tell, say*'; /hùuət/ 'to be dry' > /samŋúuət/ 'drought'; /kwaɪ/ 'to be uneasy' > /kaŋwui/ 'uneasiness'.

Closer attention to the historical development of the registers would have another pedagogical result, namely that of reassigning the primacy of the registers. This may appear to be a trivial point, but heretofore it has been the Low Register that is treated as the more original vowel subset while the High Register has been relegated to a position vaguely subordinate to it. The only discernible logic in this interpretation is that the Low Register nuclei are associated with voiceless initials while the High Register nuclei are associated with voiced initials which, according to the usual scheme, follow their voiceless counterparts. In reality it is the nuclei of the High Register, with the exceptions noted below, that stand closer to the nuclei of Middle Khmer and have undergone relatively little change. The primacy of the High Register nuclei is, moreover, borne out by the circumstance that all syllables known to me with initial clusters consisting of stops belonging to different registers have High Register nuclei: phgara /pkɔɔr/ 'thunder', thbēka /tpeək/ 'to be bare, bald', chboŋ /cɔpɔh/ 'toward', khjā'ka /kceək/ 'to spit out'.<sup>19</sup>

The vowel system of modern standard Khmer developed by degrees out of the far simpler system of Middle Khmer based on 9 long vowels (Table 1), which can be demonstrated fairly well from rhymes of the period. The 3 falling diphthongs may have had a short prior member but were themselves probably long in functional terms. The short counterparts of /ee/ and /εε/, shown in parentheses, can be neither proved nor disproved with existing evidence and are included mainly for symmetry's sake.

Front unrounded		Central		Back rounded	
<i>long</i>	<i>short</i>	<i>long</i>	<i>short</i>	<i>long</i>	<i>short</i>
ii	i	yy	y	uu	u
	iə		yə		uə
ee	(e)	əə	ə	oo	(o)
εε	(ε)	aa	a	cc	c

Table 1: The Vowel System of Middle Khmer

To all appearances, the Middle Khmer period was a time of rapid phonological development. One can only conclude that this development, reinforced and perhaps even accelerated by a multiplication of contacts with Mon, Cham, Thai and Vietnamese, was a form of readjustment to new conditions created by the gradual devoicing of the voiced stops of Old Khmer. In her doctoral dissertation my colleague Mme Saveros Lewitz registered surprising success in dating this devoicing process by exploiting data embedded in toponyms as transcribed by Portuguese and Spanish travellers of the 16th century and later. Acknowledging that devoicing set in probably early in our era and lasted many centuries, Mme Lewitz concludes that the changes in question seem to have reached completion between the 16th and 18th centuries.<sup>20</sup>

As contrast was lost in Old Khmer or early Middle Khmer between such minimal pairs as *gava* \*[gɔɔw] and *kava* \*[kɔɔw] or *dūra* \*[duur] and *tūra* \*[tuur], compensatory changes took place which maintained the old contrasts in new ways. What part resonance, tension and pitch differences played we are in no position to say, but these changes culminated in the emergence of two parallel vowel subsets: one reflecting the nuclei of syllables with originally voiced initials and showing little modification of the Middle Khmer vowel system, the other reflecting the nuclei of syllables with originally voiceless initials and developing out of the first by several types of increased openness.

Seeing that one term, *register*, has already been taken over from music, it may be permissible to adopt one more. Metaphorically, the intervals of openness from high to higher-mid, from higher-mid to lower-mid, and higher-mid from lower-mid to low may be called intervals of one *tone*; on this basis the intervals from high to lower-high, lower-high to higher-mid, and so forth may be called *semitones*, while smaller intervals can be known as *microtones*. These fine distinctions of tongue-height were one of two mechanisms by which the old contrasts were replaced, the other being the generation of onglides of several types.

Table 2 shows the emergent vowel system of a purely hypothetical early modern stage which may be helpful in following the changes from the Middle Khmer system to that of the present-day standard.

In the *front unrounded nuclei*, the Low Register counterpart of Middle Khmer /ii/ (which, being already high, remains the same) develops by a lowering of the latter by one tone. A precarious contrast is created by lowering the Middle Khmer /ee/ by a microtone, probably with microtonic raising of the tongue in the new High Register. More dramatically, the Low Register counterpart of Middle Khmer /εε/ (which may undergo raising from the higher-low to the lower-mid level) comes into being by development of a low onglide. The Middle Khmer /i/, whose contrast with

/ii/ was probably already accentuated by semitonic lowering to lower-high [ɪ] and by incipient centralisation, may have yielded a correspondingly lower and more central /e/ in the emergent Low Register.

Front unrounded				Central				Back rounded			
long		short		long		short		long		short	
HR	LR	HR	LR	HR	LR	HR	LR	HR	LR	HR	LR
i	ee	i <sup>v</sup>	e <sup>v</sup>	yy	yy <sup>v</sup>	y	y <sup>v</sup>	uu	oo	u <sup>v</sup>	o <sup>v</sup>
	iə ~ iə				yə ~ yə				uə ~ uə		
ee	ee <sup>v</sup>	(e)	(e <sup>v</sup> )	əə	<sub>a</sub> əə	ə	ə <sup>v</sup>	oo	<sub>a</sub> oo	(o)	(o <sup>v</sup> )
εε	<sub>a</sub> εε	(ε)	(ε <sup>v</sup> )	<sub>e</sub> aa	aa	<sub>e</sub> a	a	oo	aa	o	a
				<sub>o</sub> aa		<sub>o</sub> a					

Table 2: The Vowel System of Early Modern Khmer

In the central nuclei, a Low Register counterpart of Middle Khmer /yy/ is manifested only by microtonic lowering. On the mid level, however, the Low Register counterpart of /əə/ appears with the acquisition of a low onglide, while the new High Register nucleus, possibly on the mean-mid level formerly, may be raised by one microtone to reinforce the contrast. On the low level, the process is reversed: Middle Khmer /aa/, already low, can go no lower; it consequently serves as the new Low Register nucleus, while contrast is introduced by generation of a high onglide before it in the new High Register nucleus. Depending on the dialect, this onglide is front [e] or back [o], both probably mean-mid. This circumstance seems to proceed from the variable articulation of Middle Khmer /aa/ which, though generally central, must have ranged between fairly front (where it tended to blend into /εε/) and fairly back (where it tended to blend into /oo/). The same method of introducing new contrasts to replace the old is also used in the case of short /a/, which serves as the new Low Register nucleus while the same mean-mid onglides are generated before it for the new High Register. As in the case of the front vowels, Middle Khmer /y/ may have been a semitone or microtone lower than /yy/; unable to be raised without weakening its contrast with /yy/, it maintains its original level while its equivalent in the new Low Register is manifested by further lowering, nearly to the mean-mid level where contrast between it and the microtonically raised form of Middle Khmer /ə/ blends into it. The Low Register counterpart of the latter may be on the lower-mid level.

In the back rounded nuclei, the Low Register development of Middle Khmer /uu/ is manifested by dropping a full tone to the higher-mid level; that of Middle Khmer /oo/ acquires a low onglide. In both cases the new High Register nucleus is unchanged. The Middle Khmer /ɔɔ/, which may like /εε/ have been on the higher-lower level originally, is raised to lower-mid for the new High Register but is dropped to the low level for the new Low Register. Among the short nuclei, Middle Khmer /u/ was most likely on the lower-high level along with /i/ and /y/, at which point it remains for the HR, being dropped to mean-mid level for the new Low Register. A Middle Khmer /o/, shown in parentheses on Table 1, may have occurred before /-h/; by the time of our early modern Khmer it had probably been absorbed by raising into /u/ while its Low Register counterpart, /ɔ/, was also probably lost at an early date by falling together with /a/. As can be seen, the development of Middle Khmer /ɔ/ runs parallel to its long counterpart.

The 3 falling diphthongs of Middle Khmer probably underwent no change apart from a potentiality of microtonic lowering for the Low Register.

Table 3, finally, shows the vowel system of modern standard Khmer arranged by registral pairs. Two general developments may be noted first:

1. With early modern /<sub>a</sub>εε, <sub>a</sub>əə, <sub>a</sub>oo/ the peak of sonority shifts back to the onglide, which thereby becomes long and stressed, leaving the original element short and unstressed: /aaε, aaə, aao/.

2. With early modern /iə : iə, yə : yə, uə : uə/ the prior elements acquire length and establish their capacity to pattern with initials of either voiceless or voiced nature: /fiə : iə, ɣyə : yə, úuə : uə/. Eventually contrast is lost between the new /fiə/ and early modern /<sup>e</sup>aa/, which first undergoes the same shift as /<sub>a</sub>εε, <sub>a</sub>əə, <sub>a</sub>oo/ and becomes /eeə/ (still heard in some of the nonstandard dialects) and is then raised to /fiə/. On the other hand, contrast is introduced between the new /úuə/ and the new /uə/ allophone of short /u/; be it noted, incidentally, that this /uə/ may well reflect the Middle Khmer /uə/ of ambiguous length. The rare /ooə/ nucleus, which is not in allophonic relationship with /fiə/ but is a functionally distinct diphthong, was not raised along with /eeə/. The short counterparts of /fiə, ooə/, which were /<sup>e</sup>a, <sup>o</sup>a/ in the early modern stage, shift their peak of sonority and become /eə, oə/ and undergo no further change.

We may now turn to examine the less general changes shown in Table 3.

In the front unrounded nuclei, the Low Register counterpart of /ii/, namely /ðə(j)/, has been considerably centralised. The /j/ element, on the synchronic level, is the self-closure this nucleus develops in the absence of another final: pī /pəəj/ 'wind instrument', trī /trəəj/

Front unrounded				Central				Back rounded			
long		short		long		short		long		short	
HR	LR	HR	LR	HR	LR	HR	LR	HR	LR	HR	LR
ii	əə(j)	i	ɨ	ýy	(ýy)	-	-	uu	ɔo	u~	o
fiə	liə	-	-	ýyə	ýyə	-	-	úuə	ùuə	-	-
ée	èe	-	-	éə	aaə	-	-	óo	ao	-	-
εε	aaε	-	-	-	-	-	-	-	-	-	-
fiə	aa	eə	a	-	-	-	-	ɔɔ	aa	u~	ɑ
ooə		oə								uə	

Table 3: The Registral Pairs of Modern Standard Khmer

'fish', *chī* /chəəj/ 'to eat', *ktī* /kdəəj/ 'substance', *srī* /srəəj/ 'female'. However, this /j/ is hard to account for unless we postulate an independent development from Middle Khmer /ii/ parallel to the development of /aaε, aaə, ao/, namely the generation of a low front onglide which eventually took to itself the peak of sonority and moved toward a more central point of articulation. It may be mentioned in this connection that the nucleus /əə/ with other finals is uncommon in modern Khmer. The pair /ée : èe/ normally shows no registral distinction except before /-h/,<sup>21</sup> while in the Received Pronunciation its members contrast as [ɪ] and [ɛ] respectively. The central position of /aa, a/ in the early modern period seems to be demonstrated both by the fact that they acquired back as well as front onglides and by the fact that certain dialects (e.g. Sisaket and Surin) show a front onglide before /aa/ and a central (in reality, back unrounded) onglide before /a/: *māna* /mfiən/ 'to exist' = Sisaket /miən/, *mā'na* /moən/ 'fowl' = Surin /myən/. At all events /aa, a/ now advance from central to front, though the long nucleus continues to have considerable horizontal variability, to reinforce their contrast with /aa, a/. Among the short nuclei, /i/ is centralised to the point of blending into /ý/,<sup>22</sup> while its Low Register counterpart, likewise centralised, either remains on the mean-mid level or drops slightly.

In the central nuclei, apart from the fronting of /aa, a/ and their High Register counterparts, little change occurs other than a loss of length contrasts: /ý/, as has just been said, is absorbed into /i/; /ý/ merges with /ə/, and early modern /eˆ : əˆ/ appear as environmentally

determined allophones of their long counterparts. The Low Register /ỳ/, incidentally, is a doubtful nucleus but may occur in a few syllables.

In the back rounded nuclei, the long nuclei undergo little modification. Among the short nuclei, the early modern /o : ɔ/ appear as environmentally determined allophones of their long counterparts, while the /ɔ/ of Middle Khmer and the early modern period no longer provides contrast with /a/ only a semitone below and is drastically raised to lower-high /u ~ uə/, where it blends with /u/ (represented by *u*) on the one hand and /ó/ (represented by *o*) on the other. This latter nucleus, incidentally, is not seldom raised to [ɔ:] in the Received Pronunciation while the Low Register /ðo/, with which it is not historically related, frequently drops to the mean-mid level. This reinforced contrast parallels that of /é : è/ in the front nuclei.

It can be seen from the foregoing that each Middle Khmer syllable nucleus has yielded two (paired) nuclei in modern standard Khmer in such a way that the vowel system of the latter is divisible into two matching subsets (Tables 1 and 3). In general, these subsets came into being by (1) retention of the Middle Khmer inventory with relatively little modification and (2) development of a parallel inventory by three types of increase in aperture, namely (a) lowering of the pre-modern nuclei by intervals of one semitone or less, (b) lowering by intervals of three or four semitones, and (c) generation of low onglides embracing intervals of from three to five semitones. This general direction of change was perforce reversed in the case of three low vowels, specifically Middle Khmer /aa, a, ɔ/; these served as points of departure for modern Low Register /aa, a, a/, and out of them appeared new High Register reflexes by generation of high onglides in the case of /fiə ~ ooə, eə ~ oə/ and by radical closure in the case of /u ~ uə/. Just as the lowering of the shortened nuclei (/i, ə, u, uə, o/) seems to have been an original secondary feature reinforcing contrast with their long counterparts, so the further lowering of Low Register /è, ðo/ and the further raising of High Register /é, ó/ may be explained as a widening or accentuation of the contrastive interval between them.

In the modern standard, potential registral contrasts involving intervals of *less* than one tone are seven in number: /é : è/ appear to be a little over a semitone apart, /é : è/ and /ó : ðo/ are about a semitone apart, while /ỳ : ỳ/, /fiə : fiə/, /ỳə : ỳə/ and /ú : ú/ are at most only a microtone apart. Inasmuch as their resonance, tension or pitch contrasts may not be manifested, these slight intervals are the only stable feature distinguishing the members of these seven pairs. Predictably, it is precisely among these pairs that registral



contrast is most often lost. It is the presence or absence of contrast in these cases that differentiates the Received Pronunciation from the neutral or normal style of utterance. Experiments which I carried out in 1966 with various informants suggest that native perception of contrast between the members of these seven pairs is proportionate to the intervals between them: in my data, /éə : èə/ were distinguished with nearly 100% accuracy, /ée : èe/ and /óo : òo/ were distinguished with about 70% accuracy, members of the four other pairs with about 55% accuracy. These percentages are valueless in themselves, being almost certainly raised either by recognition of the forms used in my tests or by pure guesswork, and it is more than likely that Jacob is correct in maintaining that registral distinctions are inoperative in the case of /fiə : ðiə, ýyə : ýyè, úuə : ùuè/.

#### NOTES

1. H.L. Short, *A Dictionary of Modern Spoken Mon* (London: Oxford University Press, 1962), x.
2. Eugénie J.A. Henderson, "The Phonology of Loanwords in Some South-East Asian Languages", in *Transactions of the Philological Society of Great Britain*, 1951: 146 and note 1, where the author still refers to the "head" and "chest" registers.
3. Eugénie J.A. Henderson, "The Main Features of Cambodian Pronunciation", in *BSOAS*, XIV (1952). 1:149-74.
4. F. Martini, "Aperçu phonologique du cambodgien", in *BSLP*, 42 (1942-1945). 1:112-31; see especially 129 and 131.
5. See, for example, Georges Maspero, *Grammaire de la langue khmère (cambodgien)* (Paris: Imprimerie Nationale, 1915), 49 (section 39), 79 (section 83), and *passim*; Gaston Cambefort, *Introduction au cambodgien* (Paris: G.P. Maisonneuve, 1950), 20-30; Tiw-011, *Le cambodgien sans maître* (Phnom-Penh: Variya, 1957), 6.
6. Henderson, "Main Features", 151.
7. Shorto, loc. cit.
8. Henderson, "Main Features", 151.
9. *ibid.*, 159.
10. *ibid.*, 155.
11. Martini, op. cit.
12. Heinz-Jürgen Pinnow, "Sprachgeschichtliche Erwägungen zum Phonemsystem des Khmer", in *ZPAS*, 10 (1957). 4:378-91.

13. Judith M. Jacob, *Introduction to Cambodian* (London: Oxford University Press, 1968), 4.
14. *idem.*
15. Franklin E. Huffman, *Modern Spoken Cambodian* (New Haven and London: Yale University Press, 1970), 3-8; Franklin E. Huffman, *Cambodian System of Writing and Beginning Reader, with Drills and Glossary* (New Haven and London: Yale University Press, 1970), 8-11, 15-20, 24-34, 43-6.
16. See my review of Huffman's two volumes in *JAOS*, 92 (1972).4: 556-8.
17. Charles C. Fries and Kenneth L. Pike, "Coexistent Phonemic Systems", in *Lg.*, 25 (1949). 1:29-50.
18. No definite percentage can be assigned because the frequency of initials and other components of the syllable has not been reliably calculated yet.
19. This phenomenon has been expressed as an orthographical rule: if both members of an initial cluster are stops which pattern individually with nuclei of different registers, the post-initial is dominant, i.e. determines the register. The main point here is that, to the best of my knowledge, the post-initial never corresponds to an originally voiceless stop. The significance of this may be more easily understood if I mention that the reading of syllables with initial clusters is not at all difficult. If both members of the cluster pattern individually with nuclei of the same register, the syllable nucleus belongs to that register: *mrāma* /mr̄iəm/ 'finger', *lmama* /lm̄ɔm/ 'enough', *s'āta* /sq̄aət/ 'to be clean', *spēka* /sbaæk/ 'hide, skin'. Conflict arises only when the members pattern individually with opposing registers, inasmuch as one member must be dominant. In this case the rule is simple: originally voiced stops are dominant over spirants (both of which are voiceless); originally voiceless stops and spirants are dominant over liquids, semivowels and nasals (all but two of which function as voiced).
20. See Saveros Lewitz, "La toponymie khmère", in *BEFEO*, LIII (1967). 2:387-8.
21. The fit of the writing system is by no means perfect; the fact remains that a phonemic transcription destined for instructional use should incorporate as many compromises with the writing system as may be useful. The *visarga* is used to represent final /h/, part of the time

before phonemically shortened nuclei, part of the time before nuclei having no shortened counterparts in the modern vowel system. -iḥ /-ih ~ -èh/ and -uḥ /-uh ~ -oh/ pose no problem, but note -aḥ (for -āḥ) /-eəh ~ -ah/. My solution, effective for pedagogical purposes, is to accept shortened allophones of the other four graphemes with which the *visarga* occurs and to arbitrarily abbreviate their phonemic transcription: -eḥ /-éh/ [-ʏh] ~ /-èh/ [-ɛh], -ēḥ /-eh/ [-ɛh] ~ /-aəh/ [-ɛh], -æḥ /-áh/ [-ɤh] ~ /-aəh/ [-ɤh], -oḥ /-óh/ [-oh ~ -oəh] ~ /-aoh/ [-ah]. The front and back nuclei in this environment are exemplified by neḥ /néh/ [nɪh] 'this' : noḥ /nóh/ [noh] 'that', p̄unḥeḥ /ponnèh/ [pon¹neh] 'as much as this' : p̄onḥeḥ /ponnaoh/ [pon¹nah] 'as much as that', raleḥ /riéh/ [r¹lɪh] ~ ralēḥ /rieh/ [r¹lɛh] 'to drop' : broḥ /próh/ [prəh] 'cause', khmeḥ /kmèh/ [khmɛh] 'vinegar' : koḥ /kaoh/ [kah] 'island'. The main point to be borne in mind is that what was originally a low on-glide is lost entirely when /aaɛ, aaə, aao/ occur before /-h/ and that the abbreviated transcriptions /aɛ, aə, ao/ represent the graphemes used.

22. What is meant here is that functional contrast between [ɪ] and [u], both of which continue to be heard, cannot be demonstrated from actual discourse in the Received Pronunciation.

## OBSERVATIONS ON THE SURIN DIALECT OF KHMER

PHILIP N. JENNER

### 0. INTRODUCTION

It is sometimes forgotten that the southern third of Northeast Thailand, a generally arid and economically depressed territory, falls within the zone of distribution of Khmer speech.<sup>1</sup> One of the more fertile parts of this region, otherwise known as the Khorat Plateau, is the basin of the Mae Nam Mun, which rises in the Dong Phraya Yen chain west of Nakhon Ratchasima (Khorat) and flows east past Ubon to join the Mekong some 40 kilometres above Paksé. In its progress the river crosses Buriram, Surin, Sisaket and Ubon provinces, all bounded on the south by the Dangrek escarpment and northern Cambodia. Most of the inhabitants of the relatively densely populated lands watered by the Mun speak Khmer as their mother-tongue. Official figures are wanting for the total number of Khmer-speakers in the four provinces, and estimates range from an ultra-conservative 200,000 to a probably excessive 500,000.

Of movements and intercommunications between the Mun valley and the Cambodian heartland to the south we have little except inferential knowledge. It is tempting to conjecture that they could never have been important. During much of the Angkorian period the whole Khorat Plateau along with much of Central Thailand was under Khmer suzerainty. This circumstance cannot be used to prove that Khmer speech had a similar extension, for in reality the settlement of the Khorat Plateau is not known in any detail. There are nevertheless good archeological grounds for taking the confluence of the Mae Nam Mun with the Mekong as the earliest identifiable centre of Khmer power - providing Khmer was not also the common language of Fu-nan.<sup>2</sup> However this may be, it is reasonable to suppose that Khmer-speaking rice-growers have been in occupation

of the lower Mun valley for a millennium or more. Control over most of the Khorat Plateau was wrested from Angkor by Rāmādhīpati of Sukhotai by the year 1350. In the wars culminating in the abandonment of Angkor (1431-3) the territories on both sides of the Dangrek were repeatedly ravaged and depopulated. Although it is not known how they were affected, it is unlikely that the humble Khmer-speakers of the Mun came through these events unscathed. At least it can be pointed out that the usual direction of the mass deportations practised at the time was into rather than away from the sparsely settled Northeast. It may even be supposed that from the 15th century on the earlier Khmer population was increased by important accessions of transportees from the Tonlé Sap basin. In 1794 Thailand formally annexed the old Cambodian provinces of Battambang, Angkor, Mongkolborey, Sisophon and Khorat. While this whole territory was under Thai administration there were presumably ample opportunities for contacts and migrations between the Mun valley and central Cambodia, notwithstanding the difficulty of communications across the Dangrek chain. Since the mid-15th century, however, there has been no sustained motive for major contacts or movements, and the historical connection of the Khmer dialects of the Mun basin with modern standard Khmer has yet to be determined.

The town (/myyən/) of Surin,<sup>3</sup> seat of the province of the same name, lies 54 kilometers south of the Mun, the same distance due north of the Cambodian frontier, 42 kilometers by rail east of Buriram, and about 95 kilometers by rail west of Sisaket. During the spring of 1967 it was my good fortune to meet on the University of Hawaii campus a young Thai student, Miss Aree Somthawin, who had been born and raised in Surin, who spoke Khmer as her first language, and who accepted with much good grace my request for her services as an informant. In the following paragraphs I sketch out the results of our collaboration over a period of some four months, believing that the peripheral position of the Surin dialect may justify the presentation of such tentative findings to others in the field of Khmer or Mon-Khmer studies.

The Surin dialect is an unwritten vernacular which has for long, apparently, followed a line of development independent of the Cambodian mainstream.<sup>4</sup> The mutual intelligibility of Surin Khmer and the standard of Cambodia, which I had the opportunity to test, is limited. Its main areas of divergence from standard Khmer are intonation, the vowel system, and lexicon. Register<sup>5</sup> appears to be entirely absent; whereas standard Khmer has a maximum of 31 contrasting syllable nuclei, Surin Khmer has only 23. The circumflex clause terminal seems to be characteristic. The vocabulary contains a good many archaisms (/biət/ 'near'), regionalisms (/kmaat/ 'first person singular pronoun, masculine'), and loans

from Thai (/talaat/ 'market').

## 1. SEGMENTAL PHONEMES

### 1.1. VOWELS

There are 9 simple (long) vowel phonemes, namely /ii, ee, εε; yy, əə, aa; uu, oo, ɔɔ/. These combine with shortness to yield 8 (or 9) shortened counterparts, namely /i, (e), ε; y, ə, a; u, o, ɔ/; functional contrast between /ee/ and /e/ cannot be demonstrated by my data, and is only tentatively assumed for the sake of symmetry. The 3 simple (/ii, yy, uu/) and 3 shortened (/i, y, u/) high vowels combine with /ə/ to form 6 falling diphthongs. The total vowel inventory is shown in Table I. Note the absence of /aa, a/, of /ee, əə, ooə/, and of /aaε, aaə, aao/.

The phonetic quality of these syllable nuclei is not essentially different from that of standard Khmer and will not be detailed here. It may be worth mentioning, however, that as in the standard the phonetic length of the long nuclei is perceptibly decreased by voiceless finals while that of the shortened nuclei is increased by voiced finals. Again as in the standard, the shortened vowels tend to be lower than their long counterparts.

### 1.2. CONSONANTS

There are 17 consonant phonemes, /p, b, m; t, d, n; c, ɲ; k, ŋ; q; w, j; r, l; s, h/, all of which occur initially. /b, d/ are preglottalised (implosive); /q/ represents [ʔ], often theoretical in word-initial position; /w, j/ are the labial and palatal semivowels respectively; /r/ is a voiced lingual flap (occasionally a trill) with alveolar contact in syllable-initial position.

Front Unrounded		Central Unrounded		Back Rounded	
i	ii	y	yy	u	uu
iə	iiə	yə	yyə	uə	uuə
(e)	ee	ə	əə	o	oo
ε	εε	a	aa	ɔ	ɔɔ

Table 1: The Syllable Nuclei

Table 2, which probably falls short of representing all possible combinations, shows the initial consonant clusters occurring in my data.

<i>initials</i> →	p	b	m	t	n	c	k	w	r	l	s
p	x			x			x				x
b		x		x			x				x
m				x		x	x		x		x
t							x		x		x
d	x						x		x		x
n	x		x	x	x	x	x				x
c			x								
ɲ	x										
k	x			x		x					x
ŋ				x						x	x
q	x					x				x	x
w						x			x		x
j							x				
r	x		x	x		x	x				x
l	x		x	x		x	x	x		x	x
s	x						x		x		
h	x		x	x		x	x				

Table 2: Complex Initials



The most common initial elements are /k, p, s, c, t/ while the most common post-initials are /r, l, h, m, n/, in that order. A few ternary clusters with prefix /m-/ 'one' occur: /mkhεε/ 'one month'. As in the standard, transition from the initial to the post-initial may be direct or indirect, i.e. over a phonetic bridge-sound. Direct transition is normal between an initial stop and post-initial /r, h/. Bridge-sounds include a short voiced vowel, normally [ə] but not seldom [ɪ, ɔ]; a short voiceless vowel, here represented by [h]; and aspiration + [ə]. Before post-initial /b, d, q/ transition is indirect, nearly always by way of [ə]. After initial /s/ it is likewise over [ə] in deliberate speech but becomes direct in normal speech. Between stops other than /b, d, q/ aspiration + [ə] marks deliberate speech while normal speech shows only [h]. Before post-initial nasals and /w, l/ transition is indirect, with [h] being most common while [ə], with or without prior aspiration, is also heard in deliberate speech. These phonetic details are exemplified by /ppuək/ [pəpəək] 'cloud', /tbaan/ [təʔba·ʏn] 'to weave', /cmaa/ [čh̥ma·] 'cat', /knoŋ/ [kh̥noŋ] 'room', /pkuər/ [phəkəəɹ] 'thunder', /tnɔɔt/ [thəno·t] 'sugar palm'.

Final consonants, always simple, include all of the initials except /b, d, s/. Final stops are normally unreleased; as a result, contrast between /-k/ and /-q/ is so weak as to be usually lost: /srok/ 'village' is indifferently [srok ~ sroʔ] while /pluəq/ 'to taste' is indifferently [ph̥lœʔ ~ ph̥lœk]. Final /l/, like its initial counterpart, is a voiced post-alveolar lateral; final /r/, unlike its initial counterpart, is a voiced retroflex lateral with frictionless [r] colouring of the prior vowel: /tiər/ [ti·əɹ] 'duck', /kəmmɔɔr/ [kəm'mɔ·ɹ] 'lime', /skɔɔr/ [səkɔ·ɹ] 'sugar'.

Transition from the syllable nuclei to the finals is direct or indirect, i.e. effected by means of glides. The latter include (a) labial, palatal, and neutral glides and (b) voiced homorganic occlusion before nasal finals. Labial glides are heard mainly after back rounded nuclei, their frequency decreasing with the openness of the nucleus: /ruup/ [ru·ʷp] 'image', /poot/ [po·ʷt] 'corn, maize', /thɔɔm/ [thɔ·ʷm] 'to be big'. Palatal glides are heard after front unrounded nuclei before velar finals, as in /peek/ [pe·ʏk] 'too much' and /pleen/ [ph̥le·ʏŋ] 'music', and above all before palatal finals: /sac/ [saʏč] 'meat', /qoc/ [ʔoʏč] 'to kindle', /khəəp/ [kh̥v·ʏp] 'to see', /peep/ [pe·ʏp] 'to be full'. Neutral glides occur most often before voiced finals: /krɔɔm/ [krɔ·ə̯m] 'underside', /puɹ/ [po·ə̯ɹ] 'poison'. They are also occasionally heard before voiceless finals, where they signal the length of the preceding nucleus: /croot/ [č̥ɹo·ə̯t] 'to reap'. In the speech of my informant voiced homorganic occlusion was heard only before final /ŋ/:

/jyŋ/ [jw·<sup>9</sup>ŋ] 'first person plural pronoun', /roŋ/ [ro·<sup>9</sup>ŋ] 'hall, building'. Direct transition from the nucleus to the final is usual in the environments not specified above: /ciik/ [çi·k] 'to dig', /leep/ [le·p] 'to swallow', /baan/ [ʔba·n] 'to get', /chɔɔt/ [çhɔ·t] 'to be stupid'.

Before a pause, particularly in citation forms, the lengthening of final /m, n, ŋ, l/ (though not of /-p/) is common following long as well as short nuclei: /ktim/ [kətimm] 'onion', /myən/ [mwənn] 'fowl', /coŋ/ [çoŋŋ] 'tip, end', /kbaal/ [kəʔba·ll] 'head', /cræən/ [çrɛ·nn] 'much', /riiŋ/ [ri·ŋŋ] 'to be dried up', /qəmmel/ [ʔəm·mell] 'salt'.

## 2. SUPRASEGMENTALS

### 2.1. STRESS

There are 3 degrees of stress: primary or strong, secondary or moderate, and tertiary or weak. Primary and secondary stress are represented by the acute and grave respectively while tertiary stress is unmarked. As in the standard, monosyllables uttered in isolation take primary stress: /báan/ '[I] can [do it]', /srúuəl/ '[It's] easy'; disyllables of native provenance take primary stress on the ultima: /cənnəər/ 'ladder', /ləmbáak/ 'trouble'. The same pattern obtains in disyllabic compounds: /koon kóon/ 'children', /pləəw tɔɔl/ 'road, highway'. Polysyllables take primary stress on the ultima, secondary stress on the first syllable: /tɔɔrəsáp/ 'telephone', /wɪtsəwɔkɔɔn/ 'engineer'. Within the phrase qualifying elements take primary stress, qualified elements secondary or tertiary stress: /tɪɔɔj nɪh/ 'this day = today', /jùp méɪn/ 'last night', /nɪək prəçáan/ 'Mr Prachan', /ptɪəh kən lóok/ 'your house', /baaj tɪk ɪɪiəc nɪh/ 'this evening's supper', /təəw nàa mɔɔw nàa/ 'to go all over'.

### 2.2. INTONATION

My informant's speech showed 3 pitch levels, namely low, mid (normal), and high, hereafter number 1, 2, and 3 respectively. Four clause terminals stand out clearly: a rising pitch contour, /↑/, marking a question in the absence of an interrogative word; a sustained pitch, /+/, marking a momentary suspension of an utterance; a circumflex or rising-falling contour, /↘/, marking questions with the interrogative /nəə/ and the enclitic interrogative /qə/; and a falling or trailing contour, /↓/, marking the completion of a declarative utterance. These may be represented graphically as follows:

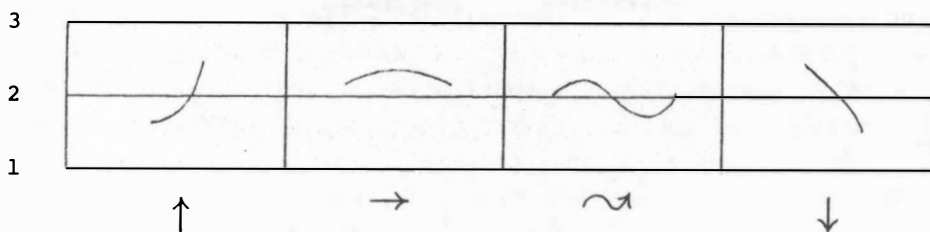


Table 3: The Clause Terminals

### 3. STRUCTURE OF THE WORD

As in the standard, the word may be defined in terms of the syllable. The latter may be expressed as (C)CV(VF), wherein F stands for a non-obligatory consonant final. Monosyllables have the shape CV(VF), exemplified by /tii/ 'place', /toq/ 'table', /bɛɛp/ 'manner'. Subdissyllables have the shape CCV(VF) and may be interpreted as phonemically monosyllabic but for the most part dissyllabic on the phonetic level: /sdam/ [səʔdam] 'right (side)', /rmyəh/ [rɔmwəh] 'itch', /pɲiəw/ [pɲiəw] 'visitor'. Dissyllables of native origin are either (1) compounds of monosyllables or subdissyllables or (2) derivatives by affixation, these last consisting of a monosyllabic or subdissyllabic main syllable together with an unstressed presyllable exhibiting one or the other of two structures: CvN- and Crv-. C in these cases represents any consonant except /m, n, ɲ, ŋ; w, j; h/ while the lower-case v represents a phonemically short, neutral vowel; N represents /m, n, ŋ/ (/ɲ/ is here included in /n/) while the lower-case r is /r/. Illustrating these various presyllables are /pənɾəə/ 'servant', /bəntooh/ 'to blame', /pɾətiəh/ 'to meet', /təmpɛɛk/ 'to be bald', /dənɾɛɛj/ 'elephant', /trəpiəŋ/ 'swamp', /cəŋkəh/ 'chopsticks', /crəmoŋ/ 'nose', /kəmmɔɔr/ 'lime', /krəbɛɛj/ 'carabao', /qənnooŋ/ 'well (puits)', /rəntiəh/ 'lightning', /ləmbaak/ 'trouble', /səmmok/ 'nest', /srəmooc/ 'ant'. Paralleling the rhotacised presyllables the Surin dialect has at least one instance of a /mrə-/, not found in the standard: /mrətiəh/ 'chili pepper', corresponding to mdesa /mtéh/ [m<sup>ə</sup>tɪh].

It is worth noting that the presyllable /qəN-/ was consistently pronounced by my informant either as a nasal with a faint vocalic onglide or as a syllabic nasal: /qəmpəw/ [q<sup>ə</sup>m<sup>1</sup>pɾw ~ ŋ<sup>1</sup>pɾw] 'sugar cane', /qəncul/ [q<sup>ə</sup>ɲ<sup>1</sup>ʒɔl ~ ŋ<sup>1</sup>ʒɔl] 'needle'. Of special interest is the dentalisation of the nasal finals of CvN- presyllables before main-syllable initial /c, r, l/: /bənriən/ 'to teach' (standard /bənɾiən/), /cənriiəp/ 'to salute, greet' (standard /cəmɾiəp/), /kənliəŋ/ 'strength' (standard

/kamlaŋ/), /sənleəŋ/ 'voice' (standard /samléəŋ/), in addition to the /qəncul/ (standard /mcul ~ qəncul/) and /pənɾə/ (standard /bəmɾaə/) cited above. Also worthy of attention is the curious assimilation of base-initial /b, d/ to the final of CvN- presyllables: standard /kam bət/ 'knife' = /kəmmit/, standard /qambət/ 'salt' = /qəmmel/, standard /sambok/ 'nest' = /səmmok/, standard /kambaaor/ [kəm<sup>1</sup>ba·o] 'lime' = /kəmmɔɔr/, standard /cundaaər/ [čɔn<sup>1</sup>ʔda·ə] 'ladder' = /cənnəər/, standard /qandəŋ/ 'well' = /qənnəŋ/. Prefix /m-/ 'one' likewise causes the assimilation of base-initial /d/ to /n/: standard /mdaŋ/ 'once' = /mnɔɔ/ (this is the sole occurrence of the replacement of /ŋ/ by /l/ in my data). Note, finally, that geminate /mm, nn/ resulting from such assimilation are sometimes reduced to simple /m, n/ and forms so affected pass from the status of dissyllables to that of subdissyllables: standard /dambəol/ 'roof' = /tmool/, standard /kandaal/ 'centre' = /knaal/, standard kaŋtura /kandol/ 'rat' = /knɔɔr/.

#### 4. SAMPLE TEXT

The following reproduces part of a breakfast-time dialogue, written and recorded by my informant, between two sisters the elder of whom manages the house, the younger being a teacher.

- A <sup>2</sup>ŋaaŋ nɪh bɔɔŋ <sup>1</sup>n <sup>2</sup>thəə <sup>3</sup>qɛɛj <sup>1</sup>klaŋ †  
*day this elder-sibling future-marker do what pluraliser*  
*What are you going to do today?*
- B <sup>2</sup>thəə qɛɛj-qɛɛj jaŋ <sup>3</sup>náa → <sup>1</sup>mjàŋ <sup>2</sup>bɔɔŋ <sup>1</sup>kɔɔ <sup>2</sup>tɔɔŋ  
*do what-what kind which one-kind elder-sibling then must*  
<sup>2</sup>təw <sup>2</sup>taláat <sup>3</sup>tɔɔn <sup>1</sup>pryk nɪh †  
*go-to market time morning this*  
*Among other things, I have to go to market this morning.*
- A <sup>2</sup>bɔɔŋ <sup>1</sup>n <sup>2</sup>təw taláat <sup>3</sup>náa ~ <sup>2</sup>təw taláat  
*elder-sibling future-marker go-to market which go-to market*  
<sup>3</sup>myət <sup>3</sup>stɪŋ † <sup>2</sup>ryy taláat <sup>3</sup>lɔɔŋ †  
*edge river or market public*  
*Which market are you going to, the one down by the river or the main one?*
- B <sup>2</sup>dèŋ jaŋ <sup>3</sup>náa <sup>2</sup>tee ~  
*know way which question-marker*  
*How am I to know? (= I'm not sure.)*

<sup>2</sup>snàam tæw talàat myət <sup>3</sup>stýŋ <sup>1</sup>həəj †  
*probably go-to market edge river completion-marker*  
*I'll probably go to the one at the river.*

<sup>2</sup>wiə <sup>1</sup>biət <sup>3</sup>cfiəŋ †  
*it close more*  
*It's closer.*

<sup>2</sup>tìŋ rbəh piir bæej jáəŋ <sup>3</sup>ponnòh<sup>2</sup> <sup>1</sup>qəeŋ †  
*buy object two-three kind that-much only*  
*I just have to buy a few things.*

A <sup>1</sup>jyyŋ <sup>2</sup>kwàh <sup>3</sup>qéej <sup>1</sup>klah †  
*we lack what pluraliser*  
*What are we out of?*

<sup>2</sup>krom kyt tháa mɛɛ <sup>1</sup>kyət <sup>2</sup>tìŋ qəej kəmpəh kəmpəŋ<sup>32</sup> mɔw  
*I think say mother she buy what all-sorts-of come*  
<sup>1</sup>həəj tæw mɛɛ <sup>2</sup>mɛŋ †  
*completion-marker go-to yesterday*  
*I thought Mama already bought plenty of things yesterday.*

B <sup>2</sup>tìŋ <sup>1</sup>qənlúək<sup>2</sup> <sup>3</sup>kɔɔ <sup>1</sup>nəŋ <sup>2</sup>sác səmráp iŋiəc nfh <sup>3</sup>tíət †  
*buy greens and-with meat use-for evening this more*  
*I want to get some more vegetables and some meat for tonight.*

<sup>1</sup>bənti<sup>2</sup> qəəw kyət qàac <sup>1</sup>nəŋ <sup>2</sup>nɔəm <sup>3</sup>pɔfəw <sup>1</sup>mɔw <sup>2</sup>ptiəh †  
*maybe father he able linker take guest come house*  
*Papa may be bringing a visitor home with him.*

A <sup>2</sup>kommɔɔj plèc tìŋ <sup>1</sup>mɔrətìh<sup>3</sup> <sup>2</sup>mɔw <sup>3</sup>nəə ~  
*don't forget buy chili come all-right?*  
*Don't forget to get some peppers, will you?*

<sup>2</sup>jup mɛŋ <sup>3</sup>mɛɛ <sup>1</sup>kyət <sup>2</sup>prap thàa → <sup>1</sup>wiə <sup>2</sup>mìn miəŋ <sup>3</sup>lúək <sup>1</sup>qii  
*night last mother she tell say one not have sell at*  
 talàat<sup>2</sup> †  
*market*

*Mama said last night they didn't have any at the market.*

<sup>1</sup>kən <sup>2</sup>jyyŋ kɔɔ hãap <sup>1</sup>nəŋ <sup>3</sup>mɔɔt <sup>2</sup>həəj †  
*thing our clause-connector near linker gone completion-marker*  
*Ours are just about gone.*

B <sup>1</sup>bəcnðh<sup>2</sup> → bɔɔŋ <sup>1</sup>n <sup>2</sup>rùuə <sup>3</sup>tíŋ <sup>2</sup>qɔɔj <sup>1</sup>bàan †  
*if-so elder-sibling will seek buy give (= benefactive) get-to*  
*In that case, I can get some for (her, us).*

A <sup>2</sup>bɔɔŋ <sup>1</sup>nɔŋ <sup>2</sup>təəw talàat myət <sup>3</sup>stýŋ → <sup>2</sup>bɔɔŋ <sup>1</sup>kɔɔ  
*elder-sibling will go-to market edge river elder-sibling then*  
<sup>2</sup>tɔɔŋ <sup>2</sup>phàan <sup>1</sup>ptiəh <sup>3</sup>lqɔɔŋ †  
*must pass house La'ong*

*If you're going to the market by the river, you'll have to pass by La'ong's house.*

<sup>2</sup>jùuə qənsyy <sup>3</sup>nìh <sup>2</sup>təəw qɔɔj kee phɔɔŋ <sup>3</sup>kɔɔ <sup>2</sup>jaan nàa †  
*take book this go-to give her on-way then way what*  
*How about taking this book to her on your way?*

B <sup>1</sup>min <sup>2</sup>qəɛj → <sup>3</sup>bàan <sup>2</sup>həəj †  
*not what get-to completion-marker*  
*It's nothing (= no trouble). I can do it.*

A <sup>2</sup>krðm <sup>1</sup>n <sup>2</sup>tùk qənsyy <sup>1</sup>qii trɔŋ myət <sup>2</sup>twi'ier <sup>3</sup>nìh <sup>2</sup>həəj →  
*I will put book at place edge door this completion-marker*

<sup>2</sup>prùəm <sup>1</sup>tiəŋ <sup>2</sup>cùutməaj <sup>3</sup>phɔɔŋ †  
*along with note as-well*  
*I'll leave it here by the door, along with a note.*

<sup>2</sup>kee soom kcəej <sup>1</sup>mɔɔw <sup>3</sup>mkhéə <sup>2</sup>həəj †  
*she ask borrow come one-month completion-marker*  
*She asked to borrow it a month ago (now),*

<sup>2</sup>kɔɔtəɛ <sup>1</sup>min <sup>2</sup>cəh <sup>1</sup>baan <sup>3</sup>jùuə <sup>2</sup>təəw qɔɔj <sup>1</sup>kəe →  
*but not know-how get-to take go-to give her*  
*but I haven't been able to get it to her.*

<sup>2</sup>qəə <sup>1</sup>bɔɔŋ <sup>3</sup>déŋ <sup>1</sup>qə → thaa <sup>2</sup>qəəw <sup>1</sup>kyət nɔŋ <sup>2</sup>nəəm niək  
*er elder-sibling know ? say father he will take person*

<sup>3</sup>náa <sup>2</sup>mɔɔw †  
*what come*

*Um-m, who's papa bringing home, do you know?*

B <sup>2</sup>lyy thaa → niək mɔ̌w <sup>1</sup>pii myyəŋ kɔ̌q<sup>2</sup> †  
 hear say person come from Bangkok  
 I understand (it's) someone from Bangkok.

<sup>2</sup>bɔ̌ŋ min skỳəl <sup>3</sup>cmúəh <sup>2</sup>kee tee †  
 elder-sibling not know name his emphasiser  
 I don't know what his name is.

<sup>1</sup>qəə <sup>2</sup>bəə <sup>1</sup>min <sup>2</sup>kəət tnaajkwəam<sup>3</sup> † kɔ̌w <sup>2</sup>kəət niək  
 er if not be lawyer then be person

<sup>3</sup>c}iəŋ <sup>1</sup>qəej <sup>2</sup>mjaəŋ mjàəŋ <sup>3</sup>n}h <sup>2</sup>həej †  
 artisan what one-kind-or-another this completion-marker  
 Well, if he's not a lawyer, I expect he's some kind of  
 technician.

A <sup>2</sup>knom səŋsəaj <sup>1</sup>thaa <sup>2</sup>kyət niək khùum thəə <sup>3</sup>tnɔ̌l <sup>1</sup>nəə →  
 I suppose say he person control make road agree?  
 I guess he (must be) in charge of building the road...

<sup>2</sup>pləəw təəw kɔ̌r}iəc<sup>3</sup> <sup>2</sup>nəə → <sup>1</sup>ryy jaəŋ <sup>2</sup>nəə <sup>3</sup>n}h <sup>2</sup>həej †  
 road go-to Khorat agree? or kind what this completion  
 you know, the road to Khorat - or something like that.

B <sup>2</sup>snəam məen <sup>3</sup>dəŋ ~ <sup>2</sup>wliiə kee mɔ̌w <sup>3</sup>kɔ̌w <sup>2</sup>jỳyŋ nɔ̌ŋ dəŋ həej †  
 probably true know time one come then we will know completion  
 You may be right. We'll find out when he gets (they get) here.

A <sup>1</sup>knom <sup>2</sup>təəw <sup>3</sup>nəə ~ <sup>1</sup>tɔ̌ŋ <sup>2</sup>triəm kluuən təəw <sup>1</sup>rooŋ <sup>2</sup>r}ən <sup>1</sup>ləəwnəh †  
 I go all-right? must fix self go-to hall learn now  
 Well, I'm going... I've got to get ready for school now.

B <sup>1</sup>l}iiəc <sup>2</sup>n}h n}iəŋ <sup>1</sup>nɔ̌ŋ <sup>2</sup>hɔ̌q mɔ̌w p}iəh <sup>1</sup>tɔ̌ŋ <sup>3</sup>nəə ~  
 evening this young-lady will come-home come house time what  
 What time will you be getting home tonight?

A <sup>2</sup>dəŋ jaəŋ <sup>3</sup>nəə tee †  
 know way which emphasiser  
 How am I to know? (= I'm not sure.)

<sup>2</sup>knom tɔ̌ŋ sɔ̌p <sup>1</sup>niək <sup>2</sup>r}iəŋ p}ɔ̌ŋ → <sup>1</sup>tɔ̌ŋ <sup>2</sup>pryk n}h †  
 I must test person study all on morning this  
 I have to give a test to the students this morning.

<sup>2</sup>qàac <sup>1</sup>nɔŋ <sup>2</sup>nəəw trùuət <sup>1</sup>kaar <sup>2</sup>sóɔp <sup>1</sup>qɔɔj ruuəc <sup>2</sup>sən †  
 able linker stay correct work test give finish first  
 I may stay on and finish correcting their work before coming  
 home.

B <sup>2</sup>kommɔɔj <sup>1</sup>səəw <sup>2</sup>nəəw lŋiɪəc <sup>3</sup>nàh nǎə ~  
 don't very stay late very agree?  
 Don't be very late, will you?

<sup>2</sup>tèc tɔɔŋ hɔq <sup>1</sup>mɔɔw <sup>2</sup>ptiəh <sup>3</sup>tqéŋ <sup>1</sup>tɔɔn <sup>2</sup>ŋyt ŋyt †  
 else must come-home come house alone in all-dark  
 Otherwise you'll have to come home all by yourself in the dark.

A <sup>3</sup>cáah † <sup>2</sup>knom nɔŋ pjàam hɔq <sup>3</sup>mɔɔw <sup>2</sup>mun pràm mɔɔŋ †  
 yes I will try come-home come before five hour  
 All right, I'll do my best to get back before five.

<sup>3</sup>mìn <sup>2</sup>qɔɔj lŋiɪəc <sup>1</sup>ciiŋ <sup>2</sup>nuh <sup>3</sup>tee ~  
 not give = let late more that emphasiser  
 I won't make it (let it be) any later than that.

B <sup>2</sup>təəw cɔh †  
 go descend  
 Run along!

<sup>3</sup>tèc <sup>2</sup>thaa bɔɔŋ thəə qɔɔj niŋ jùur  
 else say elder-sibling make give = so-that young-lady late

<sup>3</sup>káar †  
 work

Otherwise I'll make you late for work.

<sup>2</sup>ciə lqɔɔ <sup>3</sup>nǎə ~  
 be-well be-fine agree?  
 Bye-bye.



## NOTES

1. Khmer also overlaps Cambodia's western frontier into Prachinburi (Krabin), Chanthaburi, and Trat provinces.
2. On this question see the views summarised in my unpublished doctoral dissertation, *Affixation in Modern Khmer* (University of Hawaii, January 1969).
3. Surin is shown on AMS L509 (1:250,000) sheet ND 48-5 at grid coordinates UB 3946 or 14° 53' N x 103° 29' E. The town claims a population of a little over 147,000.
4. Radio Phnom-Penh is received in Surin but its broadcasts appear to be imperfectly understood. Newspapers, books, and films are not imported from Cambodia. Identification with the Khmer language community appears to be slight.
5. I use the term "register" here in its original Hendersonian sense, which emphasises contrasting oral and pharyngeal resonance and contrasting higher and lower pitch; see Eugénie J.A. Henderson, "The Main Features of Cambodian Pronunciation", in *BSOAS*, XIV (1952).1:149-74.



# PASSIVE AND ERGATIVE IN VIETNAMESE<sup>1</sup>

MARYBETH CLARK

## 0.0. INTRODUCTION

Vietnamese, which appears to be an Austroasiatic language, is spoken by approximately 28 million people in the country of Viet-Nam, along the coastal area of the Indochinese peninsula. There are three major dialect areas: North, North-Central, and South (including South-Central), though dialect differences are not great. The dialect used in this study is South Vietnamese. Some sentences used have been taken from Liem (1969); all sentences used have been checked for grammaticality with Dr Nguyễn-Đặng Liêm, a native speaker of Vietnamese.

Glosses are arbitrary in terms of tense or aspect, since these are usually situationally understood. The term *ông đó*, used extensively in this paper, can be glossed as 'that man', 'that gentleman', 'that grandfather', or 'he'. Conventional orthography is used.

## 0.1. THE PROBLEM

Vietnamese is an accusative language; i.e., in the unmarked transitive sentence, the grammatical subject (the noun phrase immediately preceding the verb) is in the agent case. In this paper, an analysis is made of the structure of transitive Vietnamese sentences which translate into English as passive sentences.

These consist of two distinct types. One is a construction which takes a so-called submissive verb which requires a verb complement which is an embedded sentence. The other is a seemingly ergative construction which takes a transitive verb and a grammatical subject which is in the object case. Verbs in Vietnamese are not morphologically marked for active or passive.

I will define passive sentences as sentences whose grammatical subject is in the object case, whose verb is in a marked form (when the language permits) or there is some marker for passive voice (such as a submissive morpheme), and whose agent when it occurs is a marked noun phrase. It is concluded from the analysis that Vietnamese has no true passive and those sentences which appear to be ergative sentences are, in fact, object-topicalised sentences.

## 1.0. THE GRAMMAR

1.1. We can write the following rules for the underlying structure of the sentences given.

- (1) 1.  $S \rightarrow NP^{\wedge}V (NP) (LP)$   
 2.  $NP \rightarrow \left\{ \left( \begin{array}{c} \{Nu\} \\ \{Pl\} \\ S \end{array} \right) (Cl) N (S) \right\} (Det)$   
 3.  $LP \rightarrow (Loc) N$

S = Sentence  
 NP = Noun Phrase  
 V = Verb  
 LP = Locative Phrase  
 Nu = Number  
 Pl = Plural  
 Cl = Classifier  
 N = Noun  
 Det = Determiner  
 Loc = Locative preposition

These rules are far from being descriptively exhaustive; it is hoped they will work for the constructions presented.

## 1.2. JUSTIFICATION OF THE RULES

### 1.2.1. $S \rightarrow NP^{\wedge}V (NP) (LP)$

My analysis of the Vietnamese sentence differs basically from those of others (cf. Liem, Manley, Taylor, Thompson) in that I consider the subject noun phrase, as well as the verb phrase, to be an obligatory element of Sentence. I do not, at this point, agree with Tim Manley that the predicate is the only obligatory element in the base component, or that such surface representations as

- (2) ngủ /sleep/ '(Someone) is sleeping.'  
 đọc sách /read book/ '(Someone) is reading a book.'

(from Thompson) are grammatical in the base component. I prefer to take the view that there is an obligatory noun phrase in the base which occurs on the surface as the subject, and, in accordance with Chomsky (1965), when it is recoverable (understood) through identical prior reference it can be deleted on the surface. I consider the sentences in (2) to be examples of this process since I believe that they cannot be discourse initial and that their subjects are anaphoric.

Sentence adverbials, negative, interrogative and other sentence elements are not touched upon, nor are verb phrase adverbs.

What, like Manley, I have called stative verbs are elsewhere called descriptive verbs (Taylor), extended state verbs (Thompson, 1965a, p. 218), and predicate adjectives (Liem, 1969, p. 107), and usually translated into English as adjectives with copulative verbs. As is to be expected, stative verbs behave differently from other verbs, mostly in that they can be attributes in the noun phrase and, I think, they can never be followed by NP or S.

Below are examples of the expanded S.

- (3) NP            V  
 ông đó        lạnh  
 man that      cold  
 'That man is cold.'
- (4) NP            V            NP  
 ông đó        mua        sách  
                   buy        book  
 'That man bought a book.'
- (5) NP            V            NP            LP  
 ông đó        mua        sách        ở Sài Gòn  
 'That man buys books in Saigon.'

$$1.2.2. \text{ NP} + \left\{ \begin{array}{l} \{ \text{Nu} \} \\ \text{S} \{ \text{Pl} \} \end{array} \right\} (\text{Cl}) \text{ N } (\text{S}) \left. \vphantom{\left\{ \begin{array}{l} \{ \text{Nu} \} \\ \text{S} \{ \text{Pl} \} \end{array} \right\}} \right\} (\text{Det})$$

This is a very simplified expansion of a very complex noun phrase, but I think it should be adequate to handle the present problem. Most nouns, not all, require a classifier when preceded by a number. With the occurrence of both a number and a classifier, the noun is optional. It seems reasonable to suppose that this is a derivational phenomenon, as Mr Starosta suggests, and that it is handled by a derivational rule:

## (6) [+Cl] &gt; [-N]

Examples of NP expansion are:

- (7) Nu Cl > N Det V N  
 hai người đó mua sách  
*two person that buy book*  
*'Those two people are buying books.'*

- (8) Nu N V N  
 hai ông mua sách?  
*'You (two gentlemen) are buying books?'*

- (9) N S: Conj... V - N Det V  
 ông mà mua sách đó đi  
           who go  
*'That man who bought the books is going.'*

The relative pronoun mà occurs before embedded sentences which are attributes of N. This sentence is uncomfortable without adverbial elements or added verbs to "soften" it, but I feel sure it is grammatical.

## 1.2.3. LP + (Loc) N

Place-nouns following verbs of direction generally occur without a locative particle. I have analysed the word ở in this paper as a locative particle meaning 'in', but since it also occurs as a verb meaning 'to be in; live at', it should probably have two entries in the lexicon.<sup>2</sup>

## 1.3. A SAMPLING OF LEXICON

ông 'gentleman, grandfather'      đại-diện 'representative'      con 'offspring'

[+N  
+Title  
...]

[+N  
+Title  
...]

[+N  
+Title  
...]

sách 'book'

đồ 'thing'

cỏ 'grass'

năm 'year'

[+N  
...]

[+N  
...]

[+N  
+Mass  
...]

[+N  
-[+Cl]\_\_  
...]

họ 'they'

tôi 'I'

Sàigòn a city

[+N  
+Pr  
-[+Pl]\_\_  
...]

[+N  
+Pr  
+([+Nu][+Pl])\_\_  
...]

[+N  
+location  
...]

bị 'suffer something  
unpleasant'

[+V  
+Subm  
...]

được 'benefit from  
something'

[+V  
+Subm  
...]

phải 'be obliged to do  
or undergo some-  
thing'

[+V  
+Subm  
...]

giết 'kill'

[+V  
+ NP  
+[animate]\_\_  
...]

mua 'buy'

[+V  
+ NP  
+[human]\_\_  
...]

chọn 'choose'

[+V  
+ NP  
+[animate]\_\_  
...]

làm 'do, make'

[+V  
+ {NP}  
S  
...]

cho 'give'

[+V  
+ {NP}  
S  
...]

đi 'go'

[+V  
- NP  
...]

lạnh 'be cold'

[+V  
+Stative  
...]

mát 'be cool'

[+V  
+Stative  
...]

vô-dụng 'be useless'

[+V  
+Stative  
...]

đó 'that, there'

[+Det  
...]

người 'person, people'

[+Cl  
+human  
...]

cuốn 'volume, roll'

[+Cl  
...]

hai 'two'

[+Nu  
...]

ở 'in, at'

[+Loc  
...]

những 'plural'

[+Pl  
- [+Pr]  
- [speaker/hearer]  
...]

chúng 'plural'

[+Pl  
+ [+Pr]  
...]

#### 1.4. CONVENTIONS AND REDUNDANCY RULES

Convention: When an NP of a constituent sentence is identical with an NP of its matrix sentence, the constituent NP is deleted.

RR.1 [+Stative] → [- {NP  
S}]

RR.2 [+Subm] → [+\_\_S]

RR.3 [+Pr] → [- [+Nu]\_\_  
- [+Cl]\_\_  
- \_\_ [+Det]  
+human]

RR.4 [+Title] → [+human  
±speaker/hearer]

RR.5 [+speaker/hearer] →  $\begin{bmatrix} -[+Cl] \_ \\ - \_ [+Det] \end{bmatrix}$

RR.6 [+Mass] →  $\begin{bmatrix} -[+Pl] \_ \\ -[+Nu] \_ \\ -[+Cl] \_ \end{bmatrix}$

RR.7 [+human] → [+animate]

## 2.0. SUBMISSIVE VERBS

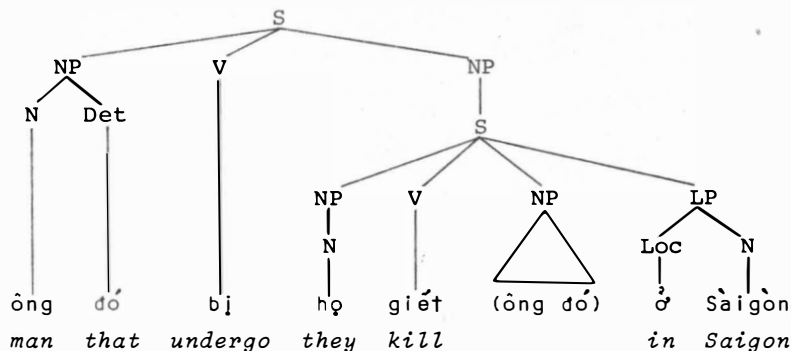
2.1. Liem (1969, p. 103) gives five words which he classifies as verbs and which have a sense of submission, "all having a general lexical meaning of 'to undergo an action' or 'to be obliged to do something' and a particular different nuance". These are as follows:

- (10) bị *under an unhappy experience*  
 mắc *undergo a setback* [Hoa: 'to be caught in']  
 chịu *undergo a burdensome experience*  
 phải *undergo an indifferent experience, or an obligation*  
 được *undergo a happy experience*

Thompson (1965a, pp. 228-9) calls these "momentary action verbs of very generalised meaning which most frequently appear with a descriptive complement" and includes *do 'be accomplished, caused, effected by, be dependent on, the result of'*.

2.2. If these items are marked in the lexicon as [+V] and [+Submissive] and Redundancy Rule 2 requires a sentential complement, application of the grammar to sentence (11) gives the underlying structure shown.

- (11) ông đó bị (họ) giết ở Sài Gòn 'That man was killed (by them) in Saigon.'





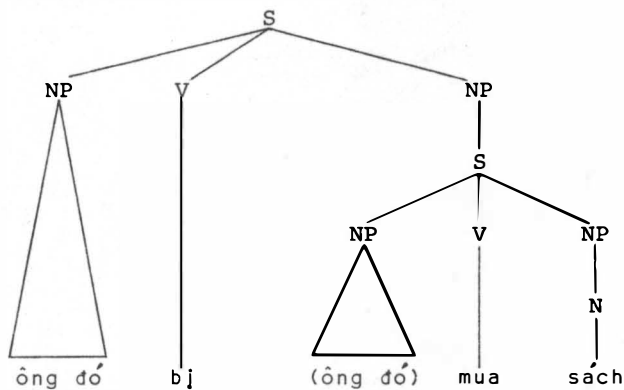
The parentheses around the embedded *ông đó* indicate obligatory deletion of the identical constituent NP, according to the convention. The embedded subject *họ* is optionally deleted (optionally chosen for the surface); i.e. (12) is also grammatical.

(12) *ông đó bị giết ở Sài Gòn* 'He was killed in Saigon.'

2.3. Submissive verbs take three types of verb complements. Sentences like (11) above easily translate as passive sentences because the grammatical subject of the matrix sentence is identical to the object of the embedded verb.

The other two types differ from the first in that the grammatical subject of the embedded verb is identical to the subject of the matrix sentence. In fact, these two types are called Submissive Actional Complements by Liem, whereas he calls the first type Submissive Passive Complements (1969, pp. 18, 103). The first of these "active" types takes a regular verb, either transitive with an object noun phrase or intransitive. The second takes a stative verb. Examples of both types are given.

(13) *ông đó bị mua sách*  
*gentleman that undergo buy book*  
 [-pleasant]  
 'He is [has the misfortune to be] obliged to buy books.'



The following can also occur.

(14) *ông đó được mua sách*  
*undergo*  
 [+pleasant]  
 'He has the good fortune to be able to buy books.'

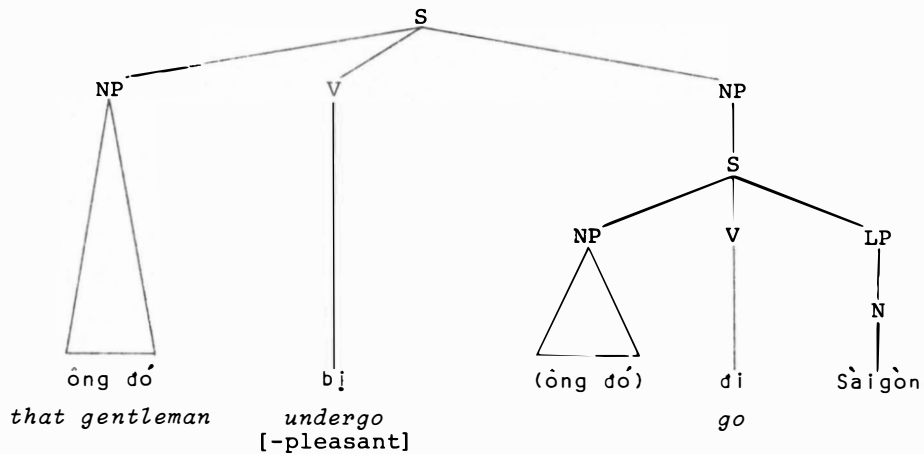
- (15) ông đó phải mua sách  
           undergo  
           [+necessary]

'He is obliged to buy books.'

Sentences with intransitive verb, (16), and with stative verbs, (17) and (18), follow.

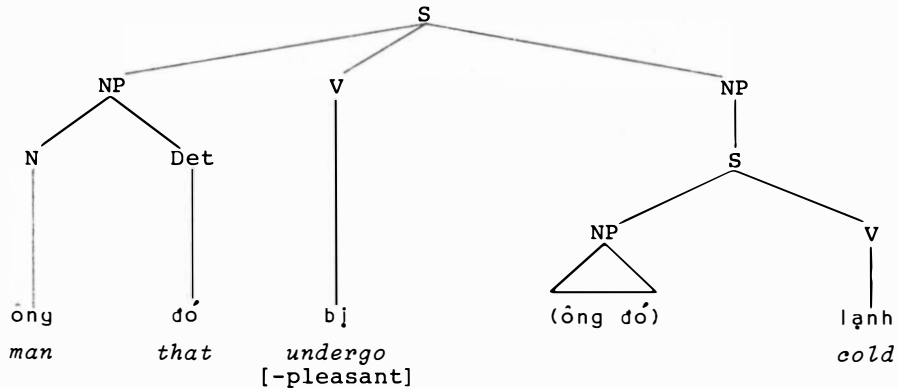
- (16) ông đó bị đi Sài Gòn

'He unfortunately had to go to Saigon.'



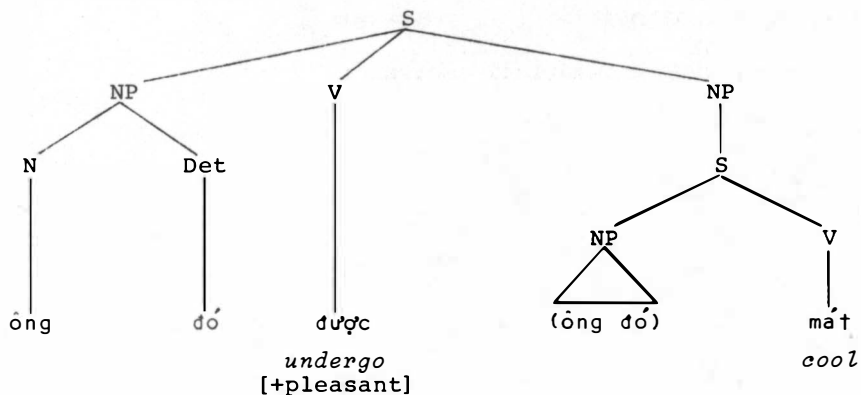
- (17) ông đó bị lạnh

'That man has the misfortune of being cold.'



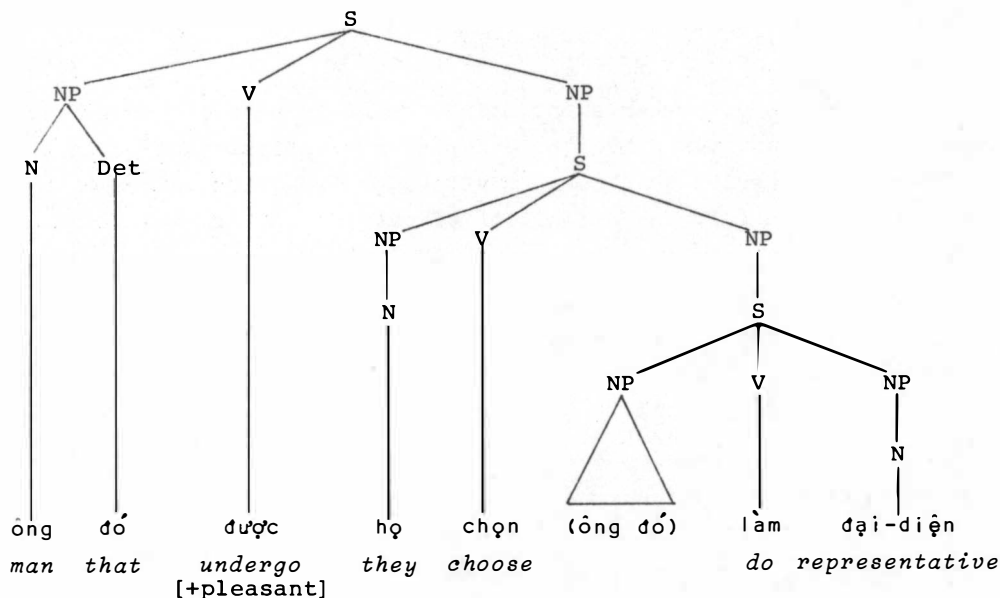
- (18) ông đó được mát

'That man has the good fortune of being cool.'



Following are more examples.

(19) (from Liem 1969, p. 19; with slight lexical changes)



'That man had the good fortune of their choosing him to be the representative.'

(20) (from Liem, 1969, p. 20)

năm ngoái ở đó nó thường bị họ cho đồ vô-dụng  
 year preceding in there he usual undergo they give thing useless  
 (inferior)

'Last year there, he was frequently submitted to the misfortune of having them give him useless things.'

## 3.0. ERGATIVE CONSTRUCTIONS

We have the following grammatical sentence:

- (21)
- 
- ```

graph TD
    S --> NP1[NP]
    S --> V[V]
    S --> NP2[NP]
    S --> LP[LP]
    NP1 --> N1[N]
    N1 --> họ[họ]
    họ --- they[they]
    V --> bán[bán]
    bán --- sell[sell]
    NP2 --> N2[N]
    N2 --> sách[sách]
    sách --- book[book]
    NP2 --> Det[Det]
    Det --> đó[đó]
    đó --- that[that]
    LP --> Loc[Loc]
    Loc --> ở[ở]
    ở --- in[in]
    LP --> N3[N]
    N3 --> Sài Gòn[Sài Gòn]
    Sài Gòn --- Saigon[Saigon]
  
```
- 'They sell that book in Saigon.'*

The only "passive" form this sentence would have is:

- (22) sách đó bán ở Sài Gòn  
*'That book is sold in Saigon.'*

Since (22) does not have an agent, since there is no overt marker for the passive voice, and yet since the sentence grammatical subject is the goal of the verb, perhaps this is an ergative construction rather than a passive construction (cf. the definition of passive in Sec. 0.1.). More sentences of this type are:

- (23) nhà này mua năm ngoái  
*house this buy year preceding*  
*'This house was bought last year.'*  
(from Liem, 1969, p. 11)
- (24) chuyện này thường nghe ở Sài Gòn  
*story this usual hear*  
*'This story is usually heard in Saigon.'*
- (25) sách đó bán nhiều  
*book that sell much*  
*'That book has sold a lot.'*
- (26) nhà này bán  
*house this sell*  
*'This house is for sale.'*

However, Liem lists (23), *nhà này mua năm ngoái*, as an Emphatic construction, and he suggests, in private discussion, that all these sentences may be topicalisation, with the object NP moved to the beginning of the sentence for focus and the subject NP deleted (a common phenomenon in Vietnamese). This view is strengthened by the fact that each of

these sentences can take an agent NP with the only change in meaning being specification of the agent, and the agent NP occurs *only in the grammatical subject (unmarked) position*, i.e. immediately preceding the verb.

(23a) nhà này tôi mua năm ngoái

*I*

*'This house I bought last year.'*

(24a) chuyện này người-ta thường nghe ở Sài Gòn

*people*

*'This story people usually hear in Saigon.'*

(25a) sách đó họ bán nhiều

*they*

*'That book they've sold a lot of.'*

(26a) nhà này tôi bán

*I*

*'This house I'm selling.'*

When ergative sentences have agent NPs, these NPs are marked (cf. Fillmore 1968 and Hohepa 1969). Furthermore, the occurrence of the agent in the subject position when it is present suggests that the logical object, instead of being the grammatical subject, is a preposed topic and there is optional occurrence of an agent.

For these reasons, I consider sentences (23)-(26) and their counterparts (23a)-(26a) to be topicalisations and not ergative sentences.

#### NOTES

1. This paper is a revision of portions of "Is bj really passive and will Vietnamese find the true ergative?", a term paper written for a class in Advanced Analysis under Stanley Starosta at the University of Hawaii, December 1969. Mr Starosta has been helpful in the examination of, but he should not be held responsible for, the ideas presented.

2. I have since revised this notion in a paper on this topic, in which I suggest that, in Vietnamese, Chinese, Thai, and Khmer, certain verbs which have certain lexical case frame features can undergo a lexical derivation rule which allows them to occur as derived prepositions marking case forms on nouns.

## REFERENCES

CHOMSKY, Noam

1965 *Aspects of the Theory of Syntax*. MIT Press, Cambridge.

FILLMORE, Charles

1966 *Toward a Modern Theory of Case*. POLA Report No. 13, The Ohio State Research Foundation, Columbus.

1968 "The Case for Case", in Emmon Bach and Robert T. Harms (eds), *Universals in Linguistic Theory*, Holt, Rinehart & Winston, Inc.

HOHEPA, Patrick W.

1969 "The accusative-to-ergative process in Polynesian languages". *Working Papers in Linguistics*, 6.1-69. University of Hawaii Department of Linguistics, Honolulu.

LIÊM, Nguyễn-Đặng

1969 *Vietnamese Grammar, A Combined Tagmemic and Transformational Approach*. *Pacific Linguistics*, Series C, No. 4, A Contrastive Analysis of English and Vietnamese, vol. 2. The Australian National University, Canberra.

forth- "Clause units in Vietnamese", *Zeitschrift für Phonetik,*  
coming *Sprachwissenschaft und Kommunikationsforschung*.

MANLEY, T.M.

1969 *Verb phrases in Vietnamese* (revised). Term paper, Ling. 650 (S. Starosta), University of Hawaii.

NGUYỄN-Đình-Hòa

1959 *Vietnamese-English Dictionary*. Bình-Minh, Saigon.

TAYLOR, Harvey

- 1968 *Vietnamese syntax, a beginning*. Term paper, Ling. 650  
(S. Starosta), University of Hawaii.

THOMPSON, Laurence C.

- 1965a *A Vietnamese Grammar*. University of Washington Press, Seattle.  
1965b "Nuclear models in Vietnamese immediate-constituent analysis",  
*Language*, 41.610-18.



# SUBMISSIVE VERBS AS ADVERSATIVES IN SOME ASIAN LANGUAGES<sup>1</sup>

MARYBETH CLARK

## 0. INTRODUCTION

The purpose of this paper is to make an initial exploration into the relationship between a certain kind of verb and adversative connotations, in Japanese, Vietnamese, Thai and Lao, Mandarin Chinese, Cambodian, and English.

Branching tree structures represent underlying structures.

## 1. PASSIVE AND ADVERSATIVE

In the preceding paper, "Passive and Ergative in Vietnamese", I have defined passive sentences as sentences whose grammatical subject is in the object case, whose verb is in a marked form (when the language permits) or there is some marker for passive voice (such as a submissive morpheme), and whose agent when it occurs is a marked noun phrase. I will elaborate on this definition to say that the grammatical (or surface) subject is a logical (or deep) Object or Dative case rather than Agent case. (Capitalisation of initial letters will indicate deep case, lower case letters will indicate surface case.) Object and Dative can be thought of in the traditional sense roughly as the direct object and indirect object of the verb, respectively. The grammatical subject can be considered to be in the logical Object/Dative case when the corresponding noun phrase (NP) is in the object/dative surface case when in a non-subject relation to the same verb.

Stanley Starosta, in class discussion, has further suggested that passive sentences are marked sentence types in accusative languages. I will assume that languages are either accusative or ergative, and that

ergative languages can be defined as those which mark the Agent in transitive sentences; i.e. they have Object subjects in unmarked sentence types - the single NP of intransitive sentences and the object NP of transitive sentences are marked the same. (Cf. Fillmore, Hohepa, Lyons, etc.)

This rather broad definition can account for a greater range of sentence types than what are considered passive sentences in English, which require the presence of the copula to carry tense and aspect.

Most languages seem to make use of passivity when they wish to emphasise the adverse effect on the NP of a verb action. Compare the emphasis in English 'he saw me' with 'I was seen (by him)' and 'I got seen (by him)'. Some languages employ particular syntactic constructions to express such adversity. The subjects of such constructions are usually - if not always - animate, and I consider them to be Dative case. All languages cited here are accusative languages and transitivity is an essential feature of adversative sentences, though the corresponding non-adversative sentence may be intransitive. This aspect will be clarified by presentation of examples.

It has been found in Japanese, Vietnamese, Thai, Lao, Cambodian, and Mandarin Chinese that adverse concepts are conveyed by the use of special verbs - sometimes called submissive verbs because of the relation of passivity between the Dative subject and the verb. These verbs require sentential complements (embedded sentences in the object slot).

## 2. THE SITUATION IN VARIOUS LANGUAGES

### 2.1. JAPANESE

Japanese has what Howard calls "pure" passive, which is an innovation in Japanese based on translation from Western languages (Howard, Sugita), which has an Object subject and the "passive" derivational affix *rare* on the verb, and which has corresponding active sentences. Japanese also has what is often called the *inflictive passive*, an indigenous sentence type using *rare* but with a construction quite different from English passive sentences: Howard, Sugita, and others analyse *rare* as a deep-structure verb which requires a sentential complement, and the construction is marked [+adversative]. (Further, *rare* is a special verb in that it doesn't require the infinitive marker *te* which customarily occurs when there is more than one verb in a sentence.) Thus, the "active" sentence (1) with intransitive verb is embedded in the higher sentence of (2).

- (1) *gaadoman ga ne- ta*  
*watchman Nominative sleep Past*  
 'The night-watchman slept.'

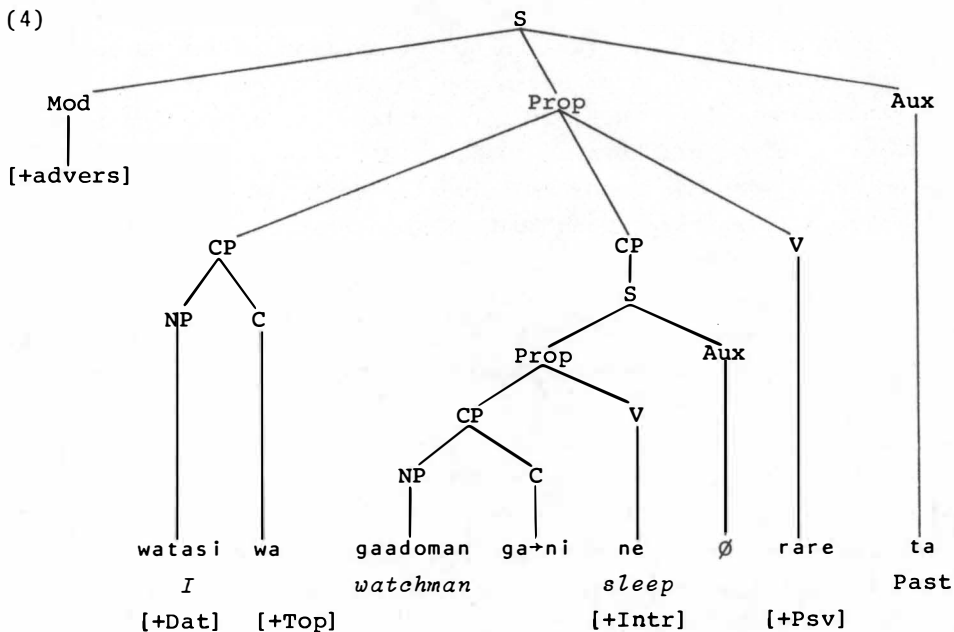
- (2) (watasi wa) gaadoman ni ne-rare-ta  
*I (humble) Topic Agt*  
*'I underwent adversely: the night-watchman slept.'*

Given the rewrite rules in (3) for all cited languages, (2) would have the underlying structure diagrammed in (4).

- (3) S → (Mod) (Aux) Prop  
 Prop → V<sup>^</sup>CP (CP) (CP)  
 CP → { (C) NP }  
           S

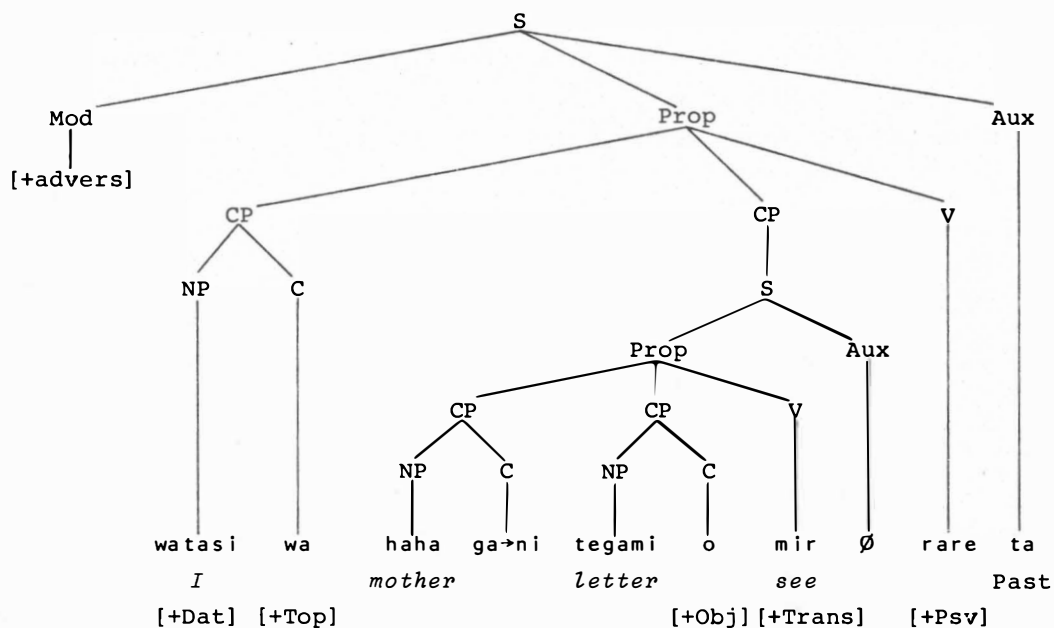
S = Sentence  
 Mod = Modality  
 Aux = Auxiliary  
 Prop = Proposition  
 V = Verb  
 CP = Case Phrase  
 C = Case Marker  
 NP = Noun Phrase

Aux and C are obligatory in Japanese, and there is a rule that states that in a sentence dominated by CP, Aux is realised as  $\emptyset$ .



The embedded sentence in (5) has a transitive verb with two NPs.

- (5) *watasi wa haha ni tegami o mir-are-ta*  
 'I underwent adversely: mother saw the letter.'  
 (from Sugita, pp.3-4)

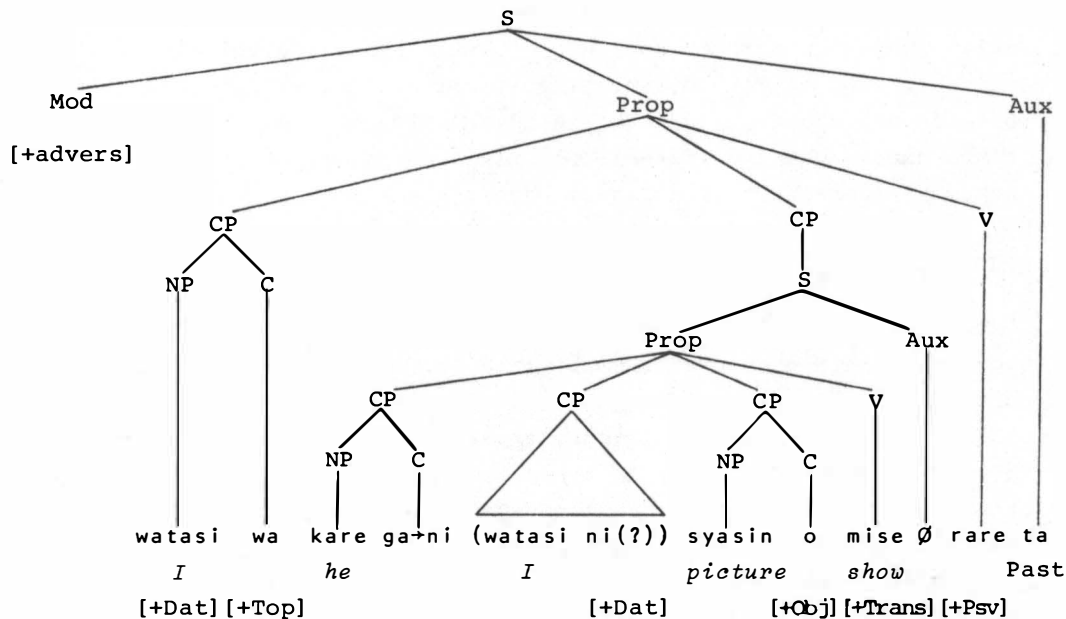


The most striking feature about the Japanese adversative is the fact that the Dative subject NP of *rare* is not equivalent to any NP in the embedded sentence, i.e. it does not stand in direct relationship to the embedded verb. For this reason, it is sometimes called the indirect passive (Sugita, p. 2; cf. Howard, 1968, p. 2). (What Howard and Sugita call the embedded verb has been traditionally analysed as the main verb with a passive suffix *-rare*, such that "the grammatical subject is indirectly, rather than directly, affected by the action of the verb". (Howard, 1968, p. 2).) The subject of *rare* is animate. If it refers to the speaker or to someone situationally understood, it is optionally selected. The subject of the embedded verb cannot be topicalised. It is always marked with the Agent marker *ni* on the surface. Howard adds, "The subject of the constituent sentence, moreover, must be dynamic, that is, an animate object, a natural phenomenon, or a machine, and it may not be identical with the subject of *-rare*." (1968, p. 4).

The construction more easily recognised by English speakers as passive (and translated as such) is one in which the higher NP is identical with the Object or Dative NP in the embedded sentence. This is called the "direct passive" by Sugita (pp. 1, 8f) and the "pure passive" by Howard (1968, pp. 1, 5f), and may or may not be adversative. Howard

(1968, p. 5, 1969, p. 42) states that this passive, "as in English", is transformationally derived from an active counterpart, but I would agree with Sugita that these sentences have the same structure as the adversative, as shown in (6).

- (6) *watasi wa kare ni syasin o mise-rare-ta* (Sugita, p. 9)  
*I he picture show*  
*'I underwent: he showed me a picture.'*  
*('I was shown a picture by him.')*



( ) = delete under identity with higher NP.

The Dative CP *watasi ni* in the embedded sentence is deleted out of identity of the NP *watasi* with the matrix NP (cf. Sec. 3, (31)). The matrix NP can also be identical with the Object in the embedded sentence, as in (7).

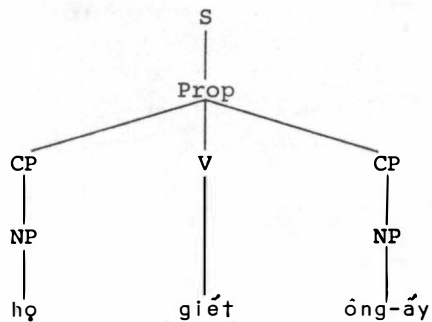
- (7) *watasi wa haha ni home-rare-ta*  
*I mother praise past*  
*'I was praised by my mother.'* (Sugita, p. 8)

The embedded sentence of (7) is (8):

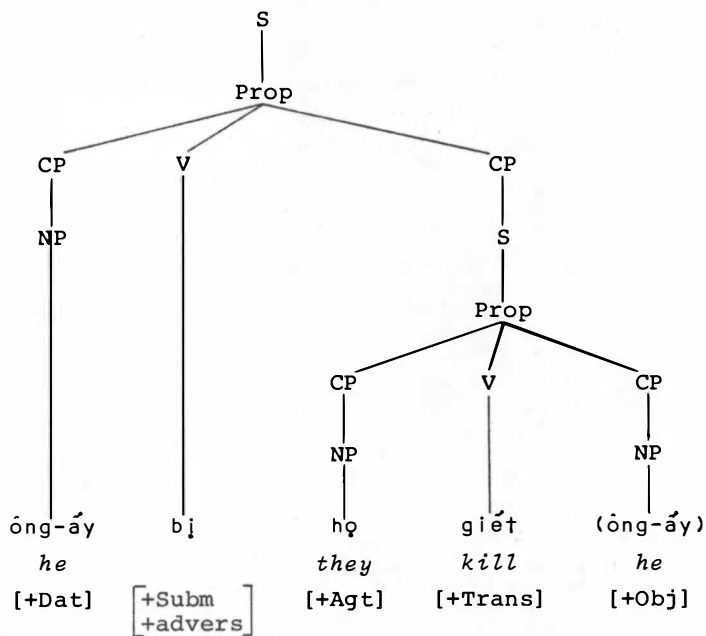
- (8) *haha ga watasi o home-ta*  
*mother Nom I Obj praise-Past*  
*'Mother praised me.'*

There is some controversy whether, if the rare form (7) is used

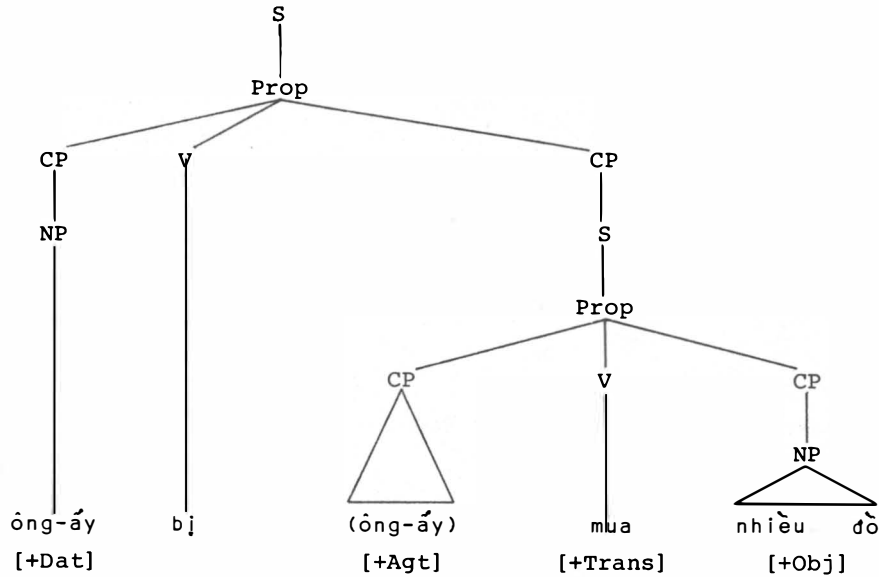




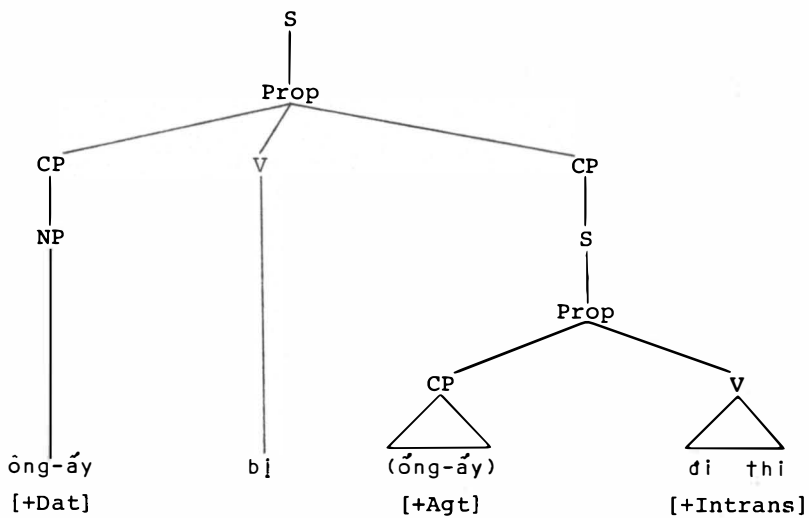
- (11) ông-ấy bị (họ) giết  
*'He underwent adversely: they kill him.'* (*'He was killed.'*)



- (12) ông-ấy bị mua nhiều đồ  
*he buy many thing*  
*'He underwent adversely: he bought many things.'*  
*('He had to buy many things.')*

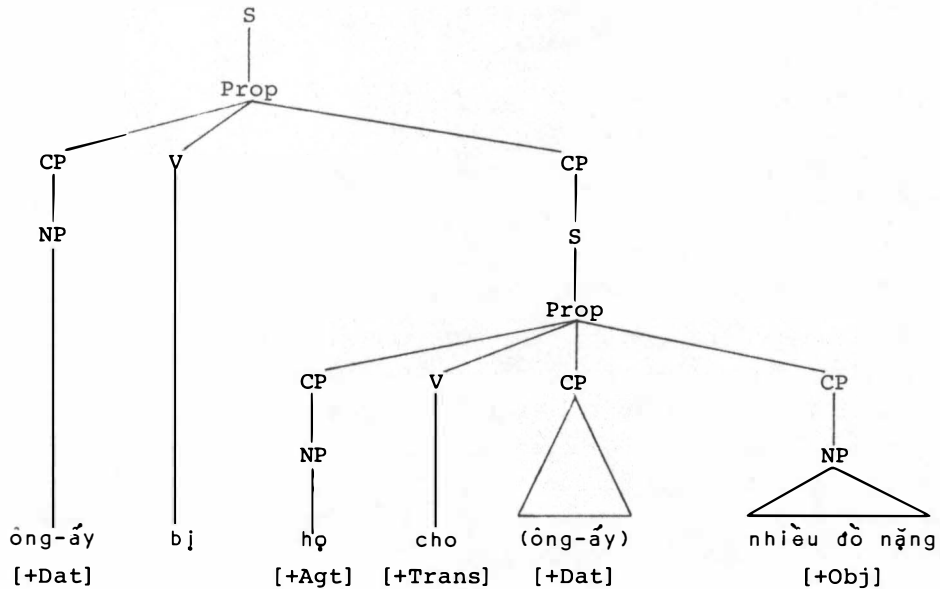


- (13) ông-ấy bị đi thi  
*he go take exam*  
 'He underwent adversely: he took exam.'  
 ('He unfortunately had to take an exam.')

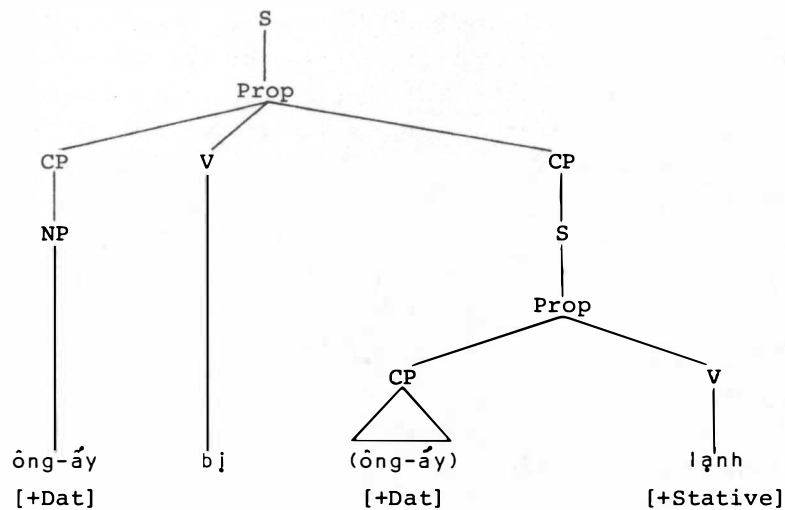


- (14) ông-ấy bị (họ) cho nhiều đồ nặng  
*they give many thing heavy*  
 'He underwent adversely: they gave him many heavy things.'





- (15) ông-ấy bị hạ  
cold  
'He undergoes adversely: he's cold.'



In the case of inalienable possession, it is possible for the matrix NP not to be identical with an embedded NP. If we say that deletion of an embedded NP is dependent on coreferentiality rather than identity, then inalienable possession is a special case of coreferentiality. (16) below was accepted only with discomfort, but (17) is clearly

acceptable. However, Vietnamese does not have anything to correspond to the indirect passive of Japanese.

- (16) ông-ấy bị họ giết vợ  
                   they kill wife  
       '*He underwent adversely: they killed his wife.*'

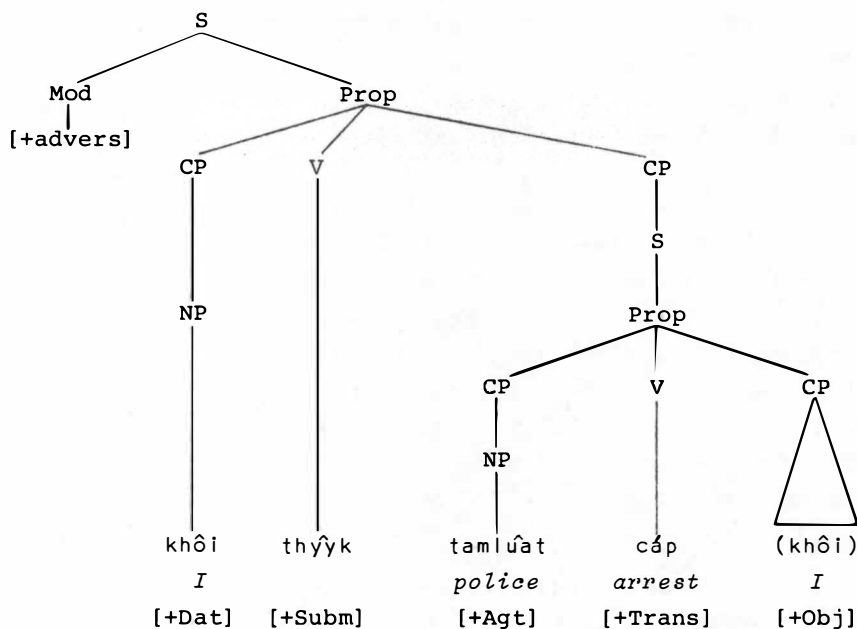
- (17) ông-ấy bị (họ) cắt tay  
                   they cut hand  
       '*He underwent adversely: (they) cut his hand.*'

Apart from sentences with Submissive verbs, Vietnamese seems not to have passive or ergative structures. (Cf. my paper "Passive and ergative in Vietnamese" in this volume.) At any rate, the Agent can occur only in the (unmarked) grammatical subject slot; Dative can be grammatical subject only as subject of Submissive verbs or of Stative verbs - which is not a marked sentence type. The Object can never be subject, though it can be topicalised.

### 2.3. THAI AND LAO

It is probable that Thai and Lao also do not have "passive" or ergative structures. They do have a Submissive verb - *thuòk* in Thai and *thỳk* in Lao - which appears to have a structure like Vietnamese *bị*. (Thai *thuòk* sentences have been analysed as passive and are, of course, translated into English passive sentences. As in Japanese, urban Thai speakers exposed to English use *thuòk* for a translation passive.) It also appears that use of the Submissive verb in Thai and Lao generally connotes adversity, though the verb itself doesn't have the inherent feature [+adversative]. So far as I have been able to tell, the matrix NP must be equivalent to the embedded Object, therefore the embedded verb must be transitive. The following examples, taken from Lao, were given to me by Arthur Crisfield and Vilai Soulatha. All of them have the same underlying structure as that given for (18).

- (18) khôi thỳk (tamluât) cáp  
       I                   police arrest  
       '*I underwent adversely: the police arrested me.*'



(19) mū khôi thýyk khacàw kháa  
*friend I they kill*  
 'My friend underwent: they killed him.'

(20) khôi thýyk sáen  
*invite*  
 'I underwent: (someone) invited me.' ('I got invited.)

(21) khôi thýyk khacàw  $\left. \begin{array}{l} \text{thám l\grave{a}aj} \\ \text{destroy} \\ \text{ñ\grave{o}ng} \\ \text{praise} \end{array} \right\} \text{kh\grave{o}p\text{kh}ua kh\grave{o}i$   
*they family I*

'I underwent: they destroyed/praised my family.'

Presumably, (21) is acceptable because the matrix NP, though not absolutely identical to the embedded Object, is included in the membership of the embedded Object (inalienable possession).

#### 2.4. MANDARIN CHINESE

The complexities of Submissive and possibly Submissive verbs in Chinese would make a very sizeable project of itself and are far beyond the scope of this paper. The reader is referred to works of Annear, Y.R. Chao, A. Hashimoto, M. Hashimoto, H. Wang, and others for more extensive research in this area. What is presented here is a small corner,

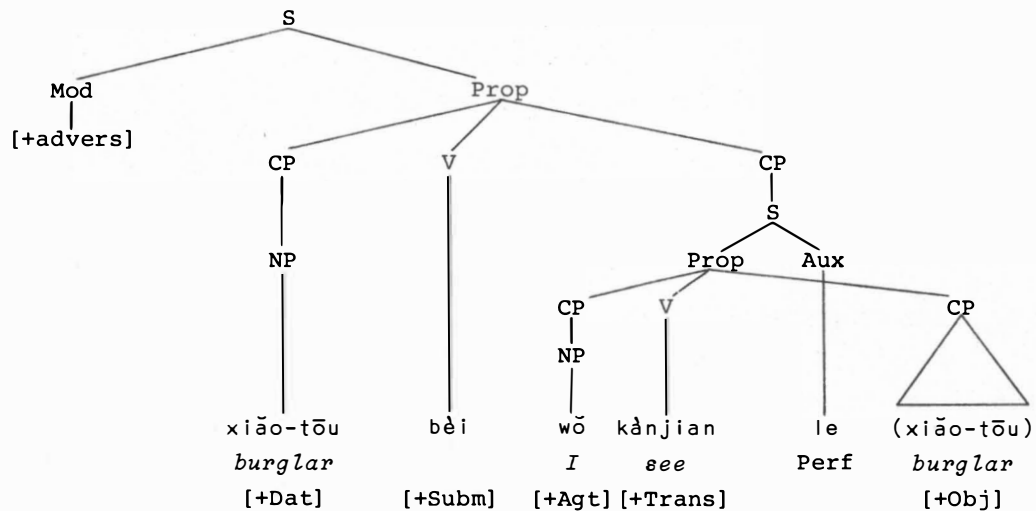
concerning the Mandarin Submissive verb *bèi*, whose structure is similar to the verbs previously discussed. Again, *bèi* sentences tend to be adversative, though *bèi* itself seems to be only [+Submissive]. In Chinese, the matrix NP must be identical with the embedded Object or Dative.

Teresa Cheng (who speaks Mandarin as a second language) would not accept *bèi* with an intransitive verb and non-identical NP, as in (22), a sentence taken from M. Hashimoto and about which Hashimoto says, "...the notion of inflictive construction...will be further strengthened, when an 'intransitive passive' can be found in modern Chinese (though not so numerous as in medieval Chinese)." (p. 64). It is not clear from this remark just how acceptable (22) is, but we are given a clue that earlier Chinese made broader and more specific use of *bèi*, perhaps more like Vietnamese *bị* or Japanese *rare*.

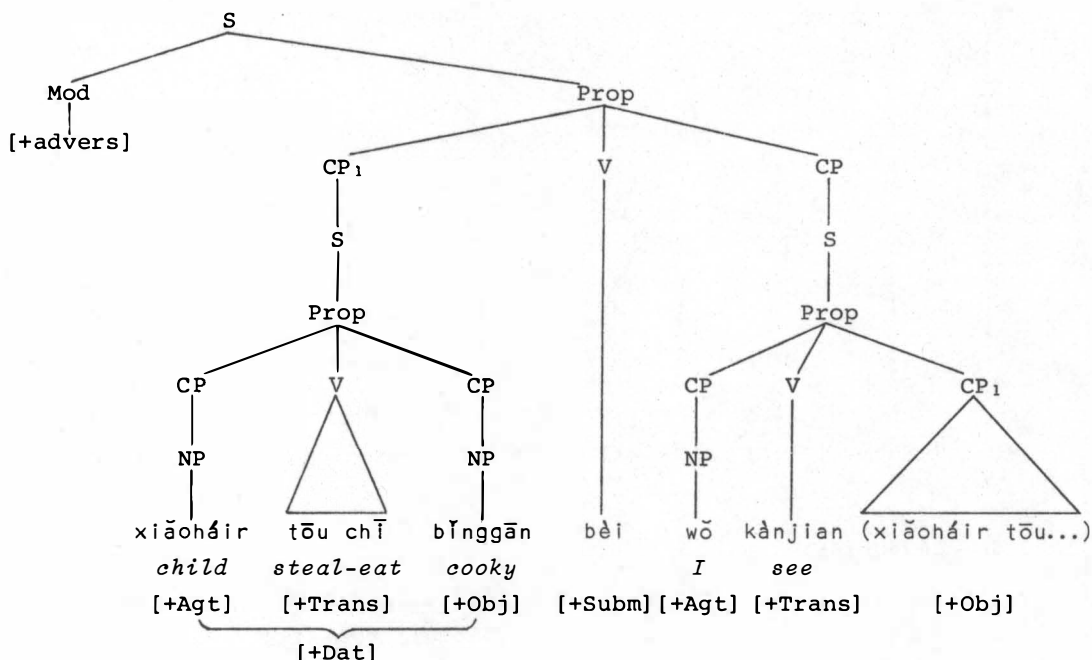
- (22) kǎnshǒu bèi fàn fēn pǎo-le  
 guard criminal run-Perf  
 'The guard underwent adversely: the criminal ran away.'

Mrs Cheng was also hesitant about accepting a single noun for the matrix NP when the embedded verb is not clearly adversative in its lexical meaning, unless the noun itself explained some activity, as in (23). Much more acceptable with a non-adversative verb is (24), a sentence whose matrix NP is an embedded sentence.

- (23) xiǎo-tōu bèi wǒ kànjian le  
 burglar I see Perf  
 'The burglar underwent adversely: I saw him.'



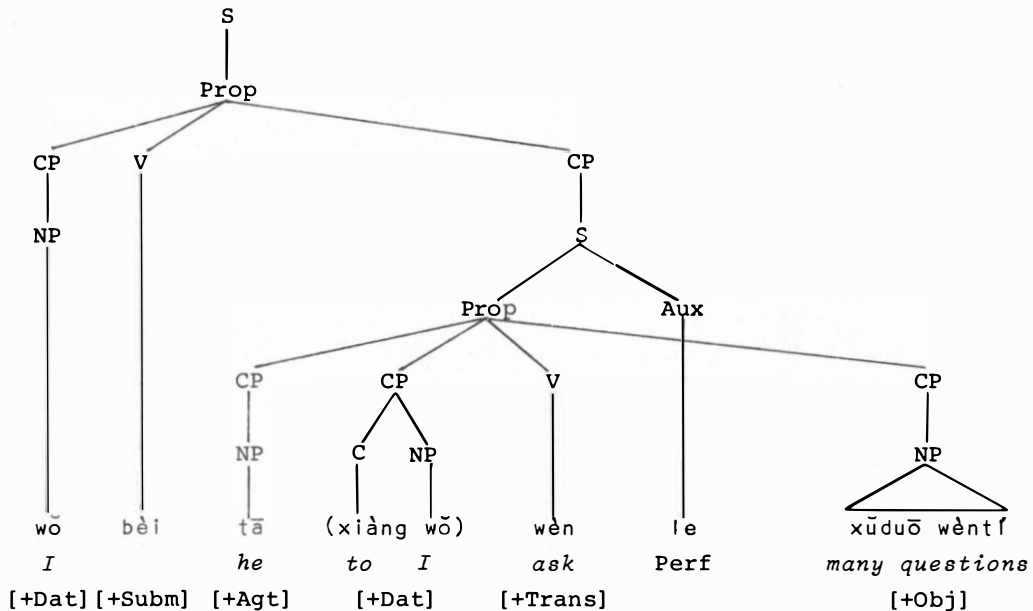
- (24) xiǎo-háir tōu chī bǐnggān bèi wǒ kànjian  
*child steal eat cooky I see*  
 'The child-taking-cookies underwent: I saw it (the incident).'



I suspect, however, that the restriction of adversative meaning in the embedded verb is situational, and that no such restriction exists grammatically.

The matrix NP may be identical to the embedded Dative, as in (25), which is perhaps more ambiguous regarding adversativity.

- (25) wǒ bèi tā wèn-le xǔduō wèn tí (Hashimoto, pp. 63-64)  
*I he ask-Perf many question*  
 'I underwent (adversely): he asked me many questions.'



## 2.5. CAMBODIAN

Huffman (1970, p. 426) gives the following functions and meanings for the Cambodian (Khmer) verb *tṛəw* (Jacob: *troðv*):

adjectival verb: *to be right, correct*

modal verb: *to have to, must*

transitive verb: *to hit, come in contact with; be subjected to, meet with*

He gives many examples of its use as a Submissive verb. As the modal verb 'must', it behaves much like the Vietnamese Submissive verb *phải* 'undergo obligation or necessity', where the matrix NP must be identical to the embedded subject NP. This is shown in the sentence, *kñom tṛəw tɨw psaa tɨay-nih* (*I, undergo, go, market, day-this*) 'I have to go to the market today.' (Huffman, p. 302), where the subject of *tṛəw* is identical to the subject of *tɨw 'go'* in the embedded sentence.

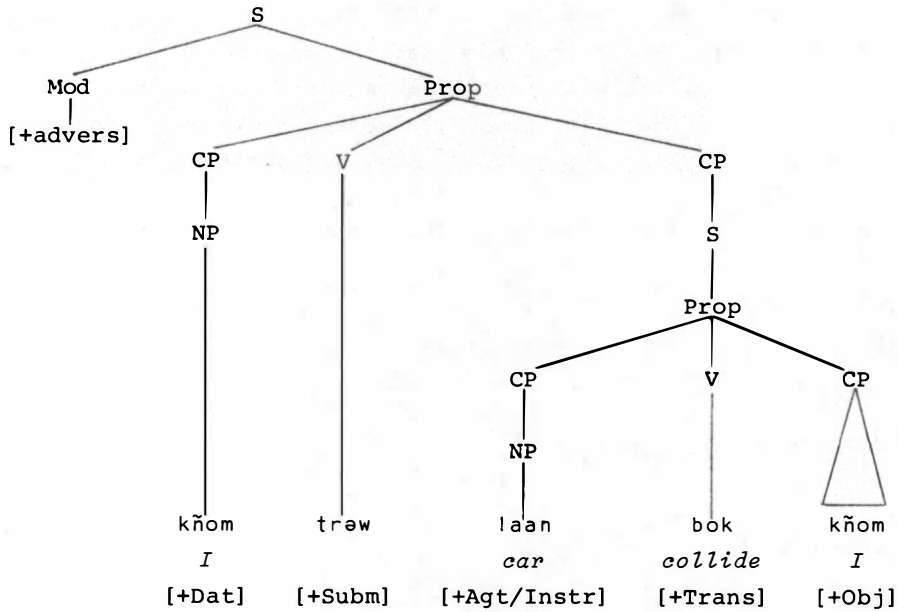
When the subject of *tṛəw* is identical with the embedded Object, as in (26), or the embedded Dative, as in (27), it appears to have an adversative meaning; at least, all such examples found indicate adverse situations. (26a), from Thach Sarun of Southern Illinois University, is an example of the embedded Object being inalienably possessed by the matrix NP.

(26) *kñom tṛəw laan bok* (kñom)

*I undergo car collide with I*

'I was hit by a car.'

(Huffman, p. 302)

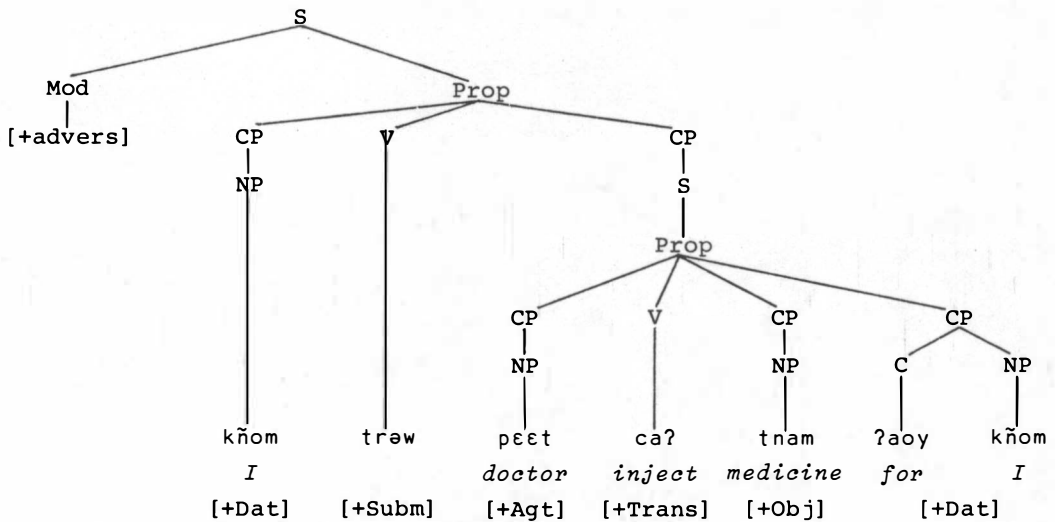


- (26a) kñom traw kee bəm-bæk kbaal (kñom)  
 I undergo they (cause-break) head I  
 break(trans)

'I underwent: they broke my head.' (because deserving of punishment)

- (27) kñom traw pɛt ca? tnam (?aoy kñom)  
 I undergo doctor inject medicine give for I  
 'I underwent: the doctor injected medicine in me.'

(Huffman, p. 302, and Thach Sarun)

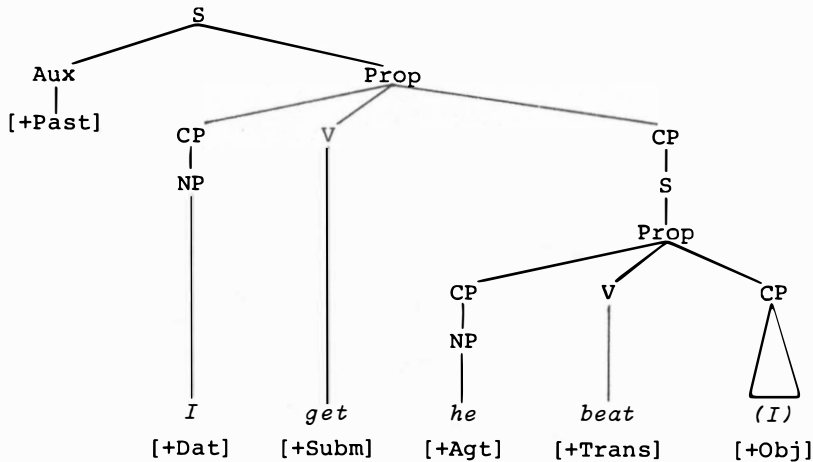


I have not attempted to analyse for this paper such sentences as *mien mənuh mənēə? trəw krūh-tna? bok laan* (*have, person, one person, undergo, accident, collide with, car*) 'There was a man who met with an automobile accident.' (Huffman, p. 302; cf. (26) above); or *kñom trəw baek kbaal* (*I, undergo, break, head*) 'I have a fractured skull.' (Huffman, p. 302), where *baek* is an intransitive verb and *baek kbaal* is perhaps an idiomatic expression (cf. (26a) above).

## 2.6. ENGLISH

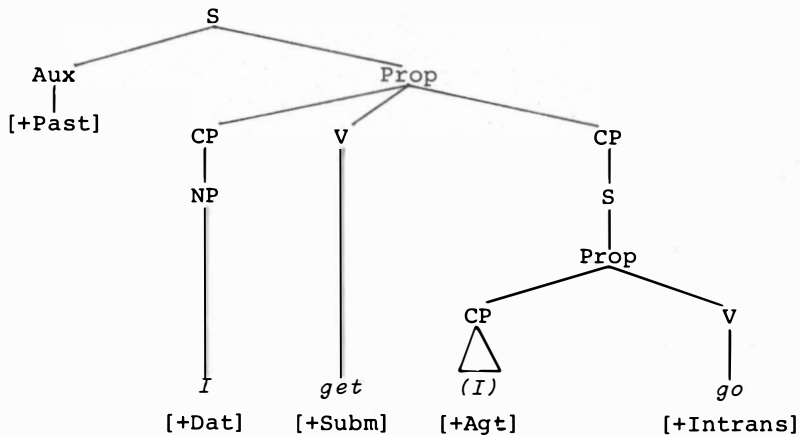
English *get*, in one of its uses, might be considered a Submissive verb. The sentence (28) could have the underlying structure indicated.

(28) *I got beaten (by him).*



Like Vietnamese *bj*, *get* can take an embedded intransitive verb, as in (29).

(29) *I got to go.*

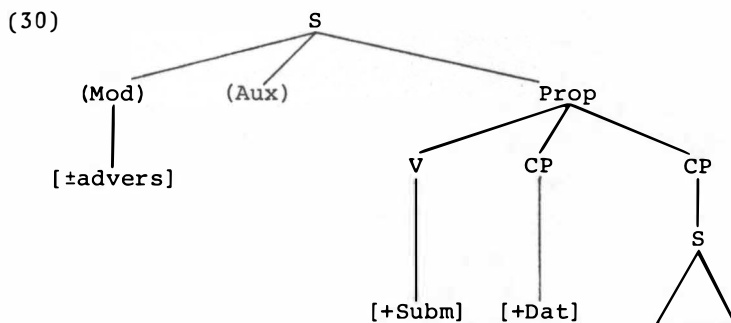




Although in this usage *get* is frequently an adversative, neither its lexical meaning nor the grammatical sentence structure it requires is in itself adversative.

### 3. A SUMMARY OF SUBMISSIVES

We can draw some general conclusions from the evidence given. The underlying structure of sentences with Submissive verbs is shown in (30).



There is a rule such that when a sentence has a Submissive verb and the NP of the matrix sentence is identical with the Object or Dative NP of the constituent sentence, the higher sentence meets the Structural Description for Equi-NP Deletion (Lakoff: ID-NP-DEL, pp. 35, 49ff.), and the entire CP dominating the identical constituent NP is deleted, according to (31), which has been formulated specifically to handle Submissive verb sentences.

$$(31) \text{ END-A} \quad \text{SD: } \left[ \begin{array}{c} Z^{\wedge} \\ \left[ \begin{array}{c} +V \\ +Subm \end{array} \right] \\ S \end{array} \right]^{\wedge} \left[ \begin{array}{c} +CP \\ +NP_i \\ +Dat \end{array} \right]^{\wedge} \left[ \begin{array}{c} X^{\wedge} \\ \left[ \begin{array}{c} +CP \\ +NP_i \\ +Obj/Dat \end{array} \right] \\ S \end{array} \right]^{\wedge} Y \right]_S$$

Structural Change: 1 2 3 4 5 6 → 1,2,3,4,∅,6

Japanese, Lao, and Mandarin use END as stated in (31). Vietnamese and Cambodian do not require the restriction on the constituent NP that it be either Object or Dative. Cambodian, furthermore, does not obligatorily undergo END-A. All five languages share Redundancy Rules 1 and 2. Only Japanese has RR.3.

$$RR.1 \quad [+Subm] \rightarrow \begin{bmatrix} +V \\ + \_ CP^{\wedge} S_C \\ + \_ [+Dat] \\ \pm Dir Psv \end{bmatrix} \quad S_C = \left[ \begin{array}{c} CP \\ S \end{array} \right]_{CP}$$

$$RR.2 \quad [+Dir Psv] \rightarrow \begin{bmatrix} +SD:END-A \\ \pm advers \end{bmatrix}$$

$$RR.3 \quad [-Dir Psv] \rightarrow \begin{bmatrix} -SD:END-A \\ \pm advers \end{bmatrix}$$

In most cases, the Dative subject of Submissive verbs is animate, but there are some exceptions which won't be discussed in this paper.

If the Vietnamese Submissive verbs other than *bj* (cf. Sec. 2.2.) - especially *đủc* 'undergo a happy experience' - truly occur as Submissive verbs, it cannot be said that it is the structure of these Vietnamese sentences which is adversative; whereas, for Mandarin, Thai, Lao, Cambodian, and especially Japanese, it is the particular structure, utilising a Submissive verb, which may mark the sentences [+adversative].

#### 4. A COUPLE OF SUBMISSIVE QUESTIONS

Three questions are raised in my mind regarding the ways languages handle adversatives, particularly the Submissive verbs. The first question is, what is the status of the occurrence of Submissive verbs in relation to the accusative-ergative distinction? Are Submissive verbs the only kind of passive sentences in accusative languages that don't have ergative sentences? (I don't think Japanese has ergative sentences.) Can ergative languages have such structures?

The second question is, to what extent are such special features as Submissive verbs areal? Except for English, all my examples are from languages spoken in East and Southeast Asia. And, except for Thai and Lao, none of the languages analysed here - Japanese (Altaic?), Chinese (Sino-Tibetan), Thai-Lao (Thai-Kadai?: Sino-Tibetan?), Vietnamese (Austroasiatic), Cambodian (Mon-Khmer: Austroasiatic) - are genetically related unless distantly. Is such a phenomenon really special? What other languages of the world exhibit it? If Submissive verbs are not typologically associated and if they are special to Southeast Asia, then it would seem clear that this is an areal phenomenon. There is strong evidence, linguistic and otherwise, of extensive movement and contact in this area. Is this feature the sort of feature that would be

readily shared? If so, what kind of bilingual situation would contribute to this particular sharing?

That brings us to the third question: How closely related to culture is the particular choice of means of passive expression? How closely related to culture is the presence of productive adversatives? Howard (1969, p. 44) has this to say about the Japanese adversative:

"...most likely these sentences are the result of sociolinguistic factors, a kind of formalized modesty which prohibits one from bragging about oneself or one's family. By expressing a positive event as if it were unfavorable, one is expressing humility and minimizing the disparity between oneself and others. Since this is a formalized usage, however, the adversative meaning may be filtered out and the positive content, the intended meaning of the message, is communicated. Perhaps this understood difference between what one is saying and what one means gives the neutral feeling to these passive sentences [with positive (-adversative) embedded verbs]."

There are many speculations regarding the influence of culture on language and the influence of language on thought. Modern psycholinguistics, sociolinguistics, and anthropological linguistics delve into these matters. I mention them here simply as questions to be considered.

#### NOTE

1. This paper is a revision of portions of "How adversative are passives?", a term paper written for a seminar in Case and Copula under Stanley Starosta at the University of Hawaii, May 1971. For data and ideas besides those culled from papers in the bibliography, I am indebted to my teachers (especially Mr Starosta) and fellow students, who have been generous of their time and very helpful, but should not be held responsible for conclusions drawn or questions raised.

\*\*\*

#### REFERENCES

BOLLINGER, Dwight

1968 *Aspects of Language*. Harcourt, Brace & World, Inc., New York.

CARLSON, Barry F.

1970 *Spokan as an ergative language*. Term paper, Ling. 651 (S. Starosta), University of Hawaii.

CHAFE, Wallace L.

1971 "Directionality and paraphrase", *Lang.*, 47.1. 1-26.

CHENG, Teresa

1971 "A Proposal concerning question words", to appear in *Proceedings of Workshop on Syntax and Semantics of Natural Languages*, Stanford University.

CHOMSKY, Noam

1965 *Aspects of the Theory of Syntax*. MIT Press, Cambridge.

CLARK, Marybeth

1969 *Is b1 really passive and will Vietnamese find the true ergative?* Term paper, Ling. 650 (S. Starosta), University of Hawaii.

FILLMORE, Charles J.

1968 "The Case for Case", in E. Bach & R. Harms (eds), *Universals in Linguistic Theory*. Holt, Rinehart & Winston, Inc., New York; pp. 1-88.

HASHIMOTO, Mantaro J.

1969 "Observations on the passive construction." *Unicorn*, 5.59-71. Chinese Linguistics Project and Seminar, Princeton University.

HOHEPA, Patrick W.

1969 "The accusative-to-ergative process in Polynesian languages." *Working Papers in Linguistics*, 6.1-69. University of Hawaii Department of Linguistics, Honolulu.

HOIJER, Harry

1951 "Cultural implications of some Navaho linguistic categories." In D. Hymes, *Language in Culture and Society*. Harper & Row, New York, 1964.

HOWARD, Irwin

1968 *The so-called Japanese passive*. ERIC, Washington.

1969 "A semantic-syntactic analysis of the Japanese passive." *The Journal Newsletter of the Association of Teachers of Japanese*, 6.1.40-46.

HUFFMAN, Franklin E.

1970 *Modern Spoken Cambodian*. Yale University Press, New Haven.

JACOB, Judith M.

1968 *Introduction to Cambodian*. Oxford University Press, London.

KIERMAN, Frank A.

1969 "Night-thoughts on the passive." *Unicorn*, 5.72-78. Chinese Linguistics Project and Seminar, Princeton University.

## LAKOFF, George

- 1970 *Irregularity in Syntax*. Transatlantic Series in Linguistics. Holt, Rinehart and Winston, Inc., New York.

## LI, Ying-che

- 1970 *An Investigation of Case in Chinese Grammar*. University of Michigan Ph.D. thesis.

## LIEM, Nguyen-Dang

- 1969 *Vietnamese Grammar, A Combined Tagmemic and Transformational Approach*. *Pacific Linguistics*, Series C, No. 4, A Contrastive Analysis of English and Vietnamese, vol. 2. The Australian National University, Canberra.

## LYONS, John

- 1968 *Introduction to Theoretical Linguistics*. University Press, Cambridge.

## PARRISH, James E., Stephen Anderson, Adrian Akmajian, Kenneth Hale

- n.d. *Remarks on pronominalization and the passive in Navajo*. MIT ditto.

## SAPIR, Edward

- 1931 "Conceptual categories in primitive languages." *Science*, 74.578.

## STAROSTA, Stanley

- 1967 *Tsou focus*. Mimeo, University of Hawaii.  
1971 Seminar lectures in linguistic theory, University of Hawaii.

## SUGITA, Emiko

- 1970 *Reconsideration of Japanese passive*. Term paper, Ling. 651 (S. Starosta), University of Hawaii.

## THOMPSON, Sandra Annear

- 1968? *The "passive" in Thai* (revised version). UCLA ditto.

## TIMBERLAKE, A.H.

- 1970 *Impersonal and passive in Maori*. MIT 23-783.

# MODERN LINGUISTIC THEORIES AND CONTRASTIVE ANALYSIS OF ENGLISH AND VIETNAMESE

NGUYEN DANG LIEM

0. Introduction
1. Tagmemic and Pre-Aspects Transformational Procedure
2. Case Grammar
3. Generative-Transformational Insights
4. Statistical Studies
5. Theoretical Considerations

## 0. INTRODUCTION

In 1933, Leonard Bloomfield (1933:496-510) discussed the ways in which the findings of linguistics could be applied to the teaching of foreign languages. Since then, throughout governmental and university language programmes cited by Carroll (1959), Moulton (1962), and Hodge (1963) for example, a concept of applying modern linguistics to the teaching of foreign languages has been developed in various countries throughout the world. This concept of applied linguistics has been systematised mainly by Charles C. Fries (1945), and Robert Lado (1957). Representative of the recognition of the development of a methodology for the application of findings of linguistics to the teaching of foreign languages are the numerous contrastive analysis publications and doctoral dissertations in the fifties and the sixties.

The central idea of the concept of applied linguistics is the "importance of the native language in foreign language learning", Reed, Lado, and Shen (1953:121-7). The foreign language learner, being so much used to the phonological and grammatical systems of his native

language, tends to transfer them to the language to be learned. This transfer creates a phenomenon of interference (Weinreich 1953:3). The power of contrastive analysis is then, according to the theory, to point out the similarities and differences in the two linguistic systems. Wherever there are differences, there are teaching and learning difficulties. Thus, contrastive analysis should naturally serve as the basis for materials preparation, since, as Fries says:

"The most efficient materials are those that are based upon a scientific description of the language to be learned, carefully compared with a parallel description of the native language of the learner."

Another important concept of applied linguistics is that language is not a self-contained system, but that linguistic aspects are closely related to total human behaviour (Pike 1954, 1955, and 1960). Since language is only one part of a larger totality of structured human behaviour, and behaviour is a set of habits, language learning should be overlearned by such ways as drill, as E. Picazo de Murray says:

"Language is skill, and skill is the result of habit."  
(UNESCO 1953:53).

The immediate corollary of such an axiom is that, as J.H. Cooper *et al* put it:

"The key to language learning is (well planned and lively) drill, drill, drill." (1963:9).

The notion of habit forming through drill can be found in the Lado-Fries series of English for foreigners (1954, 1956, 1957, and 1958), or the author's series of materials of English for Vietnamese (1962, 1963a, and 1963b).

At the same time, because the phenomenon of interference is observed not only in situations of languages in contact but also in those of culture in contact, the techniques of contrastive linguistic analysis was extended to contrastive cultural studies to find out cross-cultural differences, as Albert H. Markwardt says:

"Just so, it may be reasonably maintained that contrastive cultural analyses are equally important in terms of language study." (1963:1-4).

However nowadays, all this golden era of contrastive analysis seems to be a matter of the past forever gone. The claim that the best language-teaching materials are those based upon a contrast of the two interfering linguistic systems has been in the last few years seriously challenged by the generative-transformationalists. Whereas the Sapir-Whorf hypothesis claimed that the structure of a language subtly influences the cognitive processes of the speakers of that language, the generative-transformationalist proposes that infants are innately endowed with the ability of acquiring a natural language, and that all



they need to process the learning of a natural language is an exposure to the data of that language. By postulating such a natural language learning hypothesis, the generative-transformational school can account for language universals, as well as the ability to learn a first foreign language with ease, and by the same token, the difficulty to learn a second language after childhood. (For an evaluation of contrastive analysis, see Ronald Wardhaugh, "The Contrastive Analysis Hypothesis", *TESOL Quarterly*, vol. 4, No. 2, June 1970, pp. 123-30). Also, by postulating such a cognitive hypothesis, the generative-transformationalist can claim that contrastive analysis does not serve any purpose in foreign language teaching and learning, for it is set upon a wrong concept of language learning processes. This completely negative reaction to contrastive analysis can be found in the views of Ritchie (1967), or Wolfe (1967).

The argumentation for the dismissal of contrastive analysis is this: All natural languages have a great deal in common in their deep structure. If there are language-specific differences, these are only due to diverse realisations of language universals (which by definition are common to all natural languages) in various languages. The deductions from such a hypothesis are two-fold: on the one hand, since all languages are similar in their deep structure, a particular language speaker knows already a great deal about any other language he has to learn. On the other hand, since languages differ only in their surface structure through differing transformational processes from a similar deep structure, the power of contrastive analysis should not be in its presentation of surface structure differences between languages, but should rather be in its capability of showing the differing processes involving language universals in the deep structure to language-specific realisations in their particular surface structure. Contrastive analysis works in the fifties and sixties would not likely pass such a test to measure their power, for most of them dealt only with surface structure differences. Such recent suggestions for using generative-transformational theory into contrastive analysis have been advocated by Ritchie (1968) who tried to utilise distinctive feature hierarchies in the explanation of phonological interference phenomena.

In the midst of such a theoretical controversy, the language teacher cannot help but feel obliged to re-examine contrastive analysis as well as the theories advocating or rejecting it. On the theoretical ground, he finds that, although the Sapir-Whorf behavioural hypothesis and the generative-transformational language acquisition device hypothesis are contradictory to each other, they are both as difficult to prove correct or incorrect at the present state of the art. On the practical

ground, however, he also notices in the classroom situation that some pitfalls in foreign language learning are more difficult to overcome than others, and that the surface structure is as important, to be on the conservative side, as the deep structure in language teaching and learning. In these circumstances, the language teacher cannot help but feel like going back to planting his own garden, as Voltaire says, and try to make the most use of competing linguistic theories to improve his daily teaching and text preparation.

With the spirit of pragmatism above all, I am proposing an eclectic contrastive analysis of English and Vietnamese clause units, and trying to see how to apply tagmemics, case grammar, pre-Aspects transformational grammar, and generative-transformational grammar to present surface structure differences and deep structure similarities between the two languages. Finally, I shall try to point out that, besides grammaticalness and acceptability (Chomsky 1965:11) of a grammatical construction, the frequency of occurrence (Cook 1965, 13-44) of a grammatical construction is also an important factor to be considered in applied linguistics.

#### 1. TAGMEMICS AND PRE-ASPECTS TRANSFORMATIONAL PROCEDURE

The tagmemic model of contrastive analysis advocated here is based upon the theory of language developed by Pike (1954, 1955, and 1960), later refined by Longacre (1964), Cook (1969), and Young, Becker, and Pike (1970). It was tested in contrastive analysis by the author (Liem 1966, 1967, 1969, and 1970a). Tagmemics offers an excellent methodology for contrastive analysis because it postulates *patterning* as being central to human behaviour in general and linguistic behaviour in particular. It presents linguistic patterns in straightforward and summary fashion. Such patterns when systematically described for one language can be contrasted with similar patterns described for another language. Thanks to the notion of hierarchy in tagmemics, such patterns can be contrasted on clearly cut levels of consideration, such as on the sentence level, the clause level, or the phrase level. Furthermore, the notion of a tagmeme as a slot plus a filler-class permits us to see whether patterns in two languages differ in their tagmemes (i.e. patterns in one language have some tagmemes that patterns in the other do not), or just in the filler-classes of their tagmemic slots (i.e. patterns in two languages contain the same functional slots, but the slots are filled by a different number of distributional-subclasses).

In order to see how contrastive analysis works in a tagmemic model, I am presenting the clause units in English and Vietnamese in contrast.

There are 87 Clause Units in English, and 108 in Vietnamese. The Clause Units in either language are cast in a two-dimensional field: the Clause Class Dimension, and the Clause Type Dimension (see the two Charts, pp. 116 and 117). The Clause Type Dimension is subdivided, under four levels of consideration, into ten Clause Types in English, and is subdivided, also under four levels of consideration, into nine Clause Types in Vietnamese. The Clause Class Dimension is subdivided, under four levels of consideration, into ten Clause Classes in English, and is subdivided, under five levels of consideration, into thirteen Clause Classes in Vietnamese.

### 1.1. CLAUSE TYPES IN CONTRAST

If the Clause Types in English and in Vietnamese are set side by side, their similarities and differences can be noticed easily:

| ENGLISH                                                                                                                       | VIETNAMESE                                                                       |
|-------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------|
| E.a1. Intransitive:<br>[+S +Pr]<br><i>He went.</i>                                                                            | V.a1. Intransitive:<br>[±S +Pr]<br>Nó đi.                                        |
| E.a2. Transitive:<br>[+S +Pr +O]<br><i>He bought books.</i>                                                                   | V.a2. Transitive:<br>[±S +Pr ±O]<br>Nó mua sách.                                 |
| E.a3. Double Transitive:<br>[+S +Pr +IO +O]<br><i>He gave them books.</i>                                                     | V.a3. Double Transitive:<br>[±S +Pr +IO ±O]<br>Nó cho họ sách.                   |
| E.a4. Attributive Transitive:<br>[+S +Pr +O +AtCompl]<br><i>They elected him chairman.</i>                                    | V.a4. Attributive Transitive:<br>[±S +Pr ±O +AtCompl]<br>Họ bầu nó làm chủ-tịch. |
| E.a5. Passive Transitive:<br>[+S +PasPr ±A]<br><i>It was bought by him.</i>                                                   | Non-existent.                                                                    |
| E.a6. Passive Double Transitive:<br>[+S +PasPr +(I)O ±A]<br><i>It was given (to) her by him.<br/>She was given it by him.</i> | Non-existent.                                                                    |
| E.a7. Passive Attributive Transitive:<br>[+S +PasPr +AtCompl ±A]<br><i>He was elected chairman by them.</i>                   | Non-existent.                                                                    |

SUMMARY TABLE  
THE 87 ENGLISH COMPLETE CLAUSE UNITS

|                            |                                        |                                |                              | 18.27 | 44.00               | 2.45   | 3.55   | 2.21             | .04    | .08    | 25.65            | 2.05            | 1.70            | 100%         |            | PERCENTAGE         |       |  |
|----------------------------|----------------------------------------|--------------------------------|------------------------------|-------|---------------------|--------|--------|------------------|--------|--------|------------------|-----------------|-----------------|--------------|------------|--------------------|-------|--|
| I                          | II                                     | III                            | IV                           | a     | 463                 | 1115   | 62     | 90               | 56     | 1      | 2                | 650             | 52              | 43           | 2534       | 100%               | TOTAL |  |
| Inde-<br>pendent           | Without an<br>Interrogative<br>Tagmeme | Declarative                    | a                            | 227   | 599                 | 30     | 49     | 31               | 0      | 1      | 382              | 43              | 35              | 1397         | 55.13      |                    |       |  |
|                            |                                        | Imperative                     | b                            | 85    | 162                 | 18     | 23     | 2                | 0      | 0      | 14               | UG              | UG              | 304          | 2.00       |                    |       |  |
|                            |                                        | Yes-No Interrogative           | c                            | 23    | 55                  | 2      | 5      | 1                | 0      | 0      | 0                | 45              | 1               | 3            | 135        | 5.37               |       |  |
|                            | With an<br>Interrogative<br>Tagmeme    | From Declarative<br>Kernel     | Interrogative<br>Subject     | d     | 2                   | 3      | 1      | 0                | 0      | 0      | 0                | 30              | UG              | UG           | 36         | 1.42               |       |  |
|                            |                                        | From Inter-<br>rogative Kernel | Interrogative<br>Non-Subject | e     | UG                  | 48     | 0      | 0                | UG     | 0      | 0                | 13              | 0               | 1            | 62         | 2.45               |       |  |
|                            |                                        |                                | Extra-<br>Interrogative      | f     | 10                  | 10     | 0      | 0                | 0      | 0      | 0                | 37              | 0               | 0            | 57         | 2.25               |       |  |
| Dependent                  | Without an Extra Dependent<br>Tagmeme  | Dependent<br>Subject           | g                            | 20    | 22                  | 0      | 2      | 4                | 0      | 0      | 6                | UG              | UG              | 54           | 2.13       |                    |       |  |
|                            |                                        | Dependent<br>Non-Subject       | h                            | UG    | 70                  | 5      | 2      | UG               | 1      | UG     | 10               | UG              | UG              | 88           | 3.47       |                    |       |  |
|                            | With an Extra Dependent Tagmeme        | Dependent<br>Relative          | i                            | 42    | 74                  | 1      | 3      | 6                | 0      | 1      | 56               | 6               | 4               | 193          | 7.62       |                    |       |  |
|                            |                                        | Extra<br>Dependent             | j                            | 54    | 72                  | 5      | 6      | 12               | 0      | 0      | 57               | 2               | 0               | 208          | 8.21       |                    |       |  |
| The Clause Class Dimension |                                        |                                |                              | IV    | 1                   | 2      | 3      | 4                | 5      | 6      | 7                | 8               | 9               | 10           |            |                    |       |  |
|                            |                                        |                                |                              | III   | Intransitive        | Single | Double | Attrib-<br>utive | Single | Double | Attrib-<br>utive | Equational      | (There) Stative | (It) Stative | TOTAL      | PERCENTAGE         |       |  |
| The Clause Type Dimension  |                                        |                                |                              | II    | Transitive          |        |        | Transitive       |        |        | Equational       | (There) Stative | (It) Stative    | TOTAL        | PERCENTAGE |                    |       |  |
|                            |                                        |                                |                              | I     | Active              |        |        | Passive          |        |        |                  |                 |                 |              |            | Logical<br>Subject |       |  |
|                            |                                        |                                |                              | I     | Grammatical Subject |        |        |                  |        |        |                  |                 |                 |              |            |                    |       |  |

TOTAL: of each Clause Class or Clause Type

PERCENTAGE: of each Clause Class or Clause Type

O: a grammatical Clause Unit that has not occurred in the materials

UG: ungrammatical

SUMMARY TABLE  
THE 108 VIETNAMESE CLAUSE UNITS AND THEIR OCCURRENCES

|                                       |                                    |                       |                                          | 21.29                                | 37.25          | 4.54       | 6.81 | 2.14 | 1.56           | 8.96 | 13.69   | 3.76  | 100%       |      | PERCENTAGE |       |               |  |  |  |
|---------------------------------------|------------------------------------|-----------------------|------------------------------------------|--------------------------------------|----------------|------------|------|------|----------------|------|---------|-------|------------|------|------------|-------|---------------|--|--|--|
| I                                     | II                                 | III                   | IV                                       | V                                    | a              | b          | c    | d    | e              | f    | g       | h     | i          | j    | k          | l     | m             |  |  |  |
| Independent                           | Non-Interrogative                  | Declarative           |                                          | a                                    | 328            | 574        | 70   | 105  | 33             | 24   | 138     | 211   | 58         | 1541 | 100%       | TOTAL | 65.67         |  |  |  |
|                                       |                                    | Imperative            |                                          | b                                    | 19             | 33         | 8    | 9    | 1              | 0    | 1       | 3     | UG         | 74   | 4.80       |       |               |  |  |  |
|                                       |                                    | Alternate             |                                          | c                                    | 0              | 4          | 1    | 2    | 0              | 0    | 8       | 5     | 0          | 20   | 1.30       |       |               |  |  |  |
|                                       | Interrogative                      | With không            | Yes-No Interrogative                     |                                      | d              | 2          | 11   | 3    | 5              | 0    | 1       | 0     | 8          | 4    | 34         | 2.21  |               |  |  |  |
|                                       |                                    |                       | Right-Wrong Interrogative                |                                      | e              | 4          | 3    | 1    | 3              | 1    | 0       | 6     | 6          | 0    | 24         | 1.56  |               |  |  |  |
|                                       |                                    | With chu-a            | Realisation-Nonrealisation Interrogative |                                      | f              | 2          | 9    | 0    | 0              | 0    | 0       | 0     | 0          | 0    | 11         | .71   |               |  |  |  |
|                                       |                                    |                       | With an Interrogative Tagmeme            | With a nucleus Interrogative Tagmeme | Inter. Subject | g          | 2    | 4    | 1              | 3    | 0       | 0     | 2          | 8    | UG         | 20    | 1.30          |  |  |  |
|                                       |                                    | Inter. Non-Subject    |                                          |                                      | h              | UG         | 21   | 4    | 1              | 0    | 0       | 2     | UG         | 4    | 32         | 2.08  |               |  |  |  |
|                                       |                                    | Extra Interrogative   |                                          | i                                    | 28             | 31         | 1    | 8    | 4              | UG   | 1       | 7     | 6          | 86   | 5.58       |       |               |  |  |  |
| Dependent                             | Without an Extra-Dependent Tagmeme | Dependent Subject     |                                          | j                                    | 15             | 14         | 0    | 2    | 0              | 2    | 4       | 1     | UG         | 38   | 2.47       |       |               |  |  |  |
|                                       |                                    | Dependent Non-Subject |                                          | k                                    | UG             | 17         | 2    | 1    | UG             | UG   | 1       | UG    | UG         | 21   | 1.36       |       |               |  |  |  |
|                                       | With an Extra-Dependent Tagmeme    | Dependent Relative    |                                          | l                                    | 7              | 16         | 1    | 0    | 0              | 2    | 8       | 4     | 0          | 38   | 2.47       |       |               |  |  |  |
|                                       |                                    | Extra Dependent       |                                          | m                                    | 33             | 49         | 3    | 4    | 2              | 2    | 113     | 19    | 6          | 13   | 8.50       |       |               |  |  |  |
| The Thirteen Clause Classes Dimension |                                    |                       |                                          | IV                                   |                |            |      |      |                |      |         |       |            |      |            |       |               |  |  |  |
|                                       |                                    |                       |                                          | III                                  | Intransitive   | 1          | 2    | 3    | 4              | 5    | 6       | 7     | 8          | 9    |            |       |               |  |  |  |
|                                       |                                    |                       |                                          | II                                   | Transitive     |            |      |      |                |      |         |       |            |      |            |       |               |  |  |  |
|                                       |                                    |                       |                                          | I                                    | Active         | Submissive |      |      | Identificative |      | Stative | TOTAL | PERCENTAGE |      |            |       |               |  |  |  |
| The Nine Clause Types Dimension       |                                    |                       |                                          | Grammatical Subject Class            |                |            |      |      |                |      |         |       |            |      |            |       | Sub-ject-less |  |  |  |

TOTAL: of the Clause Class  
0: a grammatical Clause Unit that has not occurred in the material  
UG: ungrammatical

Clause Types in Contrast (continued from page 115)

| ENGLISH                                | VIETNAMESE                 |
|----------------------------------------|----------------------------|
| Non-existent.                          | V.a5. Actional Submissive: |
|                                        | [±S +SubmPr +SubmCompl]    |
| <i>he undergo-bad-experience-of go</i> | Nó bị đi.                  |
| <i>He had to go.</i>                   |                            |
| Non-existent.                          | V.a6. Passive Submissive:  |
|                                        | [±S +SubmPr +SubmPasCompl] |
| <i>he undergo...(they) reprimand</i>   | Nó bị (họ) rầy.            |
| <i>He was reprimanded (by them).</i>   |                            |
| E.a8. Equational:                      | V.a7. Equational:          |
| [+S +EqPr +EqCompl]                    | [±S +EqPr +EqCompl]        |
| <i>They are carpenters.</i>            | Họ là thợ-mộc.             |
| Same as E.a8. above.                   | V.a8. Adjective:           |
|                                        | [±S +AdjPr]                |
| <i>he fast</i>                         | Nó lẹ.                     |
| <i>He is fast.</i>                     |                            |
| E.a9. 'there' Stative:                 | V.a9. 'co' Stative:        |
| [+there +StPr +StS]                    | [+cóPr +StCompl]           |
| <i>There were two people.</i>          | Có hai người.              |
| E.a10. 'it' Stative:                   | Non-existent.              |
| [+it +StPr +StCompl]                   |                            |
| <i>It was the boys.</i>                |                            |

The tagmemic contrastive analysis advocated points out the following dissimilarities:

- (1) The nuclear Subject tagmeme is obligatory in English, and is optional in Vietnamese. This is confirmed in practical situations where Vietnamese speaking English actually omit the Subject.
- (2) The nuclear Object is obligatory in English, and is optional in Vietnamese. This is also confirmed in the same way as for (1).
- (3) The Passive Transitive Clause Types E.a5-7 in English are non-existent in Vietnamese. The following pre-Aspects surface-to-surface transformational rule is advocated to derive these Passive Clause Types from the Active Transitive Clause Types E.a3-4:

- KERNEL STRUCTURES: E.a2. [+S +Pr +O]  
*He bought it.*
- E.a3. [+S +Pr +IO +O]  
*He gave her the flowers.*
- E.a4. [+S +Pr +O +AtCompl]  
*They elected him chairman.*

## TRANSFORM RULE

## Passive

- 
- E.a5. [+S +PasPr ±A]  
*It was bought (by him).*
- E.a6. [+S +PasPr +(I)O ±A]  
*She was given the flowers (by him).*  
*The flowers were given (to) her (by him).*
- E.a7. [+S +PasPr +AtComp ±A]  
*He was elected chairman (by them).*

## 1.2. CLAUSE CLASSES IN CONTRAST

The Independent Declarative Clause Class, whose Clause Types are listed in 1.1., is the kernel Clause Class from which all the other Clause Classes can be derived in both English and Vietnamese when appropriate transformational rules are applied. Thus a contrastive analysis of the Clause Classes in the two languages consists necessarily and sufficiently of a comparison of the Clause Class Formulas, and that of their Transformational Rules:

## ENGLISH

E.b. Imperative:

[+ImpPr...]

*Go! (Mr)*

Similar to E.c. below.

*Mr go or Mr stay**(Are you going or are you staying?)*

E.c. Yes-No Interrogative:

[+↑ +S +InterPr...]*Are you going? (no?)*

## VIETNAMESE

V.b. Imperative:

[+S +ImpPr...]

*Ông hãy đi!*

V.c. Alternative Interrogative:

[+DeclCl +AlternPhrase +DeclCl]

*Ông đi hay ông ở?*

V.d. Yes-No Interrogative:

[+DeclCl +YesNoInterPhrase]

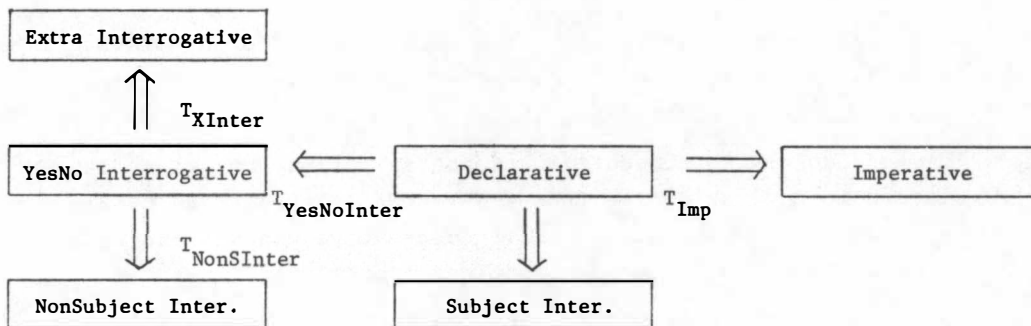
*Ông đi không?*

|                                                                                                          |                                                                                                  |
|----------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------|
| Realised as an Echo Interrogative.                                                                       | V.e. Right-Wrong Interrogative:<br>[+DeclCl +RWInterPhrase]<br>Ông đi phải không?                |
| <i>Mr went right no</i><br>( <i>You went, didn't you?</i> )                                              |                                                                                                  |
| Realised as a Yes-No Interrogative.                                                                      | V.f. Realisation-Nonrealisation<br>Interrogative:<br>[+DeclCl +RnonRInterPhrase]<br>Ông đi chưa? |
| <i>Mr go no-yet</i><br>( <i>Did you go?</i> )                                                            |                                                                                                  |
| E.d. Subject Interrogative:<br>[+InterS +Pr...]<br><i>Who went?</i>                                      | V.g. Subject Interrogative:<br>[+InterS +Pr...]<br><i>Ai đi?</i>                                 |
| E.e. Non-Subject Interrogative:<br>[+ <u>InterNonS</u> [YesNoInterCl -NonS]]<br><i>What did you buy?</i> | V.h. Non-Subject Interrogative:<br>[±S +Pr +InterNonS]<br><i>Ông mua gì?</i>                     |
| E.f. Extra Interrogative:<br>[+XInterIntroducer +YesNoInterCl]<br><i>Where did you go?</i>               | V.i. Extra Interrogative:<br>[+DeclCl +XInterIntroducer]<br><i>Ông đi đâu?</i>                   |
| E.g. Subject Dependent:<br>[ <u>DepS</u> +Pr...]<br><i>that went...</i>                                  | V.j. Subject Dependent:<br>[±DepS +Pr...]<br><i>(mà) đi...</i>                                   |
| E.h. Non-Subject Dependent:<br>[±DepNonS [+DeclCl -NonS]]<br><i>(that) you bought...</i>                 | V.k. Non-Subject Dependent:<br>[±DepNonS [+DeclCl -NonS]]<br><i>(mà) ông mua...</i>              |
| E.i. Relative Dependent:<br>[±RelDepIntroducer +DeclCl]<br><i>(that) he went.</i>                        | V.l. Relative Dependent:<br>[±RelDepIntroducer +DeclCl]<br><i>(rằng) nó đi.</i>                  |
| E.j. Extra Dependent:<br>[+XDepIntroducer +DeclCl]<br><i>When he went...</i>                             | V.m. Extra Dependent:<br>[+XDepIntroducer +DeclCl]<br><i>Khi nó đi...</i>                        |

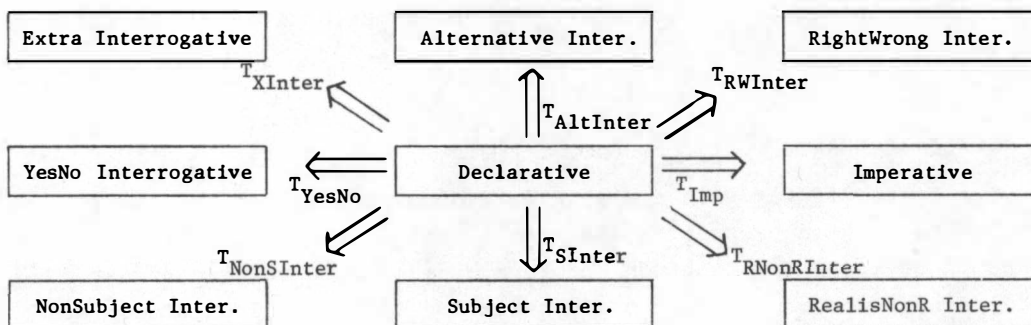
Because all these Clause Classes in both languages can be derived, on the surface structure consideration, from the Independent Declarative Clause Class when appropriate surface Clause Class Transformational Rules are applied, a contrastive analysis should point out the transformational differences as follows:



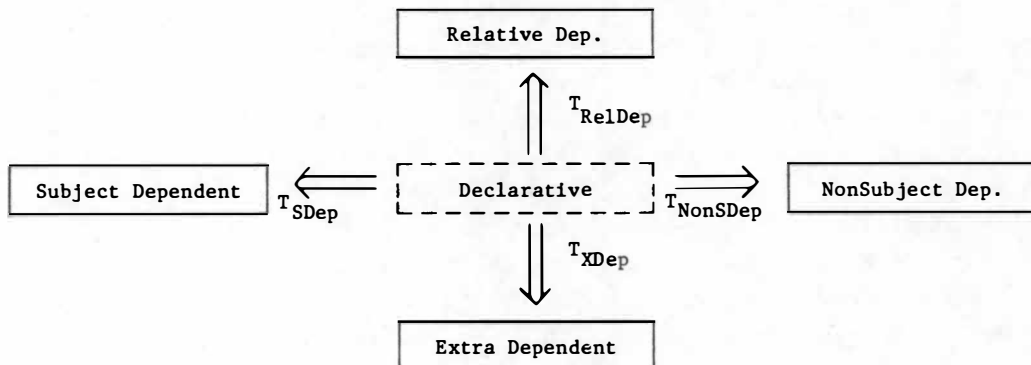
## ENGLISH INDEPENDENT CLAUSE CLASSES AND TRANSFORM RULES



## VIETNAMESE INDEPENDENT CLAUSE CLASSES AND TRANSFORM RULES



## ENGLISH OR VIETNAMESE DEPENDENT CLAUSE CLASSES AND TRANSFORM RULES



From these diagrams, the similarities and differences in the two languages can be easily seen. The dissimilarities stem only from the In-

dependent Clause Classes, while the Dependent Clause Classes are quite similar. The dissimilarities are as follows:

- (1) The Yes-No Interrogative Transform Rules are different: in English, there is an inversion of the order of the Subject and Predicate tagmemes, and in Vietnamese, there is no such an inversion.
- (2) The Non-Subject Interrogative and the Extra Interrogative Clause Classes in English derive from the Yes-No Interrogative Clause Class in English. In Vietnamese, all Interrogative Clause Classes derive directly from the Independent Declarative Clause Class.
- (3) The Right-Wrong, Alternative, and Realisation-Nonrealisation Interrogative Clause Classes in Vietnamese do not have equivalents in English.

The surface transformational rules to derive one clause class from another are also advocated to underlie transformational drills in language exercises.

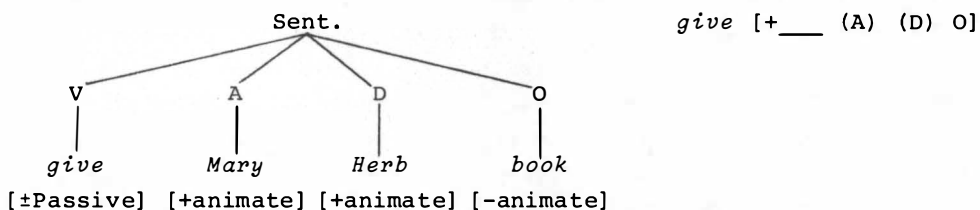
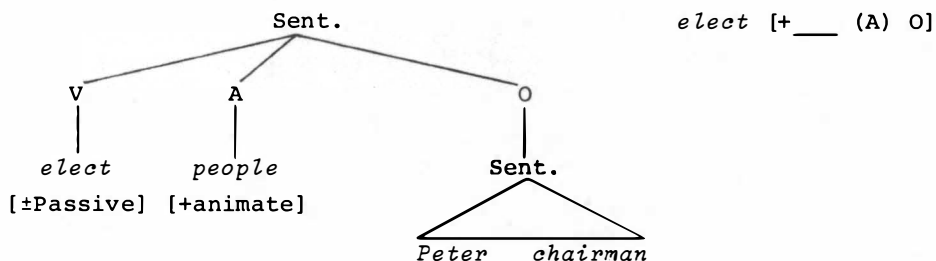
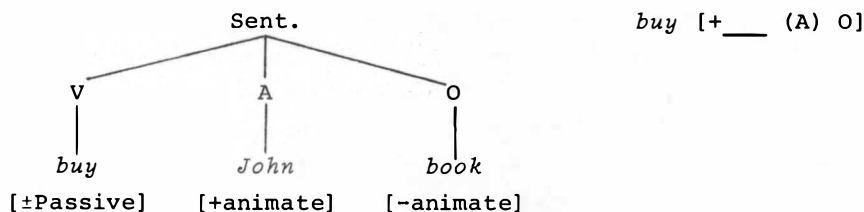
## 2. CASE GRAMMAR

The system of tagmemic analysis as demonstrated above is a well defined system in presenting the grammatical elements of surface structure in terms of the function and form of each unit being contrasted. Its primary concern is with the description and comparison of surface structure. The matrix displays and transformational rules presented above attempt to show the relationships between sentences with diverse surface structures, for example the passive clause types are conceived as related to the active ones.

It is also the belief of the author that the use of case grammar will enrich contrastive analysis a great deal, especially in the presentation of the deep structure and the relationships between different surface structure patterns that have a common deep structure. Case grammar as developed by Charles Fillmore (1968, 1970a, and 1970b), is a system which views the deep structure of sentences as a set of relations between a verb and a series of case-marked noun phrases. The series of noun phrases constitutes a set of roles which are useful in classifying verbs in terms of the case frames in which they occur. The system is particularly powerful in relating sentences with identical deep structures but diverse surface structures. The usefulness of case grammar techniques to contrastive analysis can be stated as follows:

- (1) By considering the deep structure cases, the theory can be related to different surface structures that have a common deep structure

among themselves, thus, it can provide a more powerful means of presenting transformational drills in language teaching. For example, the verbs *buy*, *elect*, and *give* which appeared in the examples above can be conceived with their case-marked noun phrases as follows:



With the deep structures realised as verbs and a series of case roles, these deep structures must then be mapped on to the surface structures by a series of transformational rules, which are called realisation rules. (In 1970, Fillmore proposed the following realisation rules: 1. Subject raising (optional); 2. Co-reference deletion; 3. Dative (or Experiencer) shunting; 4. Psych movement; 5. Accusative marking; 6. Passive rule (optional); 7. Nominative marking; 8. Subject formation; 9. Extraposition; and 10. Object formation for English.

By applying case grammar to contrastive analysis, we can see that the three verbs above are marked ±Passive in English, and they are not marked that way in Vietnamese. We can then apply the optional Passive rule to form passive constructions in English, and present

the constructions in a set of transformational drills.

- (2) The second application of case grammar to contrastive analysis can be seen as this: In case grammar, verbs are classified according to the case frames hospitable to them. Thus, a verb such as *give* has the case frames A, D and O, whereas a verb such as *elect* has the case frames A and O only. This is why *give* can have two passive realisations and *elect* can have only one. Such a classification of verbs according to their case frames would be useful in the preparation of transformational drills in eliminating ungrammatical constructions.
- (3) By using case grammar, contrastive analysis will be able to make distinctions, and consequently present in class, between sentences because of their different deep structures. For example, the two sentences *John is eager to please* and *John is easy to please* will have *John* in the first sentence marked with the Agentive role, while it will be marked with the Objective case in the second sentence. Such differences should be made in drills as well.

### 3. GENERATIVE TRANSFORMATIONAL INSIGHTS

The transformational rules presented in 1. showed the surface relationships between various surface structures. Today, we possess some other notions about language. We are now concerned with two kinds of structures, surface structure and deep structure. In contrastive analysis, we also want to relate the surface differences between languages to the language universals in the deep structure. Whether we want to do it in a case grammar framework or in a generative transformational framework (the kind posited by Chomsky), contrastive analysis will be enriched a great deal.

### 4. STATISTICAL STUDIES

Contrastive analysis points out the similarities and differences in two or more languages under consideration. Its application to foreign language teaching should be then based upon statistical studies of actual errors made by the learner in order to have a hierarchy of difficulty. Furthermore, language is communication. In foreign language teaching, the frequency of occurrence of grammatical constructions should also be considered. The statistical study of grammar as presented in the two charts would serve to establish which grammatical constructions should receive priority in a language teaching textbook. (In order to

see the statistical techniques utilised, see Liem 1970b, and 1970c.)

## 5. THEORETICAL CONSIDERATIONS

It is the belief of the author that contrastive analysis should not be rejected because one belongs to a linguistic camp. On the contrary, if one makes use of existing theories and techniques, one will be able to improve contrastive analysis and make use of it in foreign language teaching until another cognitive hypothesis is proved solid and capable of destroying the present overlearning character of language learning.

## REFERENCES

- BLOOMFIELD, Leonard  
1933 *Language*. Holt and Co., New York.
- CARROLL, John B.  
1959 *The Study of Language*. Harvard University Press, Cambridge.
- CHOMSKY, Noam  
1965 *Aspects of a Theory of Syntax*. The M.I.T. University Press, Cambridge.
- COOK, Walter A.  
1965 "The Generative Power of a Tagmemic Grammar". *Monograph Series on Languages and Linguistics, No. 20, Linguistics and Language Study*. Blansitt (ed.). Georgetown University Press, Washington D.C.  
1969 *Introduction to Tagmemic Analysis*. Holt, Rinehart, and Winston Inc., New York.
- COOPER, J.H. *et al.*  
1963 *Methods and Techniques for Teaching English*. English Language Services, Washington D.C.
- DE MURRAY, E. Picazo  
1955 "Teaching Spanish to Students from the U.S.". *The Teaching of Modern Languages*. (Proceedings of UNESCO Seminar in Ceylon, 1953.) Drukkerij, Amsterdam.
- FILLMORE, Charles  
1968 "The Case for Case". *Universals in Linguistic Theory*. Emmon Bach and Robert T. Harms (eds.), 1-88. Holt, Rinehart and Winston Inc., New York.

- 1970a "Subjects, Speakers and Roles". *The Ohio State University Working Papers in Linguistics*. No.4, pp. 31-63.
- 1970b "On Generativity". *The Ohio University Working Papers in Linguistics*. No.6, pp. 2-19.
- FRIES, Charles C.
- 1945 *Teaching and Learning English as a Foreign Language*. University of Michigan Press, Ann Arbor.
- HODGE, Calton T.
- 1963 "The Influence of Linguistics on Language Teaching". *Anthropological Linguistics*, 5:50-58.
- LADO, Robert
- 1957 *Linguistics Across Cultures*. University of Michigan Press, Ann Arbor.
- LADO, Robert and Charles C. Fries
- 1954 *English Pronunciation*. University of Michigan Press, Ann Arbor.
- 1956 *English Sentence Patterns*. University of Michigan Press, Ann Arbor.
- 1957 *English Pattern Practices*. University of Michigan Press, Ann Arbor.
- 1958 *Lessons in Vocabulary*. University of Michigan Press, Ann Arbor.
- LIEM, Nguyen-Dang
- 1962 *English Pronunciation for Vietnamese*. Southeast Asian Regional English Project (SEAREP), University of Michigan - United States Operations Missions, Saigon.
- 1963a *English Grammar for Vietnamese*. SEAREP, Saigon.
- 1963b *English Pattern Practice for Vietnamese*, vols. I and II. SEAREP, Saigon.
- 1966 *English Grammar, a Combined Tagmemic and Transformational Approach*. The Australian National University (ANU), Canberra.
- 1967 *A Contrastive Grammatical Analysis of English and Vietnamese*. ANU, Canberra.
- 1969 *Vietnamese Grammar, a Combined Tagmemic and Transformational Approach*. ANU, Canberra.
- 1970a "Clause Units in Vietnamese". *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung*. Band 23, Heft 1: 20-30.

- 1970b "Clause Unites in English". *Pacific Linguistic Studies in Honour of Arthur Capell*, pp. 97-108. ANU, Canberra.
- MARKWARDT, Albert H.  
 1963 "The Cultural Preparation of the Teacher of English as a Foreign Language". *Selected Articles from Language Learning*. No.2, pp. 1-4.
- MOULTON, William G.  
 1962 "Linguistic and Language Teaching in the United States". *Trends in European and American Linguistics 1930-1960*, pp. 82-109. Spectrum Publishers, Utrecht.
- PIKE, Kenneth L.  
 1954, *Language in Relation to a Unified Theory of the Structure of*  
 1955, *Human Behavior*. Vols. I, II, and III. Summer Institute of  
 1960 Linguistics, Glendale.
- REED, David D., Robert Lado, and Yao Shen  
 1953 "The Importance of the Native Language in Foreign Language Learning". *Selected Articles from Language Learning*. No.1, pp. 121-7.
- RITCHIE, William C.  
 1967 "Some Implications of Generative Grammar for the Construction of Courses in English as a Foreign Language". *Language Learning*, 17:45-49, 111-131.  
 1968 "On the Explanation of Phonic Interference". *Language Learning*, 18:183-197.
- WARDHAUGH, Ronald  
 n.d. "The Contrastive Analysis Hypothesis". *TESOL Quarterly*. Vol. 4, No.2, pp. 123-130.
- WEINREICH, Uriel  
 1953 *Language in Contact*. Linguistic Circle of New York, New York.
- WOLFE, David L.  
 1967 "Some Theoretical Aspects of Language Learning and Language Teaching". *Language Learning*, 17:173-188.
- YOUNG, Richard E., Alton L. Backer, and Kenneth L. Pike  
 1970 *Rhetoric: Discovery and Change*. Hartcourt, Brace and World, New York.



# CLAUSES AND CASES IN ENGLISH AND SOUTHEAST ASIAN LANGUAGES (BURMESE, CAMBODIAN, CANTONESE, LAO, THAI, AND VIETNAMESE) IN CONTRAST

NGUYEN DANG LIEM

0. Introduction
1. Clause Units in the Languages
2. Clause Units in Contrast
3. Deep versus Surface Structures
4. Philosophy of Approach

## 0. INTRODUCTION

For over a generation, the conviction that the best language-teaching materials are based upon a contrastive analysis of the language to be learned and the language of the learner has been predominant in foreign language teaching. Allied with the conviction was the hypothesis on language learning which assumed that the new linguistic system, and by extension the whole new cultural behaviour, should be established as a set of new habits by drill, drill, and drill which would ensure over-learning. Such a pedagogical philosophy was systematised mainly by Charles C. Fries (1945) and Robert Lado (1957).

However, all this firm belief in contrastive analysis seems to be in the past, at least for some people. The generative-transformational theory, which was born in 1957 with Noam Chomsky's *Syntactic Structures*, claims that language behaviour is rule-governed creative behaviour, and consequently, language learning should be in the form of a process of internalising the creative rules in the new language, and not just that

of mere habit forming. The theory is concerned not only with the actual utterances, the surface structures of a language, but even more so with meaning, the deep structure of universal language, and with the various transformational rules that map the deep structure denominations that are common to all languages to the surface structure realisations that are specific to particular languages. It explores not only the performance, but also the competence of language speakers (Chomsky 1965:3).

Before such a theoretical conflict, a language teacher may be tempted to make the most use of existing linguistic techniques to improve teaching materials as much as he can. While he may not be absolutely positive about the total efficiency of applied linguistics, he is likely to believe that surface structure is as important as deep structure in foreign language learning, for he constantly observes interference (Weinreich 1953:3) in situations of languages in contact.

With pragmatism in mind, I am trying to make use of various linguistic techniques in this contrastive analysis of English and Southeast Asian languages in this paper. Firstly, a surface structure presentation of clause units in each language will be given in tagmemic formulas (Pike 1954, 1955, 1960, Longacre 1964, Liem 1966, 1967, 1969, and 1970a, and Cook 1969) and two-dimensional matrix systems (Ray 1967). Secondly, a contrastive analysis will point out the surface structure differences between English on the one hand, and the Southeast Asian languages on the other. Thirdly, an attempt will be made to decide the deep structures of the surface structure differences found in 2. Finally, I shall conclude that contrastive analysis will continue to play a major role in language teaching and in area linguistics.

## 1. CLAUSE UNITS IN THE LANGUAGES

The tagmemic model utilised here was developed by Pike, and improved by his followers and himself (Young, Becker, and Pike 1970, and Cook 1971). It views language as hierarchically ordered. The clause hierarchy is in between the sentence and the phrase hierarchies. This paper presents an analysis of clauses because, as Longacre puts it: "In essence, the clause posits a situation in miniature (whether asserting, questioning, commanding, or equating" (1964:35). The clause tagmeme includes one or more phrase-level tagmemes, each of which has a functional slot and a filler class, and may be nuclear (i.e. essential to the clause) or satellite, obligatory or optional. This analysis will present only the minimal formulas of clauses, and will only present the functional slots in the clauses and not the filler classes of these

slots.\*

SUMMARY TABLE I  
THE 87 CLAUSE UNITS IN ENGLISH

| I                      |                                    | II                        |                           | III |  | IV |  | No.                 | 1            | 2          | 3           | 4          | 5      | 6               | 7               | 8            | 9           | 10 |   |   |
|------------------------|------------------------------------|---------------------------|---------------------------|-----|--|----|--|---------------------|--------------|------------|-------------|------------|--------|-----------------|-----------------|--------------|-------------|----|---|---|
| Independent            | Without an Interrogative tagmeme   | Declarative               |                           |     |  |    |  | 1                   | +            | +          | +           | +          | +      | +               | +               | +            | +           | +  | + |   |
|                        |                                    | Imperative                |                           |     |  |    |  | 2                   | +            | +          | +           | +          | +      | +               | +               | +            | +           | +  | + | + |
|                        |                                    | Yes-No Interrogative      |                           |     |  |    |  | 3                   | +            | +          | +           | +          | +      | +               | +               | +            | +           | +  | + | + |
|                        | With an Interrogative tagmeme      | From Declarative kernel   | Interrogative Subject     |     |  |    |  |                     | 4            | +          | +           | +          | +      | +               | +               | +            | +           | +  | + | + |
|                        |                                    | From Interrogative kernel | Interrogative Non-Subject |     |  |    |  |                     | 5            |            | +           | +          | +      |                 | +               | +            | +           | +  | + | + |
|                        |                                    |                           | Extra-Interrogative       |     |  |    |  |                     | 6            | +          | +           | +          | +      | +               | +               | +            | +           | +  | + | + |
| Dependent              | Without an Extra Dependent tagmeme | Dependent Subject         |                           |     |  |    |  | 7                   | +            | +          | +           | +          | +      | +               | +               | +            | +           | +  | + |   |
|                        |                                    | Dependent Non-Subject     |                           |     |  |    |  | 8                   |              | +          | +           | +          |        | +               | +               | +            | +           | +  | + |   |
|                        | With an Extra Dependent tagmeme    | Dependent Relative        |                           |     |  |    |  | 9                   | +            | +          | +           | +          | +      | +               | +               | +            | +           | +  | + |   |
|                        |                                    | Extra Dependent           |                           |     |  |    |  | 10                  | +            | +          | +           | +          | +      | +               | +               | +            | +           | +  | + |   |
| The ten Clause Classes |                                    |                           |                           |     |  |    |  | No.                 | 1            | 2          | 3           | 4          | 5      | 6               | 7               | 8            | 9           | 10 |   |   |
| The ten Clause Types   |                                    |                           |                           |     |  |    |  | IV                  | Intransitive | Transitive |             | Transitive |        | Equational      | (There) Stative | (It) Stative |             |    |   |   |
|                        |                                    |                           |                           |     |  |    |  |                     | Single       | Double     | Attributive | Single     | Double |                 |                 |              | Attributive |    |   |   |
|                        |                                    |                           |                           |     |  |    |  | III                 | Active       | Passive    |             |            |        |                 |                 |              |             |    |   |   |
|                        |                                    |                           |                           |     |  |    |  |                     |              |            |             |            |        |                 |                 |              |             |    |   |   |
| II                     |                                    |                           |                           |     |  |    |  |                     |              |            |             |            |        |                 |                 |              |             |    |   |   |
|                        |                                    |                           |                           |     |  |    |  |                     |              |            |             |            |        |                 |                 |              |             |    |   |   |
| I                      |                                    |                           |                           |     |  |    |  |                     |              |            |             |            |        |                 |                 |              |             |    |   |   |
|                        |                                    |                           |                           |     |  |    |  |                     |              |            |             |            |        |                 |                 |              |             |    |   |   |
|                        |                                    |                           |                           |     |  |    |  | Grammatical Subject |              |            |             |            |        | Logical Subject |                 |              |             |    |   |   |

### 1.1. ENGLISH

There are 87 Clause Units in English, which are cast in a two-dimensional field: the Clause Class Dimension, and the Clause Type

\*It is noted that the transcriptions of data in this paper are as close to phonemic transcriptions for Burmese, Cambodian, and Thai-Lao as practically acceptable to researchers in the languages, or follow the most commonly accepted transcriptions which may not be phonemic as it is for Cantonese. English and Vietnamese examples are recorded in current spelling systems of the languages. Phrase-level analyses are given only when pertinent, they occur mostly in Burmese. The author is deeply indebted to Arthur Crisfield, Thomas W. Gething, Philip N. Jenner, Julia C. Kwan, and D. Haigh Roop for their valuable data and enlightening insights into Lao, Thai, Cambodian, Cantonese, and Burmese respectively, and is without saying responsible for all the descriptive inadequacies found in this paper.

Dimension. The Clause Type Dimension is subdivided, under four levels of consideration, into ten Clause Types. The Clause Class Dimension is subdivided, under four levels of consideration, into ten Clause Classes. The total field contains one hundred possible Clause Units, but only 87 of them are grammatical and acceptable in English.

### 1.1.1. Clause Types in English

The minimal formulas of the ten Independent Declarative Clause Units in English are as follows:

- E.a1. Intransitive      [+S    +Pr]  
                               *He    went.*
- E.a2. Transitive        [+S    +Pr        +O]  
                               *He    bought    a book.*
- E.a3. Double  
       Transitive        [+S    +Pr        +IO    +O]  
                               *He    gave    her    a book.*
- E.a4. Attributive  
       Transitive        [+S        +Pr            +O    +AtCompl]  
                               *They    elected    him    chairman.*
- E.a5. Passive            [+S    +PassPr        ±A]  
                               *It    was bought    by him.*
- E.a6. Double Passive    [+S        +PassPr        +(I)O    ±A]  
                               *She    was given    a book (by him).*  
                               *It    was given    to her (by him).*
- E.a7. Attributive  
       Passive            [+S    +PasPr            +AtCompl    ±A]  
                               *He    was elected    chairman (by them).*
- E.a8. Equational        [+S    +EqPr    +EqCompl]  
                               *He    is        a student.*  
                               *He    is        intelligent.*
- E.a9. 'there'  
       Stative            [+there +StPr    +StS]  
                               *There    were    two people.*
- E.a10. 'it' Stative      [+it    +StPr    +StCompl]  
                               *It    was        the boys.*

### 1.1.2. Clause Classes in English

The minimal nuclear formulas of the remaining nine derived Clause Classes in English are as follows:

|       |                              |                                                                                                             |
|-------|------------------------------|-------------------------------------------------------------------------------------------------------------|
| E. b. |                              | [+ImpPr...]<br><i>Go!</i>                                                                                   |
| E. c. | Yes-No<br>Interrogative      | [<+S <+Pr...]<br><i>Are you going?</i>                                                                      |
| E. d. | Subject<br>Interrogative     | [+InterS +Pr...]<br><i>Who is going?</i>                                                                    |
| E. e. | Non-Subject<br>Interrogative | [+InterNonS +[+YesNoInterCl -NonS]]<br><i>What did he buy?</i><br><i>Who did he give it to?</i>             |
| E. f. | Extra Inter-<br>rogative     | [+XInterIntroducer +YesNoInterCl]<br><i>Where did he go?</i>                                                |
| E. g. | Subject<br>Dependent         | [+DepS +Pr...]<br><i>...that went...</i>                                                                    |
| E. h. | Non-Subject<br>Dependent     | [±DepNonS +DeclCl[+S +Pr -NonS]]<br><i>...(that) he bought...</i><br><i>...(who) he gave the book to...</i> |
| E. i. | Relative<br>Dependent        | [±RelDepIntroducer +DeclCl]<br><i>...(that) he went...</i>                                                  |
| E. j. | Extra<br>Dependent           | [+XDepIntroducer +DeclCl]<br><i>...when he went...</i>                                                      |

## 1.2. BURMESE

There are 75 Clause Units in Burmese. They are cast in a two-dimensional field: the Clause Type Dimension, and the Clause Class Dimension. The Clause Type dimension is subdivided, under four levels of consideration, into seven Clause Types. The Clause Class dimension is subdivided, under three levels of consideration, into twelve Clause Classes. The total field contains 84 possible Clause Units, but only 75 of them are grammatical and acceptable in Burmese.

SUMMARY TABLE II  
THE 75 CLAUSE UNITS IN BURMESE

| I                      | II                        | III         | No. | 1            | 2                   | 3          | 4           | 5          | 6         | 7       |
|------------------------|---------------------------|-------------|-----|--------------|---------------------|------------|-------------|------------|-----------|---------|
| Independent            | Non-Interrogative         | Declarative | a   | +            | +                   | +          | +           | +          | +         | +       |
|                        |                           | Imperative  | b   | +            | +                   | +          | +           |            |           |         |
|                        | Interrogative             | Yes-No      | c   | +            | +                   | +          | +           | +          | +         | +       |
|                        |                           | Echo        | d   | +            | +                   | +          | +           | +          | +         | +       |
|                        |                           | Subject     | e   | +            | +                   | +          | +           | +          | +         |         |
|                        |                           | Non-Subject | f   |              |                     |            |             |            |           | +       |
|                        |                           | Extra       | g   | +            | +                   | +          | +           | +          | +         | +       |
|                        |                           | Subject     | h   | +            | +                   | +          | +           | +          | +         |         |
| Dependent              | Nominalised               | Non-Subject | i   |              | +                   | +          | +           |            | +         |         |
|                        |                           | Relative    | j   | +            | +                   | +          | +           | +          | +         |         |
|                        |                           | Extra       | k   | +            | +                   | +          | +           | +          | +         |         |
|                        |                           | Topical     | l   | +            | +                   | +          | +           | +          | +         |         |
|                        | The twelve Clause Classes |             |     | No.          | 1                   | 2          | 3           | 4          | 5         | 6       |
| The seven Clause Types |                           |             | IV  | Intransitive | Single              | Double     | Attributive | Equational | Adjective | Stative |
|                        |                           |             |     |              | III                 | Transitive |             |            |           |         |
|                        |                           |             | II  | Active       |                     | Non-Act.   | Stative     |            |           |         |
|                        |                           |             |     | I            | Grammatical Subject |            |             | S-less     |           |         |

### 1.2.1. Clause Types in Burmese

The minimal formulas of the seven Independent Declarative Clause Units in Burmese are as follows:

#### B.a.1. Intransitive [+S +Pr]

thu thwa:de                                    he go-actual  
He went.

/-te/ is an obligatory particle indicating a statement of fact. It is one of a group of particles which mark a verb phrase as independent.

#### B.a.2. Transitive [+S ±O +Pr]

thu sa-qou' we-de                            he book buy-actual  
He bought a book.

B.a3. Double Transitive [ $\pm S$   $\pm IO$  O +Pr]  
 $\uparrow$        $\uparrow$

thu.gou sa-qou' pei:de he-goal book give-actual  
 Someone gave him a book.

The particle /-kou/ marks 'object, goal (time, place, person, etc.)'. It appears optionally with DO, and certain expressions of time to come (i.e. /nau'kou/ 'later on'. It is normally, though not always, present with expressions of place to which, and with IO if a DO is also present - it is entirely optional if the DO is not present. The order /...sa-qou' thu.gou.../ is also possible.

B.a4. Attributive Transitive [ $\pm S$   $\pm G$  +At Compl +Pr]

thu.gou nain-gan-qou' ywei: kau'te  
 he-goal nation-leader elect-actual  
 (They) elected him president.

B.a5. Equational [ $\pm S$  +EqCompl  $\pm EqPr$ ]

thu caun:dha:be: he student-emphatic  
 He is a student.

/-hpe:/ is a particle marking emphasis. In a short equation like this, the Burman feels the sentence to be incomplete without some such particle present. In longer equations such "fillers" are optional.

The negative of non-verbal equations requires the verb /hou'/'to be so', i.e. /thu caun:dha: mahou'hpu:/ (he student not-so) 'he isn't a student'. /ma=...hpu:/ is the standard negative statement particle combination.

B.a6. Adjective [ $\pm S$  +AdjPr]

thu kaun:de he good-actual  
 He is good (i.e. a good person).

B.a7. Stative [+StCompl +hyi.Pr]

caun:dha hnayau' hyi.de student two-individual have-actual  
 There are two students.

### 1.2.2. Clause Classes in Burmese

The minimal nuclear formulas of the eleven derived Clause Classes in Burmese are as follows:

## B.b. Imperative [+ImpPr ±PoliteFormula]

thwa: *Go!*  
 thwa:ba *Please go!*  
 thwa:ba-qoun: *Please go then!*

/thwa:/ 'Go!', i.e. the base form of the verb, is imperative. The bald verb, however, is impolite. In most situations it is softened with /-pa/ 'politeness, respect' and often still further with /-qoun:/ 'further, yet'.

## B.c. Yes-No Interrogative [±S +InterPr]

hkin-bya: thwa:dhala: *you go-actual-question*  
*Are you going?*

The particle /-la:/ is attached to statements to make yes-no questions. /tha/ is the allomorph of /te/ before /la:/.

## B.d. Echo Interrogative [+DeclCl +EchoInter]

hkin-bya: thwa:de, mahou'hpu:la:  
*you go-actual, not-so-question*  
*You are going, aren't you?*

In contrast to the other Southeast Asian languages in consideration, such as in Thai /châj máj/, which is positive, Burmese here uses the negative like English.

## B.e. Subject Interrogative [+InterS +Pr]

badhu thwa:dhale: *who goes-actual-question*  
*Who goes?*

/le:/ is the allomorph of /-la:/ 'question' which occurs in sentences with the interrogative nouns /ba/ 'what', /be/ 'which' and their derivatives.

## B.f. Non-Subject Interrogative [±S ±O +InterG +InterPr]

thu ba we-dhale: *he what buy-actual-question*  
*What does he buy?*  
 thu sa-qou' badhu.gou pei:dhale:  
*he book who-goal give-actual-question*  
*Who did he give the book to?*

The order /...badhu.gou sa-qou'.../ is also possible.

## B.g. Extra Interrogative [±S +XInter[Subordinate Clause] +InterPr]

thu be-gou thwa:dhale: *he where-goal go-actual-question*  
*Where did he go?*



thu ba hpyi'lou. thwa:dhale:

*he what happen-result go-actual-question*

*Why did he go?*

'why' is expressed in Burmese with a subordinate clause - here embedded in the sentence /thu thwa:de/ 'he went' - /ba hpyi'lou./ literally 'as a result of what happening'. Note that the presence of /ba/ even in a subordinate clause calls for the /-le:/ allomorph of the question particle after the main verb.

B.h. Subject Dependent [+NominalisedDepCl[+InterS...]]

badhu thwa:hman: thi.de *who go-matter know-actual*

*I know who went.*

badhu thwa:de hsou-da thi.de

*who go-actual speak-actual nominaliser know-actual*

*I know who went.*

/-ta/ or /-hta/ is a verb nominaliser (corresponding to the actuality particle /-te/ which forms action nouns from verbs. The first sentence thus means something like 'I know whose going'. In both examples, the object of /thi./ 'know' must be a noun of some sort.

Note also that the first clause in the second example occurs in statement form despite the presence of the question noun /ba/ 'what'.

B.i. Non-Subject Dependent [+NominalisedDepCl[+InterNonS]]

thu ba we-de hsou-da thi.de

*he what buy-actual speak-actual=nominaliser know actual*

*I know what he bought.*

thu ba we-hman: thi.de *he what buy-matter know-actual*

*I know what he bought.*

B.j. Relative Dependent [+NominalisedDepCl]

thu thwa:da thi.de

*he go-actual=nominaliser know-actual*

*I know he went.*

thu thwa:de hsou-da thi.de

*he go-actual speak-actual=nominaliser know-actual*

*I know he went.*

In the second example, the clause /thu thwa:de/ 'he went' is complement of the verb /hsou/ 'speak'; this whole construction is

then nominalised with /-ta/ and acts as object of /thi./ 'know'.  
The whole means something more like 'Speaking of his going, I know (it)'.

B.k. Extra Dependent [+NominalisedExtraDepCl]

thu be-gou thwa:de hsou-da thi.de  
*he where-goal go-actual speak-actual=  
 nominaliser know-actual*  
*I know where he went.*

thu be-gou thwa:hman: thi.de  
*he where-goal go-matter know-actual*  
*I know he went.*

B.1. Topical Dependent [+TopicalisedDepCl]

thu thwa:yin, kaun:me *he go-if, good-potential*  
*It would be good if he went.*

where the subordinate clause - lacking a particle of the class of /-te/, but marked with a subordinating particle /-yin/ 'if' - precedes the main clause which ends with the particle /-me/ 'possibility, potential'; /-me/ is the same class of particle as /-te/.

### 1.3. CAMBODIAN

There are 74 Clause Units in Cambodian. They are cast in a two-dimensional field: the Clause Class Dimension, and the Clause Type Dimension. The Clause Type Dimension is subdivided, under four levels of consideration, into eight Clause Types. The Clause Class Dimension is subdivided, under three levels of consideration, into eleven Clause Classes. The total field contains 88 possible Clause Units, but only 74 of them are grammatical and acceptable in Cambodian.

#### 1.3.1. Clause Types in Cambodian

The minimal formulas of the eight Independent Declarative Clause Units in Cambodian are as follows:

C.a1. Intransitive [+S +Pr]  
 koət təśw *He went.*

C.a2. Transitive [+S +Pr ±O]  
 koət tŷn s'iwphəśw (múuəj) *He bought a book. múuəj 'one'*

- C.a3. Double Transitive [ $\pm S$  +Pr  $\pm O$  +IO]  
 koət jɔk s'iəwphéəw qaəoj nfiəŋ *he take book give her*  
 koət cuun s'iəphéəw mək nfiəŋ *he offer book come her*  
*(He gave her a book.)*
- C.a4. Attributive Transitive [ $\pm S$  +Pr  $\pm O$  +AtCompl]  
 kéə réəh koət cfiə prathfiən cfiə 'to be'  
*(They elected him chairman.)*
- C.a5. Submissive [ $\pm S$  +SubmPr +SubCompl]  
 koət ttúuəl tuk kéə wfiəj koət *He was beaten by them.*  
 ttúuəl 'receive'  
 tuk 'pain'
- C.a6. Equational [ $\pm S$  +EqPr +EqCompl]  
 koət cfiə kòon səh *He is a student.*
- C.a7. Adjective [ $\pm S$  +AdjPr]  
 koət cfaat *He is intelligent.*
- C.a8. Stative [ $+mfiənPr$  +StCompl]  
 mfiən kòon səh piir neək *There were two people.*  
 neək 'person' = classifier

### 1.3.2. Clause Classes in Cambodian

The minimal nuclear formulas of the ten derived Clause Classes in Cambodian are as follows:

- C.b. Imperative [ $\pm Polite Formula$  +Pr]  
 kham təəw *Please come.*  
 kham 'try'
- C.c. Alternative Interrogative [+DeclCl +AlternInterPhrase]  
 qaəəŋ tśəw rýy nśəw *Are you going or staying?*
- C.d. Yes-No Interrogative [+DeclCl +YesNoInterPhrase]  
 qaəəŋ tśəw (rýy) tée *Are you going?*  
 qaəəŋ tśəw rýy rýy 'or'  
 tée 'no, not'
- C.e. Subject Interrogative [+InterS +Pr]  
 neək naa tśəw *Who goes?*

SUMMARY TABLE III  
THE 76 CLAUSE UNITS IN CAMBODIAN

| I                         |                       | II          |  | III |   | No. | 1            | 2      | 3 | 4 | 5 | 6 | 7 | 8 |   |
|---------------------------|-----------------------|-------------|--|-----|---|-----|--------------|--------|---|---|---|---|---|---|---|
| Independent               | Non-Interrogative     | Declarative |  | a   | + | +   | +            | +      | + | + | + | + | + | + |   |
|                           |                       | Imperative  |  | b   | + | +   | +            | +      |   |   |   |   |   |   |   |
|                           | Interrogative         | Alternative |  | c   | + | +   | +            | +      | + | + | + | + | + | + | + |
|                           |                       | Yes-No      |  | d   | + | +   | +            | +      | + | + | + | + | + | + | + |
|                           |                       | Subject     |  | e   | + | +   | +            | +      | + | + | + | + | + | + | + |
|                           |                       | Non-Subject |  | f   |   | +   | +            | +      | + | + |   |   |   |   | + |
|                           |                       | Extra       |  | g   | + | +   | +            | +      | + | + | + | + | + | + | + |
| Dependent                 | Without Extra Tagmeme | Subject     |  | h   | + | +   | +            | +      | + | + | + | + | + |   |   |
|                           |                       | Non-Subject |  | i   |   | +   | +            | +      | + | + | + | + | + | + |   |
|                           | With Extra Tagmeme    | Relative    |  | j   | + | +   | +            | +      | + | + | + | + | + | + |   |
|                           |                       | Extra       |  | k   | + | +   | +            | +      | + | + | + | + | + | + |   |
| The eleven Clause Classes |                       |             |  |     |   | No. | 1            | 2      | 3 | 4 | 5 | 6 | 7 | 8 |   |
|                           |                       |             |  |     |   | IV  | Intransitive | Single |   |   |   |   |   |   |   |
| Double                    |                       |             |  |     |   |     |              |        |   |   |   |   |   |   |   |
| III                       | Transitive            | Attributive |  |     |   |     |              |        |   |   |   |   |   |   |   |
|                           |                       | Submissive  |  |     |   |     |              |        |   |   |   |   |   |   |   |
| II                        | Active                | Equational  |  |     |   |     |              |        |   |   |   |   |   |   |   |
|                           |                       | Adjective   |  |     |   |     |              |        |   |   |   |   |   |   |   |
| I                         | Grammatical Subject   | Non-Active  |  |     |   |     |              |        |   |   |   |   |   |   |   |
|                           |                       | Stative     |  |     |   |     |              |        |   |   |   |   |   |   |   |
| The eight Clause Types    |                       |             |  |     |   |     |              |        |   |   |   |   |   |   |   |
|                           |                       |             |  |     |   |     |              |        |   |   |   |   |   |   |   |

C.f. Non-Subject Interrogative [ $\pm S$  +Pr +NonSInter]

koət tʏn naa (qəəjwan naa) *What did he buy?*  
qəəjwan 'wares, goods'

C.g. Extra Interrogative [ $\pm S$  +Pr +XInter]

koət tʃəw (tii) naa *Where did he go?*  
tii 'place'

C.h. Subject Dependent [+InterS +Pr]

(krom dʏŋ) neək naa tʃəw haəj *I know what person went.*  
(krom dʏŋ) neək daeɪ tʃəw haəj *I know the person who went.*

C.i. Non-Subject Dependent [ $\pm S$  +Pr +InterNonS]

(krom dʏŋ) koət tʏn naa *I know what he bought.*

C.j. Relative Dependent [+RelDepIntroducer +DeclCl]

(krom dʏŋ) thaa koət tʃəw haəj *I know that he went.*

## C.k. Extra Dependent [+S +Pr +XDepIntroducer]

(krom dŷŋ) koət tʰəw (tii) naa I know where he went.

(krom dŷŋ) tii naa koət tʰəw

## 1.4. CANTONESE

There are 81 Clause Units in Cantonese. They are cast in a two-dimensional field: the Clause Type Dimension, and the Clause Class Dimension. The Clause Type Dimension is subdivided, under four levels of consideration, into eight Clause Types. The Clause Class Dimension is subdivided, under three levels of consideration, into twelve Clause Classes. The total field contains 96 possible Clause Units, but only 81 of them are grammatical and acceptable in Cantonese.

## 1.4.1. Clause Types in Cantonese

The minimal formulas of the eight Independent Declarative Clause Units in Cantonese are as follows:

- Ct.a1. Intransitive [+S +Pr]  
 keúih heui He went.
- Ct.a2. Transitive [+S +Pr ±O]  
 keúih maai syù He bought a book.
- Ct.a3. Double Transitive [+S +Pr +IO ±O]  
 keúih béi keúih syù He gave her a book.  
 keúih béi syù keúih
- Ct.a4. Attributive Transitive [+S +Pr ±O +AtCompl]  
 keúih-deih syún keuih Jouh jyú-jihk They elected him chairman.  
 Jouh 'to be'
- Ct.a5. Submissive [+S +SubmPr +SubmCompl]  
 keúih béi keúih-deih dá He was beaten by them.  
 béi 'undergo a bad experience'
- Ct.a6. Equational [+S +EqPr +EqCompl]  
 keúih haih hohk-saàng He is a student.
- Ct.a7. Adjective [+S +AdjPr]  
 keúih chùng-míng He is intelligent.

Ct.a8. Stative [+yáuhPr +StCompl]  
 yáuh léuhng-go hohk-saàng *There are two students.*

SUMMARY TABLE IV  
 THE 81 CLAUSE UNITS IN CANTONESE

| I                         | II                                 | III                         | No. | 1                   | 2      | 3      | 4           | 5          | 6          | 7         | 8       |
|---------------------------|------------------------------------|-----------------------------|-----|---------------------|--------|--------|-------------|------------|------------|-----------|---------|
| Independent               | Non-Interrogative                  | Declarative                 | a   | +                   | +      | +      | +           | +          | +          | +         | +       |
|                           |                                    | Imperative                  | b   | +                   | +      | +      | +           |            |            |           |         |
|                           | Interrogative                      | Alternative                 | c   | +                   | +      | +      | +           | +          | +          | +         | +       |
|                           |                                    | Yes-No                      | d   | +                   | +      | +      | +           | +          | +          | +         | +       |
|                           |                                    | Realisation-Non-Realisation | e   | +                   | +      | +      | +           | +          | +          | +         | +       |
|                           |                                    | Subject                     | f   | +                   | +      | +      | +           | +          | +          | +         |         |
|                           |                                    | Non-Subject                 | g   |                     | +      | +      | +           |            |            |           |         |
|                           |                                    | Extra                       | h   | +                   | +      | +      | +           | +          | +          | +         | +       |
| Dependent                 | Without an Extra Dependent tagmeme | Subject                     | i   | +                   | +      | +      | +           | +          | +          | +         |         |
|                           |                                    | Non-Subject                 | j   |                     | +      | +      | +           |            |            |           |         |
|                           | With an Extra Dependent tagmeme    | Relative                    | k   | +                   | +      | +      | +           | +          | +          | +         | +       |
|                           |                                    | Extra                       | l   | +                   | +      | +      | +           | +          | +          | +         | +       |
| The twelve Clause Classes |                                    |                             | No. | 1                   | 2      | 3      | 4           | 5          | 6          | 7         | 8       |
|                           |                                    |                             | IV  | Intransitive        | Single | Double | Attributive | Submissive | Equational | Adjective | Stative |
| The eight Clause Types    |                                    |                             | III | Transitive          |        |        | Submissive  | Equational | Adjective  | Stative   |         |
|                           |                                    |                             | II  | Active              |        |        | Non-Active  |            |            | Stative   |         |
|                           |                                    |                             | I   | Grammatical Subject |        |        |             |            |            |           | S-less  |
|                           |                                    |                             |     |                     |        |        |             |            |            |           |         |

#### 1.4.2. Clause Classes in Cantonese

The minimal nuclear formulas of the eleven derived Clause Classes in Cantonese are as follows:

Ct.b. Imperative [+Pr +ImpPhrase]  
 heui là *Go!*

Ct.c. Alternative Interrogative [+DeclCl +AlterPhrase +DeclCl]  
 néi heui yíkwaahk m̄heui a *Are you going or staying?*  
 yíkwaahk 'or'  
 m̄ 'not'

- Ct.d. Yes-No Interrogative [+DeclCl +YesNoInter]  
 néi heui ñ heui a *Are you going?*  
 néi heui ma
- Ct.e. Realisation-Nonrealisation [+DecCl +RnonRInter]  
 Interrogative  
 néi heui-jó mei a *Did you go?*  
 mei 'not, not yet'
- Ct.f. Subject Interrogative [+InterS +Pr]  
 bīngo heui a *Who goes?*
- Ct.g. Non-Subject Interrogative [±S +Pr +NonSInter]  
 keufh maáí-jó māt-yéh a *What did he buy?*
- Ct.h. Extra Interrogative [±S +Pr +XInter]  
 keufh heui-jó bīnsyu a *Where did he go?*
- Ct.i. Subject Dependent [+DeclCl]  
 (ngó jǐ) bīngo heui-jó *I know who went.*
- Ct.j. Non-Subject Dependent [±S +Pr +NonSDep]  
 (ngó jǐ) keufh maáí-jó māt-yéh *I know what he bought.*
- Ct.k. Relative Dependent [+DeclCl]  
 (ngó jǐ) keufh heui-jó *I know he went.*
- Ct.l. Extra Dependent [+S +Pr +XDep]  
 (ngó jǐ) keufh heui-jó bīn-syu *I know where he went.*

### 1.5. LAO AND THAI

Lao and Thai are closely related and have the same Clause Units. There are 92 Clause Units in either Lao or Thai. The Clause Units are cast in a two-dimensional field: the Clause Type Dimension, and the Clause Class Dimension. The Clause Type Dimension is subdivided, under four levels of consideration, into eight Clause Types. The Clause Class Dimension is subdivided, under three levels of consideration, into thirteen Clause Classes. The total field contains 104 possible Clause Units, but only 92 are grammatical and acceptable in either Lao or Thai.

## 1.5.1. Clause Types in Lao and Thai

The minimal formulas of the eight Independent Declarative Clause Units in Lao and Thai are as follows, with the first examples in Lao, and the second examples in Thai:

- LT.a1. Intransitive [±S +Pr]  
 la'aw paj *He went.*  
 khǎw paj
- LT.a2. Transitive [±S +Pr ±O]  
 la'aw (dǎj) sy̌y p̌ym *He bought a book.*  
 khǎw šy̌y nǎ̌nsy̌y
- LT.a3. Double Transitive [±S +Pr ±O +IO]  
 la'aw aw p̌ym ȟaj la'aw *He gave her a book.*  
 khǎw aw nǎ̌nsy̌y ȟaj khǎw
- LT.a4. Attributive Transitive [±S +Pr ±O +AtCompl]  
 khacàw (dǎj) lyàk la'aw (pen) ȟua-nà̌a pasúm *They elected him*  
 khǎw lyàk khǎw pen ȟua-nà̌a *chairman.*
- LT.a5. Submissive [±S +SubmPr +SubmCompl]  
 la'aw thỳk khacàw tii *He was beaten by them.*  
 khǎw thuuk khǎw tii
- LT.a6. Equational [±S +EqPr +EqCompl]  
 la'aw pen nāk-hián *He is a student.*  
 khǎw pen nāk-rian
- LT.a7. Adjective [±S +AdjPr]  
 la'aw kē̌j *He is intelligent.*  
 khǎw kē̌j
- LT.a8. Stative [+miiPr +StCompl]  
 mī' nāk-hián sǒǰj khón *There are two students.*  
 mii nāk-rian sǒǰj khon

## 1.5.2. Clause Classes in Lao and Thai

The minimal nuclear formulas of the twelve derived Clause Classes in Lao and Thai are as follows:



- LT.b. Imperative [+ImpPr...]  
 paj Go!  
 paj
- LT.c. Alternative Interrogative [+DeclCl +AlterPhrase +DeclCl]  
 càw si paj lǎy si jūu Are you going or staying?  
 khun cà paj rǎy (phák)
- LT.d. Yes-No Interrogative [+DeclCl +YesNoInter]  
 càw si paj bɔɔ Are you going?  
 khun cà paj máj
- LT.e. Right-Wrong Interrogative [+DeclCl +RWInter]  
 càw si paj mēn bɔɔ You are going, aren't you?  
 khun cà paj cháj máj
- LT.f. Realisation-Nonrealisation [DeclCl +RnonRInter]  
 Interrogative  
 càw (dàj) paj lɛɛw lǎy nɔn Have you gone yet?  
 khun (dàj) paj rǎy jaŋ
- LT.g. Subject Interrogative [+InterS +Pr]  
 phǎj paj (mēn phǎj paj) Who went?  
 khraj paj
- LT.h. Non-Subject Interrogative [±S +Pr +NonSInter]  
 laáw (dàj) syý nɔn What did he buy?  
 khaw (dàj) syý ?araj
- LT.i. Extra Interrogative [±S +Pr +XInter]  
 laáw (dàj) paj sǎj Where did he go?  
 khǎw (daj) paj nǎj
- LT.j. Subject Dependent [+DepIntroducer +DepS +Pr]  
 (khôj) huù wā mēn phǎj paj I know who went.  
 (rúu) wā khraj paj
- LT.k. Non-Subject Dependent [+DepIntroducer ±S +Pr +NonSDep]  
 (khôj) huù wā laáw (dàj) syý nǎn I know what he bought.  
 (rúu) wā khǎw syý ?araj

## LT.1. Relative Dependent [+RelDepIntroducer +DeclCl]

(khôj huù) wāa laáw (dǎj) paj *I know that he went.*  
 (ruú) wāa khǎw paj

## LT.m. Extra Dependent [+DepIntroducer ±S +Pr +XDep]

(khôj huù) wāa laáw (dǎj) paj sǎj *I know where he went.*  
 (ruú) wāa khǎw paj thiî nǎj

SUMMARY TABLE V  
 THE 92 CLAUSE UNITS IN LAO OR THAI

| I                           |                    | II                          |   | III |   | No.                    | 1                   | 2          | 3 | 4 | 5 | 6 | 7 | 8 |   |  |  |  |  |
|-----------------------------|--------------------|-----------------------------|---|-----|---|------------------------|---------------------|------------|---|---|---|---|---|---|---|--|--|--|--|
| Independent                 | Non-Inter-rogative | Declarative                 |   | a   | + | +                      | +                   | +          | + | + | + | + | + | + |   |  |  |  |  |
|                             |                    | Imperative                  |   | b   | + | +                      | +                   | +          |   |   |   |   |   | + |   |  |  |  |  |
|                             | Interrogative      | Alternative                 |   | c   | + | +                      | +                   | +          | + | + | + | + | + | + | + |  |  |  |  |
|                             |                    | Yes-No                      |   | d   | + | +                      | +                   | +          | + | + | + | + | + | + | + |  |  |  |  |
|                             |                    | Right-Wrong                 |   | e   | + | +                      | +                   | +          | + | + | + | + | + | + | + |  |  |  |  |
|                             |                    | Realisation-Non-Realisation |   | f   | + | +                      | +                   | +          | + | + | + | + | + | + | + |  |  |  |  |
|                             |                    | Subject                     |   | g   | + | +                      | +                   | +          | + | + | + | + | + | + |   |  |  |  |  |
|                             |                    | Non-Subject                 |   | h   |   | +                      | +                   | +          |   |   |   |   |   |   | + |  |  |  |  |
|                             |                    | Extra                       |   | i   | + | +                      | +                   | +          | + | + | + | + | + | + | + |  |  |  |  |
|                             |                    | Subject                     |   | j   | + | +                      | +                   | +          | + | + | + | + | + | + | + |  |  |  |  |
| Dependent                   | Non-Subject        |                             | k | +   | + | +                      | +                   | +          | + | + | + | + | + |   |   |  |  |  |  |
|                             | Relative           |                             | l | +   | + | +                      | +                   | +          | + | + | + | + | + |   |   |  |  |  |  |
|                             | Extra              |                             | m | +   | + | +                      | +                   | +          | + | + | + | + | + |   |   |  |  |  |  |
|                             |                    |                             |   |     |   |                        |                     |            |   |   |   |   |   |   |   |  |  |  |  |
| The thirteen Clause Classes |                    |                             |   |     |   | No.                    | 1                   | 2          | 3 | 4 | 5 | 6 | 7 | 8 |   |  |  |  |  |
|                             |                    |                             |   |     |   | IV                     | Intransitive        | Single     |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   |                        |                     | Double     |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   |                        | Transitive          |            |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   | III                    | Submissive          | Equational |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   |                        |                     | Adjective  |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   |                        |                     | Stative    |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   | II                     | Active              | Non-Active |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   |                        |                     |            |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   | I                      | Grammatical Subject | S-less     |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   |                        |                     |            |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   | The eight Clause Types |                     |            |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   |                        |                     |            |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   |                        |                     |            |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   |                        |                     |            |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   |                        |                     |            |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   |                        |                     |            |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   |                        |                     |            |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   |                        |                     |            |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   |                        |                     |            |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   |                        |                     |            |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   |                        |                     |            |   |   |   |   |   |   |   |  |  |  |  |
|                             |                    |                             |   |     |   |                        |                     |            |   |   |   |   |   |   |   |  |  |  |  |

## 1.6. VIETNAMESE

There are 108 Clause Units in Vietnamese, which are cast in a two-dimensional field: the Clause Type Dimension, and the Clause Class Dimension. The Clause Type Dimension is subdivided, under four levels of consideration, into nine Clause Types. The Clause Class Dimension is

subdivided, under five levels of consideration, into thirteen Clause Classes. The total field contains 117 possible Clause Units, but only 108 of them are grammatical and acceptable in Vietnamese.

### 1.6.1. Clause Types in Vietnamese

The minimal formulas of the nine Independent Declarative Clause Units in Vietnamese are as follows:

|       |                        |                                      |
|-------|------------------------|--------------------------------------|
| V.a1. | Intransitive           | [±S +Pr]                             |
|       | Nó đi.                 | <i>He went.</i>                      |
| V.a2. | Transitive             | [±S +Pr ±O]                          |
|       | Nó mua sách.           | <i>He bought books.</i>              |
| V.a3. | Double Transitive      | [±S +Pr +IO ±O]                      |
|       | Nó cho họ sách.        | <i>He gave them books.</i>           |
| V.a4. | Attributive Transitive | [±S +Pr ±O +AtCompl]                 |
|       | Nó làm họ sợ.          | <i>He made them afraid.</i>          |
| V.a5. | Actional Submissive    | [±S +SubmPr +SubmActlCompl]          |
|       | Nó bị đi.              | <i>He had (was forced) to go.</i>    |
| V.a6. | Passive Submissive     | [±S +SubmPr +SubmPasCompl]           |
|       | Nó bị (họ) rầy.        | <i>He was reprimanded (by them).</i> |
| V.a7. | Equational             | [±S +EqPr +EqCompl]                  |
|       | Họ là sinh-viên.       | <i>They are students.</i>            |
| V.a8. | Adjective              | [±S +AdjPr]                          |
|       | Họ giàu.               | <i>They are rich.</i>                |
| V.a9. | Stative                | [+c <sup>o</sup> Pr +StCompl]        |
|       | Có hai người.          | <i>There were two people.</i>        |

### 1.6.2. Clause Classes in Vietnamese

The minimal nuclear formulas of the twelve derived Clause Classes are as follows:

|      |                           |                                  |
|------|---------------------------|----------------------------------|
| V.b. | Imperative                | [±S +ImpPr...]                   |
|      | Ông hãy đi.               | <i>Go! (Mr had better go.)</i>   |
| V.c. | Alternative Interrogative | [+DeclCl +AlterPhrase +DeclCl]   |
|      | Ông đi hay (ông) ở?       | <i>Are you going or staying?</i> |



|      |                           |                               |
|------|---------------------------|-------------------------------|
| V.g. | Subject Interrogative     | [+InterS +Pr...]              |
|      | Ai đi?                    | <i>Who went?</i>              |
| V.h. | Non-Subject Interrogative | [±S +Pr +InterNonS]           |
|      | Nó mua gì?                | <i>What did he buy?</i>       |
| V.i. | Extra Interrogative       | [±S +Pr... +XInter]           |
|      | Nó đi hồi nào?            | <i>When did he go?</i>        |
| V.j. | Subject Dependent         | [±DepS +Pr...]                |
|      | ...(mà) đi...             | <i>...that went...</i>        |
| V.k. | Non-Subject Dependent     | [+DepNonS ±S +Pr...]          |
|      | ...(mà) nó mua...         | <i>...(that) he bought...</i> |
| V.l. | Relative Dependent        | [+RelDepIntroducer +DeclCl]   |
|      | ...(rằng) nó đi...        | <i>...(that) he went...</i>   |
| V.m. | Extra Dependent           | [+XDepIntroducer +DeclCl]     |
|      | Khi nó đi...              | <i>When he went...</i>        |

## 2. CLAUSE UNITS IN CONTRAST

Since the Clause Units in the languages are cast in a two-dimensional field (the Clause Type Dimension, and the Clause Class Dimension), a contrastive analysis of the units can also be presented bi-dimensionally.

### 2.1. CLAUSE TYPES IN CONTRAST

On the Clause Type Dimension, the following differences between English and the Southeast Asian languages under consideration can be noticed from a comparison of the charts and formulas:

- (1) In English, the Subject, and the Object tagmemes are nuclear and obligatory. In the Southeast Asian languages, the same tagmemes are also nuclear but they are optional in the sense that they can be omitted when the context permits it.
- (2) Only English has passive transitive constructions such as E.a5, E.a6, and E.a7.
- (3) Except Burmese, all the other SEA languages have a submissive construction such as C.a6, Ct.a5, LT.a5. Vietnamese has two submissive constructions, the Actional Submissive Clause Type V.a5, and the Passive Submissive Clause Type V.a6.
- (4) In the SEA languages, the Adjective Clause Type does not require

a copula, something like 'be' in English. In Burmese, the Equational Clause Type B.a5 has only an optional equational predicate which is obligatory only when the clause type is in a negative form.

- (5) Although not apparent in the formulas, the SEA languages can have any tagmeme topicalised by being placed at the beginning of a clause, for example:  
Sách ày, tòi mua hôm qua. Lit.: *That book, I bought yesterday.*
- (6) Burmese is the only language that has the tagmeme ordering S O Pr as in B.a2, B.a3, and B.a4.
- (7) The 'it' Stative Clause Type E.a10 is unique in English.

## 2.2. CLAUSE CLASSES IN CONTRAST

On the Clause Class Dimension, the following dissimilarities can be noticed:

- (1) In English, the Yes-No Interrogative, Non-Subject Interrogative, and Extra Interrogative Clause Classes E.c, E.e, and E.f have their S Pr order reversed. In the SEA languages, the tagmeme ordering is always S Pr whether it is a declarative or interrogative clause class.
- (2) In the SEA languages except Burmese, the Dependent Clause Classes have the same form as the Interrogative Clause Classes in that the tagmeme ordering is always S Pr, whereas in English, the Dependent Clause Classes have the S Pr tagmeme ordering, and the Interrogative Clause Classes (except the Subject Interrogative Clause Class E.d) have the reversed ordering.
- (3) In Burmese, all Dependent Clause Classes are nominalised.

## 3. DEEP VERSUS SURFACE STRUCTURE

Since in foreign language learning, performance is what counts even more than competence, and since performance can be measured in terms of an individual's capability of making only grammatical utterances, the surface structure is as important as, if not more important, than the deep structure.

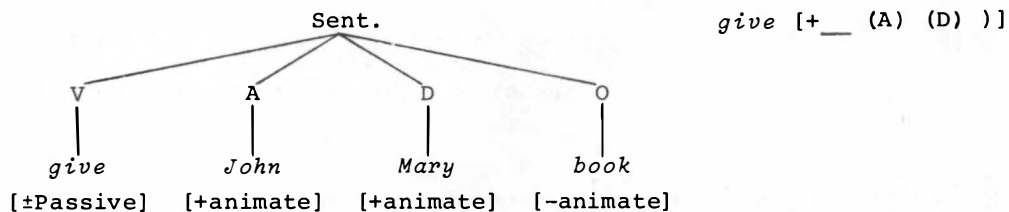
The system of tagmemic analysis as demonstrated above is particularly powerful in presenting the grammatical elements of surface structure in terms of the function and form of each unit being contrasted. The matrix displays presented above can show the surface relationships

between various clause units. Surface structure transformational rules can also be conceived to relate the Passive Clause Types E.a5, E.a6, and E.a7 to the Active Clause Types E.a2, E.a3, and E.a4 in English for example. They are as well capable of showing the differences between the Extra Interrogative and Non-Subject Interrogative Clause Classes in English and in the SEA languages in that in English, an Extra Interrogative Clause such as *'When did he go?'* or a Non-Subject Interrogative Clause such as *'What did he buy?'* can be conceived as going through two transformational stages, the first one from *'He went yesterday'* or *'He bought a book'* to become *'Did he go yesterday?'* or *'Did he buy a book?'*, and the second one from the two Yes-No Interrogative Clauses to the Extra Interrogative or Non-Subject Interrogative Clauses in question, and in that in the SEA languages the same Clauses would be derived directly from their kernel Declarative Clauses.

However, it is also the belief of the author that contrastive analysis should explore the deep structure in order to present the relationships between different surface structure patterns that have a common deep structure, and to distinguish similar surface structure patterns that have diverse deep structure cases. In order to show how contrastive analysis can be enriched by considerations on the deep structure, I shall use Case Grammar as developed by Charles Fillmore (1968, 1970a, and 1970b) to explain some of the differences found in 2 above.

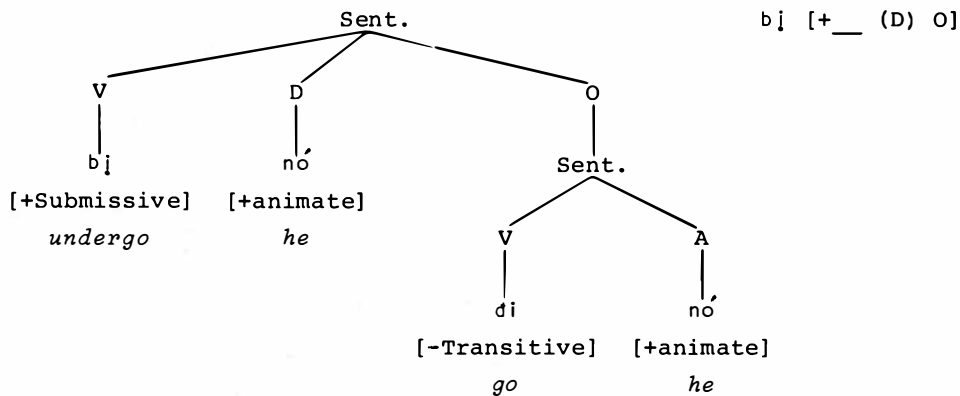
Case grammar is a system which views the deep structure of sentences as a set of relations between a verb and a series of case-marked noun phrases. The series of noun phrases constitutes a set of roles which are useful in classifying verbs in terms of the case frames in which they occur. The system is particularly powerful in relating sentences with identical deep structures but diverse surface structures, and in differentiating sentences with identical surface structures but diverse deep structures.

Let us take the verb *'give'* which appeared in E.a3 and E.a6, and see how contrastive analysis can benefit from Case Grammar. The deep structure of the two clauses exemplified in E.a3 and E.a6 can be represented as follows:

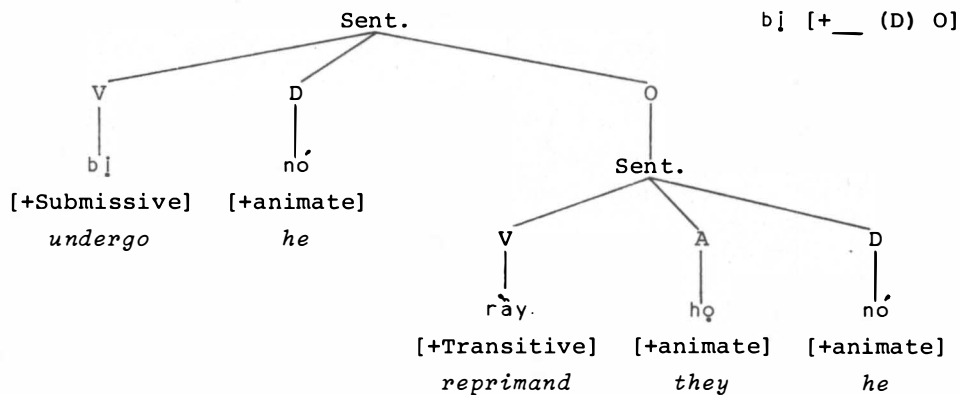


The verb '*give*' is marked with ±Passive. Thus, by applying the optional realisation rule of passivation, we can get the sentences in E.a6. Furthermore, because it is hospitable to both an Agentive case A and a Dative case D, we can have two passive sentences, one with the Dative case as surface Subject and one with the Objective case as surface Subject. Such an insight to the deep structure would be invaluable in transformational drill preparation.

In the previous example using '*give*' we could see how two diverse surface structures such as the active and the passive forms could be related, let us see how two similar surface structures can be differentiated in the two Submissive Actional and Submissive Passive Clauses V.a5 and V.a6 in Vietnamese:



V.a5. Nó bị đi. 'He had to go.'



V.a6. Nó bị (họ) rày. 'He was reprimanded (by them).'



Such a differentiation of similar surface structures that have diverse case frames will help the textbook writer eliminate the construction of drills that would form such sentences like '*John is eager to please*' and '*John is easy to please*' on the same basis, and would give him insights to the semantic differences between these sentences.

#### 4. PHILOSOPHY OF APPROACH

Contrastive analysis points out the similarities and differences on the surface structure of languages. Thus, it plays an important role in area linguistics. Furthermore, since in language learning the surface structure is as important as the deep structure, contrastive analysis should not be rejected because one belongs to a linguistic camp. On the contrary, if one is prepared to make use of existing theories and techniques in an eclectic way, one will be able to make contrastive analysis more powerful and make use of it in foreign language teaching until another cognitive hypothesis can offer a more efficient language acquisition technique than drill and overlearning.

Contrastive Analysis can be assisted by error analysis and statistical studies of the frequency of occurrence of grammatical constructions. Error analysis, such as the one done for English-Vietnamese contrastive studies by the present author (Liem 1970a) will enable the applied linguist to set up a hierarchy of difficulty upon which he will be able to develop a philosophy of approach to teaching and learning problems. Finally, since the aim of foreign language learning is communication, the relative frequency of occurrence of grammatical as advocated by the present author for English (Liem 1970b) and Vietnamese (1970c) should serve to establish which grammatical constructions should receive priority in a language textbook.

## REFERENCES

CHOMSKY, Noam

- 1957 *Syntactic Structures*. Mouton and Co., The Hague.  
1965 *Aspects of a Theory of Syntax*. The M.I.T. Press, Cambridge.

COOK, Walter A.

- 1969 *Introduction to Tagmemic Analysis*. Holt, Rinehart and Winston, Inc., New York.  
1971 "Case Grammar as a Deep Structure in Tagmemic Analysis". *Languages and Linguistics: Working Papers No.2:1-9*. Georgetown School of Languages and Linguistics, Washington D.C.

FILLMORE, Charles J.

- 1968 "The Case for Case". *Universals in Linguistic Theory*. Emmon Bach and Robert T. Harms (eds). 1-88. Holt, Rinehart and Winston, Inc., New York.  
1970a "Subjects, Speakers and Roles". *The Ohio State University Working Papers in Linguistics*, No.4:31-63.  
1970b "On Generativity". *The Ohio State University Working Papers in Linguistics*, No.6:2-19.

FRIES, Charles C.

- 1945 *Teaching and Learning English as a Foreign Language*. University of Michigan Press, Ann Arbor.

LADO, Robert

- 1957 *Linguistics across Cultures*. University of Michigan Press, Ann Arbor.

## LIEM, Nguyen Dang

- 1966 *English Grammar, a Combined Tagmemic and Transformational Approach.* The Australian National University (ANU), Canberra.
- 1967 *A Contrastive Grammatical Analysis of English and Vietnamese.* ANU, Canberra.
- 1969 *Vietnamese Grammar, a Combined Tagmemic and Transformational Approach.* ANU, Canberra.
- 1970a *A Contrastive Phonological Analysis of English and Vietnamese.* ANU, Canberra.
- 1970b "Clause Units in Vietnamese". *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung.* Band 23, Heft 1:20-30.
- 1970c "Clause Units in English". *Pacific Linguistic Studies in Honour of Arthur Capell.* 97-108. ANU, Canberra.

## LONGACRE, Robert E.

- 1964 *Grammar Discovery Procedures.* Mouton and Co., The Hague.

## PIKE, Kenneth L.

- 1954, *Language in Relation to a Unified Theory of the Structure of*  
 1955, *Human Behavior.* Parts I, II, and III. Summer Institute of  
 1960 Linguistics, Glendale.

## RAY, Punya Sloka

- 1967 *Linguistic Matrices.* Indian Institute of Advanced Studies, Simla.

## WEINREICH, Uriel

- 1953 *Language in Contact.* Linguistic Circle of New York, New York.

## YOUNG, Richard E., Alton L. Becker, and Kenneth L. Pike

- 1970 *Rhetoric: Discovery and Change.* Harcourt, Brace and World, Inc., New York.



## THE VALUE OF AU AND AI IN MIDDLE KHMER

PHILIP N. JENNER

The years from the abandonment of Angkor in 1431 down through the eighteenth century were a time of radical phonological, morphological and syntactic changes which formed the transition from late Old Khmer to early modern Khmer. The phonological changes in question affected the consonant system and the vowel system alike. The chief development in the consonantism was the devoicing of the old voiced stops while that in the vocalism was the generation, out of the original inventory, of two parallel subsets of syllable nuclei now generally known as the "registers". Describing the various changes which took place and fixing their order in time is an enterprise which no one has yet attempted, presumably because of the absence of a reliable method or of evidence recognisable as such. While the historical development of Khmer is still only very imperfectly understood, it now appears that at least some of the changes that occurred, together with their sequence, are discoverable from Middle Khmer metrical texts.

The purpose of the present paper is to suggest that the graphemes au and ai had the value in pre-modern Khmer (before the development of the registers) of /əw/ [ɿw] and /əj/ [ɿj], respectively, rather than the commonly assumed /aw/ [aw] and /aj/ [aj]. Because this claim is likely to run counter to the views of most Cambodian and other scholars who consider the matter, these two symbols provide a particularly welcome opportunity for me to describe the evidence and the method of analysis on which my conclusion is based.

As a preface to what follows it may be mentioned that the modern view of au and ai as representing original /aw/ and /aj/ apparently arises from the precedence assigned to the Low Register (LR) generally<sup>1</sup> and, in

particular, from the circumstance that the "natural" register<sup>2</sup> of syllable nuclei following initial /q/ [ʔ ~ Ø] is low. Initial 'ī and 'ū, for example, are read as LR /qəəj/ and /qəow/. The corresponding High Register (HR) readings, being "unnatural", require use of the *trāsabda*, a diacritic functioning as the diacresis: 'ī̄ /qii/, 'ū̄ /quu/. Similarly, the natural readings of 'ē and 'o are LR /qaaɛ/ and /qaaɔ/, not HR /qee/ and /qoo/. The HR and LR readings of au and ai are, respectively, /əw ~ əj/ and /aw ~ aj/. The thesis here, therefore, is that the Middle Khmer value of au and ai was virtually the same as that of the modern HR.

The question of what values of au and ai were introduced into Cambodia from India and Ceylon, while by no means irrelevant, is unanswerable in the present state of our knowledge. The modern writing system is known to have developed from a prototype attributable to the early *grantha* systems in use under the Pallava kings of Kañcī (Coromandel coast) and brought into Fu-nan by the second century A.D. However, the problem is decidedly complex. Among other things, we have only inferential knowledge regarding the phonology of Sanskrit as it was current in Cambodian territory. For the time being, we know nothing of the constancy of brahmanic literary control over this Cambodian Sanskrit. We can only guess at the conditions surrounding the adaptation of the writing system, previously used only for Sanskrit, to Khmer.<sup>3</sup> Knowing next to nothing of the phonology of the Pāli introduced into Cambodia around the fourteenth century,<sup>4</sup> we cannot even surmise whether this confirmed or conflicted with the Sanskrit tradition. We are in no position to assess possible Dravidian influence in the establishment of written Khmer. The parallels with Sino-Japanese are striking.

The data base from which the present analysis is made consists of 5164 rhymes from eighteen Middle Khmer metrical texts. Seventeen of the latter belong to the *cpā'pa* /cbap/ genre, comprising short homiletic works, while the eighteenth is the *Grande inscription d'Angkor Vat* (IMA 38).<sup>5</sup> The *cpā'pa* are undated texts but have been placed in chronological sequence on the basis of the phonological divergence from modern Khmer revealed in their rhymes; IMA 38, a longer devotional text, is dated in correspondence with A.D. 1702 or 1701, and the same chronological analysis places it immediately after the five earliest *cpā'pa*. Most of the relevant information on Middle Khmer au and ai is shown on the accompanying tables.

Table 1 shows the number of au rhymes and ai rhymes<sup>6</sup> in relation to the total number of rhymes in each text. As can be seen, there are 153 au rhymes and 154 ai rhymes, representing 2.96% and 2.98% of the total respectively. In addition to indicating the relatively small segment of

the corpus with which we are concerned here, the table shows the absence of any plausible correlation of the frequency of rhymes in each text with the age of the texts in which they occur. While it is undeniable that au rhymes are lacking in the *Cpā'pa krama*, one of the earliest works, and reach peak frequencies in the later *Kūna cau lpæka* (B) (7.09%), the *Ariyasatthā* (6.70%), the *Bākya cā'sa* (6.19%) and the *Dūnmāna kūna* (5.98%), the overall frequency profile (which averages 3.32%) is so erratic as to preclude any connection with the date of the texts. Such a view is confirmed by the pattern of ai rhymes: these have an average frequency of 2.71% and shows peaks descending from the late *Cpā'pa hai mahājana* (II) (6.17%) to the earlier IMA 38 (5.88%) and the even earlier *Kūna cau* (5.49%). At least for this group of texts, therefore, it may be supposed that the incidence of au rhymes and ai rhymes is fortuitous and dependent only on the needs and talents of the successive poets.

|                           | TOTAL RHYMES | au RHYMES |       | ai RHYMES |       |
|---------------------------|--------------|-----------|-------|-----------|-------|
| <i>Ker(tti) kāla</i>      | 105          | 1         | 0.95% | 1         | 2.86% |
| <i>Kūna cau</i>           | 328          | 15        | 4.57% | 18        | 5.49% |
| <i>Rājaneti</i>           | 214          | 10        | 4.67% | 4         | 1.87% |
| <i>Krama</i>              | 154          | 0         | -     | 8         | 5.19% |
| <i>Hai mahājana</i> (I)   | 337          | 8         | 2.37% | 13        | 3.86% |
| IMA 38                    | 493          | 2         | 0.41% | 29        | 5.88% |
| <i>Kūna cau lpæka</i> (A) | 297          | 8         | 2.69% | 0         | -     |
| <i>Vidhūrapaṇḍita</i>     | 225          | 7         | 3.11% | 9         | 4.00% |
| <i>Paṇṭām pitā</i>        | 295          | 8         | 2.71% | 0         | -     |
| <i>Kūna cau lpæka</i> (B) | 127          | 9         | 7.09% | 0         | -     |
| <i>Trīneti</i>            | 380          | 7         | 1.84% | 14        | 3.68% |
| <i>Dūnmāna khlwna</i>     | 175          | 6         | 3.43% | 9         | 5.14% |
| <i>Bākya cā'sa</i>        | 97           | 6         | 6.19% | 0         | -     |
| <i>Hai mahājana</i> (II)  | 324          | 7         | 2.16% | 20        | 6.17% |
| <i>Paṇṭām 'ū buka</i>     | 291          | 8         | 2.75% | 0         | -     |
| <i>Dūnmāna kūna</i>       | 251          | 15        | 5.98% | 4         | 1.59% |
| <i>Srī</i>                | 579          | 14        | 2.42% | 13        | 2.25% |
| <i>Prusa</i>              | 298          | 9         | 3.02% | 9         | 3.02% |
| <i>'Ariyasatthā</i>       | 194          | 13        | 6.70% | 1         | 0.52% |
|                           | 5164         | 153       | 2.96% | 154       | 2.98% |

TABLE 1: GENERAL FREQUENCIES

In weighing these au and ai rhymes as evidence and sifting the usable from the unusable it is convenient to distinguish "perfect" and "im-

perfect" rhymes. Perfect au and ai rhymes are those in which both members of a rhyming pair show graphic au or ai in their relevant syllables. Imperfect au and ai rhymes are those in which only one member shows graphic au or ai, which rhymes with some other orthographic form. Examples of both classes of rhyme will be given presently. The perfect rhymes in our group of texts represent 66.8% of all au and ai rhymes, imperfect rhymes accounting for the remaining 33.2%.

Table 2 shows that perfect au rhymes number 118 (77.1% of all au rhymes) while perfect ai rhymes number 88 (57.1% of all ai rhymes). Perfect au rhymes, mirroring the frequencies given in Table 1, rise and fall at random. With an average frequency of 2.59%, they show maxima in the case of the *Kūna cau lpøka* (B), the *Bākya cā'sa*, the *Rājaneti*, the *Kūna cau*, and the *Ḍūnmāna kūna*; minima are registered for the *Krama*, IMA 38, the *Ker(tī) kāla*, the *Trīneti*, and the *Hai mahājana* (II).

|                           | TOTAL RHYMES | au RHYMES |       | ai RHYMES |       |
|---------------------------|--------------|-----------|-------|-----------|-------|
| <i>Ker(tī) kāla</i>       | 105          | 1         |       | 2         | 1.90% |
| <i>Kūna cau</i>           | 328          | 15        | 4.57% | 10        | 3.05% |
| <i>Rājaneti</i>           | 214          | 10        | 4.67% | 1         | 0.47% |
| <i>Krama</i>              | 154          | 0         | -     | 4         | 2.60% |
| <i>Hai mahājana</i> (I)   | 337          | 6         | 1.78% | 3         | 0.89% |
| IMA 38                    | 493          | 0         | -     | 25        | 5.07% |
| <i>Kūna cau lpøka</i> (A) | 297          | 6         | 2.02% | 0         | -     |
| <i>Vidhūrapaṇḍita</i>     | 225          | 7         | 3.11% | 8         | 3.56% |
| <i>Paṇṭām pitā</i>        | 295          | 6         | 2.03% | 0         | -     |
| <i>Kūna cau lpøka</i> (B) | 127          | 7         | 5.51% | 0         | -     |
| <i>Trīneti</i>            | 380          | 5         | 1.32% | 9         | 2.37% |
| <i>Ḍūnmāna khlwna</i>     | 175          | 5         | 2.86% | 6         | 3.43% |
| <i>Bākya cā'sa</i>        | 97           | 5         | 5.15% | 0         | -     |
| <i>Hai mahājana</i>       | 324          | 5         | 1.54% | 6         | 1.85% |
| <i>Paṇṭām 'ū buka</i>     | 291          | 6         | 2.06% | 0         | -     |
| <i>Ḍūnmāna kūna</i>       | 251          | 9         | 3.59% | 0         | -     |
| <i>Srī</i>                | 579          | 10        | 1.73% | 6         | 1.04% |
| <i>Prusa</i>              | 298          | 8         | 2.68% | 7         | 2.35% |
| <i>'Ariyasatthā</i>       | 194          | 7         | 3.61% | 1         | 0.52% |
|                           | 5164         | 118       | 2.26% | 88        | 1.53% |

TABLE 2: FREQUENCY OF "PERFECT" RHYMES

Perfect ai rhymes, though more independent of the frequencies given in Table 1, are no less erratic. With an average frequency of 1.52%, they



show peaks for IMA 38, the *Vidhūrapaṇḍita*, the *Dūnmāna khlwna*, and the *Kūna cau*, but drop to zero for six texts. The absence of perceptible increase or decrease in these frequencies shows again that the incidence of perfect rhymes is dictated by the chance needs and abilities of the poets and cannot be correlated with the age or modernity of the texts.

Table 3 shows that imperfect au rhymes number 35 (22.9% of all au rhymes) while imperfect ai rhymes number 67 (43.5% of all ai rhymes). Inasmuch as these are a function of perfect rhymes, what has already been said regarding the fortuitous nature of the latter is equally applicable to them.

|                           | TOTAL RHYMES | au RHYMES | ai RHYMES |
|---------------------------|--------------|-----------|-----------|
| <i>Ker(tī) kāla</i>       | 105          | 0 -       | 1 0.95%   |
| <i>Kūna cau</i>           | 328          | 0 -       | 8 2.44%   |
| <i>Rājaneti</i>           | 214          | 0 -       | 3 1.40%   |
| <i>Krama</i>              | 154          | 0 -       | 4 2.60%   |
| <i>Hai mahājana (I)</i>   | 337          | 2 0.59%   | 10 2.97%  |
| IMA 38                    | 493          | 2 0.41%   | 4 0.81%   |
| <i>Kūna cau lpæka (A)</i> | 297          | 2 0.67%   | 0 -       |
| <i>Vidhūrapaṇḍita</i>     | 225          | 0 -       | 1 0.44%   |
| <i>Paṇtām pitā</i>        | 295          | 2 0.68%   | 0 -       |
| <i>Kūna cau lpæka (B)</i> | 127          | 2 1.57%   | 0 -       |
| <i>Trīneti</i>            | 380          | 2 0.53%   | 5 1.32%   |
| <i>Dūnmāna khlwna</i>     | 175          | 1 0.57%   | 3 1.71%   |
| <i>Bākya cā'sa</i>        | 97           | 1 1.03%   | 0 -       |
| <i>Hai mahājana (II)</i>  | 324          | 2 0.62%   | 14 4.32%  |
| <i>Paṇtām 'ū buka</i>     | 291          | 2 0.69%   | 0 -       |
| <i>Dūnmāna kūna</i>       | 251          | 6 2.39%   | 4 1.59%   |
| <i>Srī</i>                | 579          | 4 0.69%   | 7 1.21%   |
| <i>Phusa</i>              | 298          | 1 0.34%   | 2 0.67%   |
| <i>'Ariyasatthā</i>       | 194          | 6 3.09%   | 0 -       |
|                           | 5164         | 35 0.73%  | 66 1.18%  |

TABLE 3: FREQUENCY OF "IMPERFECT" RHYMES

Having defined the range of our corpus as precisely as seems warranted, we may turn now to examine its value.

Since the phonological value of au and ai is unknown in Middle Khmer but known in modern Khmer, it is permissible to propose that the perfect rhymes in our texts are divisible into three types according as their

members have HR or LR reflexes in modern Khmer. In one type, both members of the rhyme presuppose HR reflexes, e.g. nau /nəw/ : dau /təw/ (KC 16ce) and brai /prəj/ : nai /nəj/ (V 67d/68b).<sup>7</sup> In a second type, both members of the rhyme presuppose LR reflexes, e.g. smau /smaw/ : phau /phaw/ (R 36ab) and thlai /tlaj/ : kraï /kraj/ (IMA 5bc). It can be seen readily enough that neither of these types of rhymes furnishes evidence of an earlier value of au and ai. In the third type of rhyme one member is reflected by a HR form while the other is reflected by a LR form, e.g. ɽau /daw/ : dau /təw/ (KCLB 13d/14b), nau /nəw/ : praɽau /pradaw/ (KK 1bc) and prabai /prapəj/ : thhài /tɽaj/ (KC 5lef), kraï /kraj/ : vai /wəj/ (R 50ce). It is clear that rhymes of this type are no less silent regarding the pre-modern value of au and ai. It can hence be said that perfect rhymes have nothing to contribute to a solution of the problem.

This means that such evidence as we have must come from imperfect rhymes.

Imperfect au rhymes (Table 4) number 35 items and have an average frequency of only 0.73%. These consist of rhymes in which one member has orthographic au or a pre-modern variant while the other member has ūva or a variant in the relevant syllable. These are so few as to be worth listing in full. With the exception of the two rhymes in IMA 38, forms are cited in the modern orthography for the reason given above. To help the reader's understanding of what is involved, the order of the rhyming members is inverted whenever necessary so that HR members stand first.

|      |         |                 |                      |
|------|---------|-----------------|----------------------|
| HMI  | 83ab    | nau /nəw/       | : trūva /tròow/      |
|      | 94ab    | dau /təw/       | : trūva /tròow/      |
| IMA  | 31ab    | dova /təw/      | : nau /nəw/          |
|      | 138ab   | do /təw/        | : no /nəw/           |
| KCLA | 47g/48c | nau /nəw/       | : cūva /còow/        |
|      | 48fe    | dau /təw/       | : ra'ūva /rqòow/     |
| PP   | 47g/48c | nau /nəw/       | : cūva /còow/        |
|      | 48fe    | dau /təw/       | : ra'ūva /rqòow/     |
| KCLB | 14bc    | dau /təw/       | : 'āsruva /qaasròow/ |
|      | 36b/35d | dau /təw/       | : trūva /tròow/      |
| T    | 53g/54c | nau /nəw/       | : phlūva /plòow/     |
|      | 64ed    | nau /nəw/       | : kūva /kòow/        |
| DKh  | 52ba    | 'ambau /qampəw/ | : srūva /sròow/      |
| B    | 13bc    | dau /təw/       | : 'āsruva /qaasròow/ |

|                           | TOTAL RHYMES | FORM       | ITEMS  | FREQUENCY      |
|---------------------------|--------------|------------|--------|----------------|
| <i>Ker(ti) kāla</i>       | 105          | -          | 0      | -              |
| <i>Kūna cau</i>           | 328          | -          | 0      | -              |
| <i>Rājaneti</i>           | 214          | -          | 0      | -              |
| <i>Krama</i>              | 154          | -          | 0      | -              |
| <i>Hai mahājana (I)</i>   | 337          | -ūva       | 2      | 0.59%          |
| <i>IMA 38</i>             | 493          | -ova<br>-o | 1<br>1 | 0.20%<br>0.20% |
| <i>Kūna cau lpæka (A)</i> | 297          | -ūva       | 2      | 0.67%          |
| <i>Vidhūrapaṇḍita</i>     | 225          | -          | 0      | -              |
| <i>Paṇṭam pitā</i>        | 295          | -ūva       | 2      | 0.68%          |
| <i>Kūna cau lpæka (B)</i> | 127          | -ūva       | 2      | 1.57%          |
| <i>Trīneti</i>            | 380          | -ūva       | 2      | 0.53%          |
| <i>Ḍūnmāna khlwna</i>     | 175          | -ūva       | 1      | 0.57%          |
| <i>Bākya cā'sa</i>        | 97           | -ūva       | 1      | 1.03%          |
| <i>Hai mahājana (II)</i>  | 324          | -ūva       | 2      | 0.62%          |
| <i>Paṇṭam 'ū buka</i>     | 291          | -ūva       | 2      | 0.68%          |
| <i>Ḍūnmāna kūna</i>       | 251          | -ūva       | 6      | 2.39%          |
| <i>Srī</i>                | 579          | -ūva       | 4      | 0.69%          |
| <i>Prusa</i>              | 298          | -ūva       | 1      | 0.34%          |
| <i>'Ariyasatthā</i>       | 194          | -ūva       | 6      | 3.09%          |
|                           | 5164         |            | 35     | 0.73%          |

TABLE 4: FREQUENCY OF "IMPERFECT" au RHYMES

|      |           |                 |   |                   |
|------|-----------|-----------------|---|-------------------|
| HMII | 41d/42b   | nau /nəw/       | : | phīūva /pīðow/    |
|      | 42cb      | nau /nəw/       | : | phīūva /pīðow/    |
| PUB  | 47g/48c   | nau /nəw/       | : | cēcūva /caæccðow/ |
|      | 48fe      | dau /təw/       | : | ra'ūva /rɔðow/    |
| DKn  | 22c/23a   | praḷau /pradaw/ | : | trūva /trðow/     |
|      | 23ba      | bau /pəw/       | : | trūva /trðow/     |
|      | 38a       | dau /təw/       | : | trūva /trðow/     |
|      | 39a/38c   | nau /nəw/       | : | satrūva /satrðow/ |
|      | 39ab      | nau /nəw/       | : | mūva /mðow/       |
|      | 92a/91c   | bau /pəw/       | : | trūva /trðow/     |
| S    | 118c/119a | dau /təw/       | : | phīūva /pīðow/    |
|      | 119ba     | dau /təw/       | : | phīūva /pīðow/    |
|      | 145c/146a | dau /təw/       | : | kūva /kðow/       |
|      | 224c/225a | dau /təw/       | : | trūva /trðow/     |

|   |       |             |   |                  |
|---|-------|-------------|---|------------------|
| P | 83ab  | dau /tśw/   | : | trūva /trðow/    |
| A | 6g/7c | bhau /phśw/ | : | nūva /nśw/       |
|   | 7fe   | dau /tśw/   | : | phlūva /plðow/   |
|   | 9ec   | dau /tśw/   | : | sñūva /sɲðow/    |
|   | 12ce  | jrau /crśw/ | : | 'i4ūva /qəəjðow/ |
|   | 20ba  | nau /nśw/   | : | 'ū /qðow/        |
|   | 40ab  | dau /tśw/   | : | 'i4ūva /qəəjðow/ |

At first glance one is struck by the monotony of these rhymes. For example, *dau* /tśw/ 'to go' occurs 19 times, *nau* /nśw/ 'to be in or at' 13 times, *trūva* /trðow/ 'ought, must' 9 times, and *phlūva* /plðow/ 7 times. Closer examination, however, reveals a pattern which is less than monotonous: all but four of the above rhymes oppose a HR member to a LR member. Both of the rhymes in IMA 38 have a HR member opposed to another HR member. The first rhyme in the *Ḍūnmāna kūna* (22c/23a) has a LR member opposed to another HR member. The first rhyme in the *Ḍūnmāna kūna* (22c/23a) has a LR member opposed to another LR member. The first rhyme in the 'Ariyasatthā (6g/7c) again has a HR member opposed to another HR member. We shall return to these exceptions shortly. The fact of immediate importance is that 31 out of the 35 rhymes in question (88.6%) have a HR member with graphic *au* opposed to a LR member with graphic *ūva*.

In interpreting these rhymes and applying them to the problem of determining the pre-modern value of *au*, we have at our disposition four main types of evidence: (a) conclusions to be drawn from the orthography itself, (b) such historical information as we possess, (c) conclusions resulting from comparison of *ūva* rhymes not involving graphic *au*, and (d) conclusions resulting from comparison of *au* rhymes involving *ūva*.

From all we know of Indic writing systems (including the Khmer) and the languages they represent, it seems reasonable to suppose that prior to the emergence of the registers the grapheme *au* had one of, at most, four possible values: [aw], [ɤw], [ow] or [o(:)]. The last of these is dismissed from further consideration both because Khmer has had an /oo/ represented by orthographic *o* at all of its known stages and because a marked tendency to close this phoneme with the labial semivowel in otherwise open syllables is seen to be operative in Middle Khmer and modern Khmer. Hence if *au* ever represented [o(:)], the latter probably blended at an early date into [ow] or [ɤw], which differ only in degree of lip-rounding and, at least in the light of modern Khmer speech habits, would also be likely to fall together. Our options may therefore be narrowed down to [aw] and [ɤw ~ ow].

The same tendency toward diphthongisation is seen at work in the case

of ūva. On the basis of all we know, it is hard to believe that this grapheme ever represented [uw]. The majority of modern orthographic forms in -ūva correspond to Old Khmer forms in -ū, not seldom written -u but lengthened in open syllables. For example, modern srūva /sr̥ɔw/ 'unhusked rice' reflects Old Khmer srū ~ sru /sru: ~ sruw]; modern phlūva /pl̥ɔw/ 'way, path, road' appears in Old Khmer as phlū ~ plū ~ plu /pluu/ [p̥hlu: ~ p̥hluw]; modern 'īṭūva /q̥əj̥ɔw/ 'now' appears as Old Khmer 'īlū /qiilu/ [(?)i:lu: ~ (?)iluw]. However, it is during the Old Khmer period that we begin to encounter, side by side with these orthographic forms in -ū, such forms as srūv, plūv, 'ilūv for the items just cited. It seems likely that this addition of -v (now -va) to the earlier forms reflects an increasing awareness of diphthongisation in the syllable nucleus, which by this period must have begun to undergo perceptible unrounding and become [uw]. While the graphic form phlū, for example, can be interpreted as [p̥hlu: ~ p̥hluw], plūv can be interpreted as [p̥hluw]. It may in fact have been this new phonetic shape that constituted the point of departure for the development of the modern LR ū into [ɤw] /ɔw/ by regular lowering of the vowel element.<sup>8</sup> Be this as it may, it is not without reason that all but one of the 31 ūva items listed above are reflected by LR forms in modern Khmer. Table 4, moreover, shows fairly clearly that the frequency of these -au : -ūva rhymes increases toward the present. In the four earliest texts no such rhymes are found at all; after a slow start frequencies rise, albeit irregularly, and reach a peak with the most recent text. This can only mean that as the -ūva forms in question moved from [uw] to [ɤw] they became increasingly suited to the rhyme in question, namely with -au forms. The single exception in the 31 ūva items, alluded to above, is the form nūva occurring in the 'Ariyasatthā (7c). This orthographic shape would normally represent \*/nuu(w) / [nu: ~ nuw]; this particular item, however, is interpreted by Khmer scholars as pronounced nau /nəw/ [nɤw];<sup>9</sup> this is tantamount to saying that it is equivalent to LR \*ṇūva /n̥ɔw/ [nɤw].

The case of nūva accounts for one of the four exceptions to the pattern of registral contrast noted above in connection with our 35 au rhymes. In the *Dūnmāna kūna* (22c/23a) we have the rhyme praṇau : trūva, of which both members presuppose LR forms in modern Khmer. When this rhyme is compared with the other imperfect rhymes in the same text, all five of which are reflected in a HR : LR opposition, we can only conclude that praṇau rhymed with bau, dau and nau at the time this relatively late work was composed. From this detail it can be inferred that by the date of this text the grapheme ūva represented a syllable nucleus which had been lowered from earlier [uw] to its modern LR level of [ɤw], whereas au had not yet been lowered to its modern LR level of [aw] but was still

at its earlier level of [ɣw]. For if au represented an original [aw] from which the modern HR [ɣw] developed by increased closure the five other rhymes in this same text would have been impossible: nau and satrūva (39a/38c), for example, could have rhymed only if both contained the nucleus [ɣw].

The remaining exceptions to the pattern of registral contrast are the two rhymes found in IMA 38, both members of which correspond to modern HR forms. The two rhymes are the same, notwithstanding their orthographic difference: modern dau and nau. The use of o for au is not surprising in view of what has already been said of the diphthongisation of the former; this is shown clearly by the fact that dova is also written do. In view of the evidence from the other texts, earlier as well as later, the conclusion is inescapable that o(va) had rounded and unrounded allophones: [ow ~ ɣw].

Our texts include 217 rhymes (4.2% of the total) based on graphic ū, including ūva. These show an average frequency of 4.24% and frequencies for each text which, as would be expected, rule out any chronological correlation: maxima are registered for the *Ḍūnmāna kūna*, the *'Ariyasatthā*, the *Kūna cau lpæka* (A), the *Paṇṭām 'ū buka*, the *Paṇṭām pitā* and the *Keṭ(i) kāla*, in that order, and these are plainly fortuitous. Like rhymes with graphic au, those in ū fall into three main groups according as their members are reflected by modern HR or LR forms. One group, comprising 45 rhymes (20.7%), has both members of each rhyme corresponding to HR forms; these are exemplified by *yūra /juur/ : gūra /kuur/* (KK 2d/3b) and *grū /kruu/ : 'abhirū(ha) /qaphiruu/* (A 32ab). Another group, comprising 78 rhymes (35.9%), has both members of each rhyme corresponding to LR forms; these are exemplified by *paripūrṇa /barəbḍor/ : pramūla /pramḍol/* (KC 33ef) and *sūnya /sḍon/ : ṭūna /dḍon/* (P 71d/72b). In the third group, comprising 94 rhymes (43.3%), one member has a HR reflex while the other has a LR reflex; these are exemplified by *madhūra /mathuur/ : cūla /cḍol/* (R 48g/49c) and *yūra /juur/ : sampūrṇa /sambḍor/* (S 216ab). In examining such orthographic forms it cannot be seriously doubted that all rhymed by virtue of a common [u:], from which the modern LR reflexes developed by a general lowering affecting all but those nuclei which were already on the low level of openness.

Included in our 217 ū rhymes are the 33 -au : -ūva rhymes already discussed; it has already been noted that, with the exception of the anomalous *nūva* occurring in the *'Ariyasatthā* (7c), all of these forms in -ūva correspond to modern LR forms. What has not been mentioned is that orthographic ūva is found in rhyme not only with itself but also with ū. In -ūva : -ūva rhymes, 32 in number, both members of each rhyme correspond without exception to modern LR forms; these are exemplified by

skūva : 'āsrūva (IMA 93ab), trūva : tamrūva (T 80ab), and cēcūva : ra'ūva (PUB 48ce). It may be mentioned that these rhymes, showing an average frequency of 0.51%, are unevenly distributed through our texts: not found at all in the *Ker(tī) kāla*, the *Vidhūrapaṇḍita*, the *Kūna cau lpōka* (B), the *Dūnmāna khlwna*, the *Bākya cā'sa*, the *Dūnmāna kūna* or the *Prusa*, they reach a peak frequency of 2.11% in the *Tṛāneti*. In -ū : -ūva rhymes, 12 in number, the graphic form with ū is reflected by HR while that with ūva corresponds to LR; the single exception is the rhyme 'ū : 'āsrūva (P 9ab), where 'ū (modern /qðow/) is the graphic equivalent of LR 'ūva. These are exemplified by byū(ha) : trūva (PP 8g/9c) and trājū : raḷūva (KK 35b/34d). Rhymes of this type show an average frequency of only 0.29%; not found at all in ten of our texts, they register peak frequencies of 1.90% and 0.91% in the two earliest texts, the *Ker(tī) kāla* and the *Kūna cau*. This dovetails nicely with what was said above in connection with Table 4 and the increasing frequencies of -au : -ūva rhymes as we progress toward the modern period.

Adding our 32 -ūva : -ūva rhymes and 12 -ū : -ūva rhymes to the 33 -au : -ūva rhymes discussed previously, we obtain a total of 77 rhymes which include the grapheme ūva. Of these all but one, or 98.7%, are reflected by modern LR forms. That -ū : -ūva rhymes in the earliest texts were based on [u: ~ uw] in both members is to be inferred both from their orthography and from the presence of loanwords in some of them. For example, the rhyme trājū : raḷūva (KK 35b/34d) consists of a loan through Malay from Persian tarāzū 'balance, scales' and a loan from Sanskrit ṛtu 'fixed time, season'. To suppose that these two forms did not have a common syllable nucleus at the date of composition or that trājū : phlūva (KK 35bc) did not rhyme, merely because their modern reflexes do not rhyme, would be unwarranted. Nevertheless, with the gradual development of the registers during the Middle Khmer period and the emergence of a LR [ɾw] contrasting with a HR [u: ~ uw], these -ū : -ūva rhymes became increasingly dissonant. It is not surprising that after the *Cpā'pa kūna cau* they have only minimal incidence and do not occur at all in nine of the later texts under study. Indeed, the occurrence of grū : saṅkūva in such a late text as the *Cpā'pa sṛī* (134a/133c) is surprising, and suggests not so much deliberate archaizing on the part of the poet as borrowing from earlier texts. This very process can be seen in operation in the case of byū(ha) : trūva, which occurs first in the *Kūna cau lpōka* (A) (8g/9c), again in the *Paṅtām pitā* (8g/9c), and again in the *Paṅtām 'ū buka* (9c/8g). After the *cpā'pa cā'sa* or four oldest texts we see, as in Table 4, a gradual displacement of the grapheme ūva into rhyme with au and the eclipse of ū : ūva rhymes. The circumstance that the new -au : -ūva rhymes increase from earlier

to later texts is a direct reflection of the emergence of a LR.

|                          | TOTAL RHYMES | AMBIVALENT RHYMES |    | EVIDENTIAL RHYMES |       |          |
|--------------------------|--------------|-------------------|----|-------------------|-------|----------|
| <i>Ker(tī) kāla</i>      | 105          | -āya              | 1  | 0.57%             |       |          |
| <i>Kūna cau</i>          | 328          | -āya              | 8  | 2.44%             |       |          |
| <i>Rājaneti</i>          | 214          | -āya              | 2  | 0.93%             | -ī    | 1 0.47%  |
| <i>Krama</i>             | 154          | -āya              | 4  | 2.60%             |       |          |
| <i>Hai mahājana'(I)</i>  | 337          | -āya              | 5  | 1.49%             | -ī    | 5 1.49%  |
| <i>IMA 38</i>            | 493          |                   |    |                   | -ī    | 2 0.41%  |
|                          |              |                   |    |                   | -iya  | 1 0.20%  |
|                          |              |                   |    |                   | -āya  | 1 0.20%  |
| <i>Vidhūrapaṇḍita</i>    | 225          | -āya              | 1  | 0.44%             |       |          |
| <i>Trīneti</i>           | 380          | -āya              | 3  | 0.79%             | -eyya | 2 0.53%  |
| <i>Dūnmāna khlwana</i>   | 175          | -āya              | 1  | 0.57%             | -ī    | 2 1.14%  |
| <i>Hai mahājana (II)</i> | 324          | -āya              | 7  | 2.16%             | -ī    | 7 2.16%  |
| <i>Dūnmāna kūna</i>      | 251          | -āya              | 1  | 0.40%             | -ī    | 2 0.80%  |
|                          |              |                   |    |                   | -iya  | 1 0.40%  |
| <i>Srī</i>               | 579          | -āya              | 2  | 0.35%             | -ī    | 5 0.86%  |
| <i>Prusa</i>             | 298          | -āya              | 1  | 0.34%             | -ī    | 1 0.34%  |
|                          | 5164         |                   | 36 | 0.72%             |       | 30 0.47% |

TABLE 5: FREQUENCY OF "IMPERFECT" ai RHYMES\*

\*No imperfect ai rhymes are registered for the *kūna cau lpa* (A and B), the *Bākya cā'sa*, the *Paṇṭam 'ū buka*, and the *'Ariyasatthā*.

Turning now to consider the grapheme ai, we find that imperfect ai rhymes (Table 5) are divisible into two types according to the orthographic shape of their non-ai members. One type includes rhymes in which one member has ai while the other has āya<sup>10</sup> in the modern regularised orthography. Our texts contain 36 -ai : -āya rhymes, which have an average frequency of 0.72%. The other type includes rhymes in which one member has ai while the other has a grapheme other than ai or āya. Our texts contain 30 rhymes of this type, which show an average frequency of 0.47%. Since the value of āya is exactly the same as that of ai, rhymes which include it are ambivalent and must be reinterpreted as perfect rhymes divisible into three types according as their āya members have HR or LR reflexes in modern Khmer. In one type, both members of the rhyme presuppose HR reflexes, e.g. dai /táj/ : vināya /wináj/ (KC 50g/51c) and bai(ra) /páj/ : bhāya /pháj/ (HMII 80cb). In a second type, both members of the rhyme presuppose LR reflexes, e.g. thāi /tāj/ : biniscāya /piniscāj/ (KK 31b) and ʔai /dāj/ : 'āsrāya /qaasraj/ (V 31b/30d). In



the third type one member is reflected by a HR form while the other is reflected by a LR form, e.g. brai /práj/ : visāya /wisaj/ (KC 65fe) and nai /náj/ : prasrāya /prasraj/ (HMI 70cb) or thñai /tñaj/ : haḥdāya /harýtáj/ (K 1lce) and trai /traj/ : vināya /wináj/ (HMI 75ba). None of these three types of rhyme is capable of throwing light on the earlier value of ai. Yet it is not without interest that their 36 ai members are divided equally, in terms of the present, between HR and LR while their 36 āya members include only 9 HR items (25%) but 27 LR items (75%). Thus there are 6 HR : HR rhymes, 15 LR : LR rhymes, 15 HR : LR rhymes or LR : HR rhymes.

The remaining 30 rhymes (Table 5) are evidential. Not found in ten texts ranging from the earliest to the latest, these register peak frequencies for the Cpā'pa hai mahājana (II), the Hai mahājana (I), the *Ḍūnmāna khlwna* and the IMA and show no discernible increase or decrease with the date of the texts. Rhymes of this group have one member with orthographic ai and the other member with ī in 25 cases (80.7%), with iya and eyya in two cases each, and with āya (aic) in one case. As was done with imperfect au rhymes, -ai : -ī rhymes are here listed in full, the original order being inverted whenever need be so that the ai members stand first.

|     |         |                  |   |                      |
|-----|---------|------------------|---|----------------------|
| R   | 1led    | samcai /samcaj/  | : | srī /srəəj/          |
| HMI | 19ab    | nai /náj/        | : | srī /srəəj/          |
|     | 24ab    | brai /práj/      | : | ktī /kdəəj/          |
|     | 37b/36d | vai /wáj/        | : | ktī /kdəəj/          |
|     | 50cb    | nai /náj/        | : | ktī /kdəəj/          |
|     | 66cb    | thlai /tlaj/     | : | smāratī /smaardəəj/  |
| IMA | 6dc     | traiya /traj/    | : | śrī /srəəj/          |
|     | 25b/24d | didaiya /tiitáj/ | : | śrī /srəəj/          |
| DKh | 4cb     | nai /náj/        | : | khcī /kcəəj/         |
|     | 38bc    | dai /táj/        | : | smāratī /smaardəəj/  |
| HMI | 4b/3d   | krai /kraj/      | : | srī /srəəj/          |
|     | 4cd     | phdai /ptáj/     | : | thmī /tməəj/         |
|     | 6bc     | dai /táj/        | : | srī /srəəj/          |
|     | 52d/53b | cañrai /cañraj/  | : | jinasrī /cineəsrəəj/ |
|     | 80cd    | bai ra /páj/     | : | prañī /prañəəj/      |
|     | 82ba    | nai /náj/        | : | prañī /prañəəj/      |
|     | 84bc    | dai /táj/        | : | smāratī /smaardəəj/  |
| DKn | 9cb     | prabai /prapáj/  | : | ḥī /dəəj/            |
|     | 86ba    | kansai /kansaj/  | : | metrī /méetrəəj/     |

|   |         |                  |                    |
|---|---------|------------------|--------------------|
| S | 9ba     | nai /náj/        | : samt̄ī /samd̄əj/ |
|   | 22ba    | ɬadai /daatáj/   | : pt̄ī /pd̄əj/     |
|   | 59c/60a | nai /náj/        | : p̄ī /b̄əj/       |
|   | 114a    | cañrai /co.ŋraj/ | : p̄ī /b̄əj/       |
|   | 149a    | krai /kraj/      | : p̄ī /b̄əj/       |
| P | 96bc    | mai /máj/        | : kt̄ī /kd̄əj/     |

Without going into a full analysis, it can be said that orthographic  $\bar{i}$  represented [i:] in Middle Khmer but that with the development of the registers this value bifurcated into the modern HR /ii/ [i:] and the modern LR /əj/ [ɤj]. As has been described elsewhere, this latter nucleus presupposes an intermediate value consisting of the original [i:] preceded by a mid-level, typically front onglide: [<sub>e</sub>i: > e:i > ɤj]. In any case  $\bar{i}$  :  $\bar{i}$  rhymes are divisible into the same three types as other nuclei. One type consists of rhymes of which both members are reflected by HR forms (which therefore show no phonological change), e.g. dharañt̄ī /dhəərənii > thəəreənii/ : r̄ddh̄ī /rytdhii > r̄ȳtt̄hii/ (HMI 75bc) and pañj̄t̄ī /ʔbən̄jii > banciai/ : s̄t̄ī /s̄ii/ (A 5g/6c). Another type consists of rhymes both members of which are reflected by LR forms (which do show phonological change), e.g. p̄t̄ī /ʔbii > b̄əj/ : sr̄t̄ī /sr̄ii > sr̄əj/ (KK 5ab) and metr̄ī /meetrii > méetr̄əj/ : srāt̄ī /srāʔdii > srad̄əj/ (HMII 87cd). The third type consists of rhymes of which one member is reflected by a HR form, the other by a LR form, e.g. dhūl̄t̄ī /dhuulii > thuulii/ : t̄t̄ī /ʔdii > d̄əj/ (KK 18bc) and p̄āil̄t̄ī /ʔbaalii > baal̄əj/ : gamb̄tra /gomb̄iir > kumpiir/ (HMI 8cd).

Examination of our 25 ai :  $\bar{i}$  rhymes shows that the  $\bar{i}$  member is in all cases reflected by a LR form. It is reasonable to conclude that the value [i:] is not in question in these rhymes but that [ɤj] is. On the other hand, examination also shows that the ai members break down into 8 items corresponding to LR forms (e.g. samcai, thlai, traiya) and 17 items corresponding to HR forms (e.g. nai, brai, vai). From the fact that we find ai members presupposing different registers in rhyme with  $\bar{i}$  members presupposing only the LR we can only infer that the LR reflex of Middle Khmer /ii/ ( $\bar{i}$ ) had developed before registral contrast was established for Middle Khmer /əj/ (ai). It is also clear that if ai originally represented [aj] these 25 rhymes would not have been possible.

Our two imperfect ai rhymes with iya are didaiya (modern d̄īdai) /tiitáj/ : dhibvatiya (modern adhipati) /thipd̄əj/ (IMA 25bc) and 'ai /qaj/ : lokiya /lókəj/ (DKn 76a/75c). These only confirm what has just been said of ai :  $\bar{i}$  rhymes, final -iya being no more than an

allograph of  $-\bar{i}$ . The two ai : eyya rhymes are nai /náj/ : ñeyya /rėj/ (T 11ef) and thlai /tłaj/ : ñeyya /rėj/ (T 19ce). Here ñeyya is to nai and thlai as English *love* is to *drove*, and it is not without reason that the rhyme with thlai is made to continue on to ñeyya : haḥdāya (T 19ef).

The other imperfect ai rhyme is of doubtful value, since it fills an "optional" position in the meter. This is nai /náj/ : srāya /sraaj/ (IMA 145ed), which almost certainly does not involve a misapprehension of the original orthography given here. This, in any case, is the sole datum in the corpus suggesting that ai may have been [aj] at any period before the development of the LR.

Internal evidence, then, shows fairly convincingly that au and ai were [ɤw] and [ɤj] in Middle Khmer and not [aw] and [aj]. As a by-product of the analysis we uncover grounds for supposing that the LR reflexes of  $\bar{u}$  and  $\bar{i}$  developed before the LR reflexes of au and ai. The first of these conclusions conforms with the pattern of development of the modern vowel system as a whole. The second corroborates what has been said regarding the gradual manifestation of registral contrasts in the language.

## NOTES

1. I refer to this matter in "The Development of the Registers in Standard Khmer", in this volume.
2. "Natural" register is that which conforms to the general pattern: originally voiced consonants presuppose HR, originally voiceless consonants presuppose LR.
3. The earliest dated monument of Khmer is an inscription (K.557 and K.600) dated in correspondence with 611 A.D.
4. The earliest Pāli text from Cambodia is an inscription (K.759) dated in correspondence with 1308 A.D.
5. The corpus is the same as that which was developed for my "The Relative Dating of Some Khmer *Cpā'pa*", to appear in *Austroasiatic Studies*. Oceanic Linguistics Special Publications, No. 13 (Honolulu: The University Press of Hawaii, 1974).
6. By au rhymes and ai rhymes I mean rhymes in which figures at least one orthographic form of which the main syllable contains the grapheme au or ai. It is not necessary for both members of the rhyming pair to have the same grapheme.
7. Since this paper is concerned exclusively with a phonological problem, no glosses are furnished for the forms cited. The latter are given first in an Indianist transliteration of the Khmer character, then in a phonemic transcription. For fourteen syllable nuclei of modern Khmer, the acute (´) accent marks HR, the grave (`) accent marks LR. Since the analysis uses data known in the present to solve an historical problem, the transliteration reflects the modern standardised orthography with the exception of forms found in IMA 38, the text of which is available

only in a transcription of the original orthography. In this paper I write [ɾw] and [ɾj] for what I have elsewhere written [ɾ:w] and [ɾ:j], inasmuch as the accompanying phonemic forms rule out any misapprehension. Although the length of the nuclei represented by au and ai is not at issue, it must be mentioned that the traditional view is that both are short, contrasting with -āva /-aaw/ and -āya /-aaj/. While I respect this view in my phonemic transcription, nearly all of the rhymes adduced here suggest that it is no longer tenable.

8. This possibility conflicts with the interpretation advanced in "The Development of the Registers", according to which LR ū /ðow/ [ɾw] developed from \*[<sub>ə</sub>u:] by generation of a lower onglide before original /uu/ and in parallel with LR ī /ðəj/ [ɾj], from \*[<sub>ə</sub>i:] by generation of a lower onglide before original /ii/.

9. *Vacanānukrama khmṛa* (Phnom-Penh: Institut Bouddhique, 1968), I: 519a.

10. The tilde corresponds to the *saṃyogasaññā*, a diacritic borrowed from Thai to mark certain unpredictable vowel qualities.



## THE WORD $\bar{a}c$ IN KHMER: A SEMANTIC OVERVIEW<sup>1</sup>

SAVEROS POU

1.1. Modern Khmer (mod.K.)  $\bar{a}c$  /ʔaac/ is not what would normally be called a problem word. It is a verb pure and simple, and as such is likely to be modified by adverbs and, in particular, post-posed resultative verbs covering the ideas of '*to be possible*' and '*to be not possible*'. It belongs to the written language more because of its refined associations than because of any obsolescence.

1.2. Turning to our two Khmer dictionaries, we find  $\bar{a}c$  defined as follows:

VK, 1719b:  $\bar{A}c$ . a) *to be brave, courageous, fearless*; b) *to be entitled to; to be possible, can*.

DCF, 35b:  $\bar{A}j$ .<sup>2</sup> *to dare; daring, bold*.

Although there is no apparent semantic problem, in reality these glosses and their illustrations in the two dictionaries are far from adequate.

The best-known derivative of  $\bar{a}c$  is  $\text{am}\bar{n}\bar{a}c$  /ʔəmnaac/, which is glossed as follows:

VK, 1848a:  $\text{Am}\bar{n}\bar{a}c$ . a) *courage, boldness*; b) *capacity, title*;  
c) *strength, force, power*.

DCF, 80b:  $\text{Am}\bar{n}\bar{a}c$ . a) *power, authority*; b) *right, title, liberty, permission*.

It is clear enough from the foregoing that  $\text{am}\bar{n}\bar{a}c$  is a nominal derivative of  $\bar{a}c$ , but once again the glosses provided by our two dictionaries are far from satisfactory and far from complete.

Another derivative, formed with prefix /bən-/, is glossed as follows:

VK, 533b:  $\text{Pa}\bar{n}\bar{a}c$ . a) *capacity, right*; b) *to encourage, make bold* (causative derivative).

DCF, 915a: Pañ-āc. *daring, fearless.*

Here we note an obvious discrepancy between the two dictionaries; in my view, this points up the obsolescence of this derivative. The editors of the dictionaries have definitely failed to pinpoint the full sense of the item, and have therefore failed to assign it to the same wordclass.

1.3. Borrowed by Siamese, probably before the appearance of the earliest written document (end of thirteenth century), āc seems to have followed the same lines of development as in Khmer, as is suggested in

TED, 989b: Āc. *daring; bold; fearless; valiant; undaunted...; ...capable; proficient; qualified; fitted.*

Siamese also has the same two derivatives of āc as does Khmer, and I am inclined to believe they were borrowed at approximately the same period as the wordbase.

TED, 996b: Āmnāc. *power; authority; right; jurisdiction; ability; capability; ...licence.*

TED, 475a: Pāñ-āc. *to be very bold, daring or fearless; to overstep the bounds of propriety.*

The striking similarity between the Khmer and Siamese glosses is significant insofar as it shows that their editors' investigations were limited to the modern languages.

1.4. This much leads us into problems raised by the interpretation of āc as it occurs in older documents, namely of the Old Khmer and Middle Khmer periods. It must be mentioned here and now that these problems are not recognisable as such in translations of the Old Khmer inscriptions edited successively by Aymonier, Finot and Cœdès (cf. 2.2.).

For example, I myself was quite satisfied with the definitions of āc given in various dictionaries until I began, some five years ago, to analyse the Middle Khmer texts. It was only then that it occurred to me that the semantic range of āc was markedly broader and richer than I had suspected, and that it required further consideration and a reworking of its definition. The next step was intensive analysis of the Old Khmer inscriptions independently of their translations. Here again all occurrences of āc pointed to a breadth of meaning previously unknown to myself and previously unrecorded in the dictionaries. But the problem is even larger: Professor Alexander B. Griswold, the well-known student of the Old Siamese inscriptions, has recently called my attention to the fact that the various dictionaries cited (cf. 1.2., 1.3.) likewise do not cover all the senses that āc has in Siamese texts.



1.5. Sensing a need for external evidence, I next turned to Mon<sup>3</sup> with a view to finding a cognate which would throw light on the problem. It was a disappointment at first to find that  $\bar{a}c \sim \bar{a}t$  in Old Mon and modern Mon alike is confined to the notions of 'to ask, to request', which seemed to rule out any possibility of comparison. However, it was not long before I was brought back to this question, for my most recent comparative studies<sup>4</sup> seem to show the importance of the pre-inscriptional stage of Khmer and Mon, when these languages probably had more in common than is visible through written documents. This supposition led me to ask myself, "Could Khmer  $\bar{a}c$  ever have meant 'to request', as in Mon?"

2. The answer to this question calls for re-examination of representative texts of all three periods. In presenting the results of my analysis I choose, for the sake of clarity and the reader's convenience, to work forward from the earliest texts rather than in the contrary direction. At the same time, data from modern Khmer are adduced whenever they seem to throw light on data from earlier periods.

2.1.  $\bar{a}c$  is one of the most current lexical items in the Old Khmer inscriptions, being particularly common in the Angkorian inscriptions.

Analysis calls first of all for consideration of the kinds of speech involved. Briefly, occurrences of  $\bar{a}c$  in a) "actual" speech (narrative, statement) must be distinguished from occurrences of  $\bar{a}c$  in b) "virtual" speech (wish, injunction, potentiality). We turn first to consider virtual speech, the more complicated of the two types, since any light thrown on it will elucidate actual speech, whereas the contrary does not apply.

Occurrences of  $\bar{a}c$  in virtual speech conform to the following pattern:

| I. Actual speech | II. Virtual speech                |
|------------------|-----------------------------------|
| A. _____         | _____                             |
|                  | $\bar{a}c \text{ ti}$             |
|                  | $vva\text{m} \bar{a}c \text{ ti}$ |

This means that  $\bar{a}c$  introduces a clause in virtual speech (II) connected with a clause in actual speech (I) that precedes it. To pinpoint its sense and function it may be helpful to consider similar patterns of virtual speech in Old Khmer and modern Khmer:

| I. Actual speech | II. Virtual speech         |
|------------------|----------------------------|
| B. (OK) _____    | _____                      |
|                  | $le\text{ñ}$               |
|                  | $kam\text{pi } le\text{ñ}$ |
| C. (OK) _____    | _____                      |
|                  | $j\bar{a}$                 |
|                  | $vva\text{m} j\bar{a}$     |



In other words, he expresses his way of thinking and/or his desire regarding the object of the preceding clause (I).

The case of  $\bar{a}c$  (A) is a particular problem, since no known modern senses of it (cf. 1.2., 1.3.) fit the contexts; hence the need for comparative investigation.

2.4. In sorting out the data of virtual speech in modern Khmer, I chanced upon constructions with  $s\bar{u}m$  (D), a marker of the optative but primarily a verb meaning 'to ask, request, ask for a favour, wish'. This brought me back to Mon  $\bar{a}c \sim \bar{a}t$ , of similar meaning, and led me to posit 'to ask, to request' for proto-Khmer / $\bar{a}ac$ / in order to account for the semantic and grammatical development of  $\bar{a}c$  in historical times. Before entering into details on this point, it may be useful to sum up the foregoing comparative considerations:

Pre-inscriptional Khmer

|              |            |
|--------------|------------|
| Optative     | $\bar{a}c$ |
| Desiderative | $j\bar{a}$ |
| Injunctive   | ?          |

Inscriptional Khmer

|              |            |
|--------------|------------|
| Optative     | $leñ$      |
| Desiderative | $j\bar{a}$ |
| Injunctive   | $\bar{a}c$ |

Modern Khmer

|              |              |
|--------------|--------------|
| Optative     | $s\bar{u}m$  |
| Desiderative | $guor$       |
| Injunctive   | $tr\bar{u}v$ |

3. While the above table helps elucidate the changes which took place in the system, they emphasise in particular the change which occurred with  $\bar{a}c$  early in the history of Khmer, probably within or before pre-Angkorian times (sixth to eighth century A.D.). We turn now to consider how  $\bar{a}c$  passed from virtual speech to actual speech.

3.1. We first posit for virtual speech  $\bar{a}c$  'to wish':

|            |                                  |                                |
|------------|----------------------------------|--------------------------------|
| (Optative) | $\bar{a}c$ ti _____              | <i>'wishes that...'</i>        |
|            | $vva\bar{m}$ $\bar{a}c$ ti _____ | <i>'does not wish that...'</i> |

This optative notion, taking on force, blends into an injunctive or jussive encompassing "orders" of various nuances. Thus:

(Injunctive)      āc ti \_\_\_\_\_ 'orders that...'  
                   vvaṃ āc ti \_\_\_\_\_ 'does not order that...'

Inasmuch as the Khmer verb has no formal distinctions of voice, the above glosses could also be 'it is [not] desired that...' and 'it is [not] ordered that...'. In the two negative constructions, moreover, vvaṃ may negate the dependent verb: 'wishes that...not...' and 'order that...not...'.  
 'wishes that...not...'  
 'order that...not...'

Illustrating this injunctive function are numerous instances to be culled from the Angkorian inscriptions, the majority of which are concerned by their nature with the "orders" of sovereigns or high-ranking officers. I confine myself to a few of these dating mostly from the tenth and eleventh centuries.

a. K.444, IC II, p.64, A (19-21):

nā            paṃnvās        gus    gi    nā    thve    vraḥ    rājakāryya    vvaṃ    āc  
 as for religious only who do Royal Corvée not ordered  
 ti                    oy    vrīhi    paryyañ    viṣaya        cañcūli    phon  
 passive marker give rice oil district ricinus plural

'Only the religious serving in the Royal Corvée are ordered not to provide rice, local oil, and ricinus [oil].'<sup>7</sup>

b. K.152, IC V, p.192 (13-16):

ri            anak        ta        vvaṃ    thve    rañka            caṃnāṃ            sañkalpa  
 as for person linker not do rice ritual supply ritual supply  
 vvaṃ            āc ti        nau    ta        bhumyāsrama        noḥ.  
 not (are) ordered stay linker land hermitage that

'Those who do not provide ritual rice are ordered not to remain in the hermitage of this land.'

3.2. Here we come to a kind of turning-point where, in the mind of Khmer speakers, an actual transformation takes place. It is but a short step from 'I am desired or ordered to...' to 'I am obligated or entitled to...'. Conversely, 'I am not ordered to...' easily passes into 'I am not required or allowed to...'. This first step in the shift clearly runs parallel to the initial semantic enlargement. Thus, to use our previous examples:

a. It is ordered to

(VP<sub>1</sub>)

only the religious serving in the Royal Corvée  
 (recipient)

not to provide rice, etc.

(VP<sub>2</sub>)

= *Only the religious, etc.*  
(recipient)

|                        |                                |
|------------------------|--------------------------------|
| <i>may not</i>         |                                |
| <i>are allowed not</i> | [to] <i>provide rice, etc.</i> |
| (VP <sub>1</sub> )     | (VP <sub>2</sub> )             |

b. *It is ordered to*  
(VP<sub>1</sub>)

*persons not providing ritual rice*  
(recipient)

*not to stay in the hermitage of this land*  
(VP<sub>2</sub>)

= *Those who do not provide ritual rice*  
(recipient)

|                        |                                           |
|------------------------|-------------------------------------------|
| <i>may not</i>         |                                           |
| <i>are not allowed</i> | [to] <i>remain in the hermitage, etc.</i> |
| (VP <sub>1</sub> )     | (VP <sub>2</sub> )                        |

The first semantic shift may hence be summarised as follows: '*it is requested of X to do*' > '*it is ordered of X to do*' > '*X must, is allowed (entitled), has the right, to do*', or the corresponding negative.

3.3. From the idea of '*to be entitled*' the semantic development proceeds one step further:

c. K.33, IC III, p.150 (30-32):

|                       |               |                    |                            |                       |          |               |
|-----------------------|---------------|--------------------|----------------------------|-----------------------|----------|---------------|
| ti                    | añ            | svaṃ               | leñ                        | āc ti                 | añ       | ta            |
| <i>passive marker</i> | <i>I</i>      | <i>ask</i>         | <i>optative marker</i>     | <i>be entitled</i>    | <i>I</i> | <i>linker</i> |
| paṃnvas               | pvās          | paripāla           | vvaṃ                       | āc ti                 |          |               |
| <i>religious</i>      | <i>enter</i>  | <i>clergy</i>      | <i>take care</i>           | <i>not be allowed</i> |          |               |
| āyatva                | ta            | kaṃloñ             | phoñ                       |                       |          |               |
| <i>depend on</i>      | <i>linker</i> | <i>dignitaries</i> | <i>collective particle</i> |                       |          |               |

'*It is my request, I who am a religious in the clergy, to be entitled to take care of [this foundation], not to have [it] depend on the mandarins.*'

d. K.933, IV, p.48, B (15-16):

|               |                  |               |             |               |                  |                      |
|---------------|------------------|---------------|-------------|---------------|------------------|----------------------|
| nau           | anak             | pvās          | ta          | mān           | dharmma          | pi                   |
| <i>as for</i> | <i>religious</i> | <i>linker</i> | <i>have</i> | <i>virtue</i> | <i>marker of</i> | <i>coming action</i> |

āc            paripālana                            pi                            vvaṃ    pidā            noḥ  
*be up to    take care    marker of coming action    not    oppress    that*  
 gi ta                            āc                            stap                            vartamāna    caṃnyar                            dau  
*linker    be entitled    listen to                            news    future    translative marker*

*'Religious endowed with virtue who would be up to taking care of [the āśrama] and would not do harm to it would be entitled to receive official information from now on.'*

- e. K.693, IC V, p.205, B (27-28):

vraḥ    sabhā    tantyañ            daha                            āc                            vyavahāra    nu                            vāp  
*Royal Court    inquire    whether    be up to                            sue    then    male title*  
 vrahmaputra                            kathā                            man                            vvaṃ                            āc  
*personal name    declare    particle                            not    be up to*

*'The Court asked vāp Vrahmaputra if he meant [or was up] to bringing suit; he then declared that he did [or was] not.'*

The following is a typical verbal phrase found in descriptions of land or ricefields:

- f. K.817, IC V, p.201 (12-13):

bhūmi    noḥ                            ti                            duñ    ta                            anak    ta                            roḥh    noḥ  
*land    that    passive marker    buy    linker    person    linker    like    that*  
 phoñ                                                    āc                            ti                            vroḥ    bhay    mvay    tap  
*plural marker    capable linker    sow    twenty one ten*

*'The land bought from those people has a capacity of thirty sowings of grain.'*

- g. K.258, IC IV, p.179, A (11-12):

bhūmi    noḥ    nu    vraī    vnur                            mvat                                                    āc                            ta                            vroḥ  
*land    that    and    forest    hill    put together    capable    linker    sow*  
 śata                            1  
*hundred    one*

*'The land together with the forests and hills has a capacity of a hundred sowings of grain.'*

3.4. Another aspect of the semantic development of āc involves the moral connotation growing out of the idea of capability, which is to be borne in mind if we are to account for āc in Middle Khmer texts.

- h. K.227, BEFE0 XXIX, p.309 (22):

cāmpa sroṃ saṃtac ayat anak mvay ta āc  
*Cham surround [His] Majesty without person one linker capable*  
 chpañ ley  
*fight ever*

*'The Cham surrounded His Majesty [Yasovarman] without any of his men being capable [or having the courage] to fight back.'*

- i. K.659, IC V, p.146 (28-29):

nau sādhu ta āc varddheya neḥ  
*as for good linker have courage make prosperous this*  
 pūnya neḥ ge mān svargga  
*good deed this they have heaven*

*'Good people who have the courage [or are determined] to take good care of this deed [the cult of a śivaliṅga] will obtain the heavenly reward.'*

- j. K.705, IC V, p.199 (10-11):

ri ta āc paripālana ge svey divyaloka  
*as for linker have courage take care they enjoy celestial world*  
 aṃval nu gaṇa phoñ  
*together with group plural marker*

*'Those who have the courage [or have the merit] to take care of [the cult] will enjoy [the bliss] of the celestial world together with their folk.'*

Before moving on to the Middle Khmer texts, it seems worthwhile to consider a few of the more involved instances of āc illustrating its semantic effect.

- k. K.292, IC III, p.208 (lines unnumbered). This is the famous oath of allegiance of royal officers to Sūryavarman I.

daha āyuḥ yeñ āc dan ta vraḥ rājakāryyavelā lvoḥ ta  
*if life we can reach linker time of Royal Corvée arrive linker*  
 kāla nau prakāra leha yeñ slāp hetu bhakti  
*death as for kind whatever we die cause devotion*  
 gi pi yeñ thve  
*future marker we do*

*'In the event that our lives can [fortunately] last as long as the*

*time of the Royal Corvéé, whatever manner of death befall us, we shall accept it out of devotion [to His Majesty].'*

The best of these instances are provided by a text engraved at different places (K.444, IC II, pp.63-5; K.868, IC VI, pp.170-2; K.175, IC VI, pp.175-8) reporting the creation of two varṇa or 'corporations of Royal Officers' by Jayavarman V and consequently containing a number of regulations expressed by means of āc or vvaṃ āc. I choose only one case:

1. K.444, IC II, B (9-13):

nau rū kule ta anak si man ka āc ti  
as for relative linker person male linker capable

paryyann hoñ nāṃ mok oy ācāryya... pre paryyann  
teach indeed bring to teacher order teach

siksā āy nagara pi pre nā vraḥ  
learn at city marker of coming action order at Royal

rājakāryya nā pamvās vvaṃ āc ti bhāgavata  
Royal Corvéé at religious not be entitled Bhāgavata

*'Male members of the family capable of receiving education [will be] brought to the ācāryya. These are instructed to teach and train the former in the City and to employ them as religious in the Royal Corvéé [but] not as Bhāgavata.'*<sup>8</sup>

3.5. The results of the foregoing analysis (c to l) may be summarised as follows:

āc has the following nuances:

*'to have the ability, capability, competence, to; to be up to';*

*'to have the power, right, title, mandate, to';*

*'to have the intention of, mean to';*

*'to have the courage to';*

*'to have the merit, good fortune, to';*

*'to have the (supernatural) power to'.*

4. Bearing the last value in mind, we may now move on to the Middle Khmer texts.

m. IMA 31 (between 1684 and 1687 A.D.), lines 11-13:

braḥ aṅg jā bvūṃnāk Trey trān rohān sambah  
sacred being body be refuge shore shelter like junk

ratn ruñ ūtuñ samār ṅā nū āc nām satv  
gem wide noble capable that is capable lead creature



phoñ                      chloñ    leñ    vattasañsār  
*plural marker    cross    leave    cycle of transmigration*

*'The Lord [Buddha] was the refuge, the sheltering shore like unto a precious, broad, and majestic sea-going junk capable of transporting all creatures across and beyond the [Ocean] of transmigrations...'*

This passage illustrates the type of equational expression that arises in a community of people employing two languages, a learned and a popular idiom - in the case of Cambodia, Sanskrit and Khmer. The group samār ñā nū āc is tantamount to saying or writing samār = āc or, better yet, samār (= āc). In other words, samār (< Skt. samartha 'capable') is synonymous with our āc.

Here are instances from the cpāp', ethical and didactic texts in verse composed by religious teachers for their disciples:

n. C.Kr., stanza 4:

mtāy      ūbuk      sot ñā  
*mother    father    moreover*

---



---

āc                      mak                      paṃpuos  
*have courage    come    cause to enter clergy*

hetu      cañ      sambhār  
*because    want    accumulation*

---



---

*'Your parents, moreover, ...had the courage [or the virtue or merit] to bring you here and make you a religious because they wanted to accumulate [merits for themselves and for you].'*

o. C.Kr., stanza 33:

rī              grū      pādhyāy  
*as for    guru    preceptor*

praṭūc      nūv      mtāy  
*compare with    mother*

ūbuk                      eñ hoñ  
*father                      indeed*

āc                      oy      ktī cpā'p  
*have capacity give principles*

praṭau      dharm              phoṅ  
*instruct Dharma collective marker*

*'As for your preceptors, they are comparable to your parents, [since] they are qualified [or have the merit or virtue] to teach you all the principles and the Dharma...'*

Slightly different are the two following cases:

p. C.K.C., stanza 67:

kra              bāky              biroḥ  
*rare              word              harmonious*

kra      citt      āc      smoḥ  
*rare      heart      capable      devoted*

---

*'Rare are sweet [agreeable] words, rare is the heart brave [bold] enough to be devoted [to others in distress]...'*

q. C.K.C., stanza 68:

kra      man      baky      saty  
*rare      be      word      true*

kra      citt              āc      kāt'  
*rare      heart      capable      cut off*

sralaḥ      bī              sneh  
*free              from              love*

---

*'It is hard to find true words and a heart bold enough to get free from attachment...'*

4.1. I have so far concerned myself with the wordbase rather than any derivatives of āc, on the grounds that the occurrences of the former greatly outnumber occurrences of the latter. Moreover, I have assumed that the semantic expansion of āc over the centuries has to some extent had a bearing on the development of the derivatives. Here I shall confine myself to the most common of these, amṇāc (cf. 1.2.) and mention one aspect of its development.

Taking it as understood that amṇāc has progressed from the sense of \*'request' to that of 'title, right; ability, capacity; strength, energy,

(personal, political, administrative) power' and on to 'virtue, merit; courage, boldness', I find still another connotation which goes beyond the moral and religious implications of the term. In its sense of 'power', *aṃṇāc* also includes the domain of the supernatural and conveys the idea of force emanating either from the divine world above us or from magic. One can possess *aṃṇāc* 'potential force' because (a) he has accumulated merit in the past, even in previous existences, or because (b) he has acquired magical knowledge. This personal *aṃṇāc* enables him, consciously or otherwise, to perform quasi-miracles, and in this sense the word is equated with Skt. *tejas* 'hot and piercing power'. This point is illustrated by the following passages:

r. IMA 39 (1747 A.D.), p.113 (29-31):

|           |             |          |             |          |           |        |        |     |    |      |
|-----------|-------------|----------|-------------|----------|-----------|--------|--------|-----|----|------|
| noḥ       | vā          | cor      | kaṃṇāc      | pāk      | papp      | khlap  | khilāc | toy | nū | braḥ |
| then      | bandits     | evil     | routed      | submit   | thanks to | sacred |        |     |    |      |
|           |             |          |             |          |           |        |        |     |    |      |
| aṃṇāc     | braḥ        | pāda     | anak        | jā       | aṃcass    | noḥ    | pān    |     |    |      |
| power     | His Majesty | person   | be          | lord     | then      | obtain |        |     |    |      |
|           |             |          |             |          |           |        |        |     |    |      |
| parapūr   | sruk        | khsem    | khsān       | sāpāy    | dūlāy...  |        |        |     |    |      |
| abundance | country     | pacified | comfortable | well off |           |        |        |     |    |      |

*'These evil bandits were then routed and submitted, thanks to the supernatural power of His Majesty our Lord, and the country resumed abundance, peace, comfort, and prosperity...'*

s. IMA 39, p.114 (40):

[*The General came out victorious and dispersed the enemy*]

|           |       |       |             |        |     |      |    |           |
|-----------|-------|-------|-------------|--------|-----|------|----|-----------|
| toy       | nū    | tejaḥ | aṃṇāc       | braḥ   | pād | anak | jā | aṃcass... |
| thanks to | power | power | His Majesty | person | be  | lord |    |           |

*'...thanks to the supernatural power of His Majesty our Lord...'*

Both of these texts refer to victories won by the same outstanding general. Both, moreover, reflect the collective mind of the times which interpreted all happenings as manifesting an immanent supernatural power.

4.2. To bring our inquiry to a close, I propose the following statement of the semantic range of *āc* and its derivative:

|           |                                                                                                                        |
|-----------|------------------------------------------------------------------------------------------------------------------------|
| <i>āc</i> | <i>to ask, request, require; (impersonal) to be desired, ordered to do;</i>                                            |
|           | <i>to be required, obliged, expected (supposed) to do; to be allowed (entitled), have the right, to do; to be able</i> |

*(capable), competent, qualified to do; to mean (intend), be determined to do; to have the courage, have the merit (virtue) or good fortune to do; to have the power (authority) or supernatural power to do.*

**amṇāc** *request, wish;*

*liberty, permission, right, title, capacity;*

*force, (personal, political, administrative, judicial, supernatural) power, (personal, local, national) authority.*

## NOTES

1. All extracts from texts are given in transliteration, indicated by bold ("Courier") type, thus: **āc ti**; **vvaṃ āc ti**, etc.
2. Obsolete spelling.
3. H.L. Shorto, *A Dictionary of Modern Spoken Mon*. London, Oxford University Press, 1962; and, in particular, his *A Dictionary of the Mon Inscriptions from the Sixth to the Sixteenth Centuries*. London, Oxford University Press, 1971.
4. Cf. my attempt to reconstruct the semantic evolution of Mon-Khmer /siŋ/ in a forthcoming article, "Recherches sur le vocabulaire cambodgien: VIII. Du vieux khmer au khmer moderne", in *Journal Asiatique*.
5. It is not my intention to blame the translators of the Old Khmer inscriptions, most of whom were eminent scholars, for having rendered **āc** so vaguely. Their chief aim, after all, was to make these texts available for historical use. For this we must be thankful. My point here is that we must seek to improve upon their translations by applying the techniques of linguistic analysis.
6. It is sometimes forgotten that the so-called particles of Khmer are originally lexical items with full meaning and that they tend to carry part of their original meaning even when grammaticalised. Exceptions to this general rule are a very few particles such as **nu ~ nau** 'and', **dañ ~ doñ ~ dāṃṃ** 'and, together with', and **pi ~ pī** 'marker of coming action'.
7. In the glosses I use the English passive to show the rôle of **ti** as a passive marker.
8. My translation, which differs from that of Cœdès (1942), is intended to achieve greater accuracy by reflecting the structure of this passage.

Cœdès's comment (1942:67, note 5) on bhāgavata is worth reproducing:  
"The Bhāgavata represented, together with the Pāñcaratra and Sāttvata, one of the Viṣṇuite sects of Cambodia (cf. BEFE0, XXXII, p.106). According to this last text, they were not considered regular religious (paṃnvās), probably because monkhood was the privilege of the Śaivite sects."

#### ABBREVIATIONS

- BEFEO *Bulletin de l'Ecole Française d'Extrême-Orient.*
- C.K.C. Cpāp' kūn cau, included in Cpāp' phseñ-phseñ, 4th edition. Phnom-Penh, Institut Bouddhique, 1970.
- C.Kr. Cpāp' kram, included in Cpāp' phseñ-phseñ (as above).
- DCF J. Guesdon, *Dictionnaire cambodgien-français*, 2 volumes. Paris, 1930.
- IC G. Coedès, *Inscriptions du Cambodge*, 8 volumes. Paris and Hanoi, 1937-1968.
- IMA *Inscriptions modernes d'Angkor*, 2nd edition. Phnom-Penh, Institut Bouddhique, 1958.
- K. Inventory number of Old Khmer inscriptions in IC.
- Skt. Sanskrit.
- TED G.B. McFarland, *Thai-English Dictionary*. Stanford, 1956.
- VK *Vacanānukram khmaer*, 5th edition. Phnom-Penh, Institut Bouddhique, 1968.





## A CLASSIFICATION OF VERBS IN VIETNAMESE AND ITS PEDAGOGICAL IMPLICATIONS\*

NGUYEN DANG LIEM

### INTRODUCTION

For over a generation, there was the conviction that the best language teaching materials are based upon a contrastive analysis of the language to be learned and the language of the learner. Allied with the conviction was the hypothesis that the new linguistic system, and by extension, the new cultural behaviour, should be established as a set of new habits by drill, drill which would ensure overlearning. Such a pedagogical philosophy was systematised mainly by Charles C. Fries (1945) and Robert Lado (1957).

However, all this firm belief in contrastive analysis seems to be in the past, at least for some people. The generative-transformational theory claims that language behaviour is rule-governed behaviour, and consequently that language learning should be in the form of a process of internalising the creative rules in the new language, and not just that of mere habit forming.

Before such a theoretical dilemma, a language teacher may be tempted to make the most use of existing linguistic theories to improve teaching materials as best as he could. With this pragmatism in mind, I am trying to present a contrastive analysis of Vietnamese and English on the case,

---

\*Paper presented in absentia at the Second Annual Meeting of the American Council of Teachers of Uncommonly-taught Asian Languages, Boston, Massachusetts, November 22-25, 1973.

This paper is an excerpt from *Cases, Clauses, and Sentences in Vietnamese, A Contrastive Vietnamese-English Analysis* which was submitted to the Culture Learning Institute of the East-West Center as the final report on a contrastive analysis project sponsored by the CLI/EWC.

clause, and sentence levels in this study.

On the theoretical linguistics ground, this analysis of cases, clauses, and sentences in Vietnamese in contrast with English will make an attempt to utilise both tagmemics (Brend 1970, 1972a, and 1972b, Cook 1969, Longacre 1964, Pike 1954, 1955, 1960, 1966, and 1971) and a case grammar model (Fillmore 1968, 1969, and 1971) called lexibase (Manley 1972, Starosta 1971a, 1971b, 1972a, 1972b, 1972c, 1973, Taylor 1972). Such an eclectic combination of tagmemics and case grammar is not new either in the field of general linguistics (Cook 1970, 1971a, 1971b, and 1972, Heidi Platt 1970, John T. Platt 1971) or in Southeast Asian linguistic analysis (Liem 1971a, 1971b, 1972a, 1972b, 1973a, and 1973b). Case grammar studies on Vietnamese can be found in Clark 1971a and 1971b, Ha 1970, Manley 1969, Taylor 1968, and Tran 1971, and a tagmemic analysis of Vietnamese can be seen in Liem 1969.

In that this study is tagmemically oriented, it recognises firstly the hierarchical levels in syntax and secondly the grammatical unit or tagmeme as composed of a slot and a filler class. The concept of hierarchies in syntax makes it possible to single out the clause and sentence levels for this study, and consequently it is unnecessary to write, at the very beginning, phrase structure rules of the entire base component, a matter of mandatory nature in a transformational analysis (Chomsky 1965:88ff.). The grammatical slot of the tagmeme at the clause level has been pointed out by Becker (Becker 1967a and 1967b), Fries (Fries 1970), and Young, Becker, and Pike (Young, Becker, and Pike 1970) as having not only the overt syntactic relationships (case realisations or case forms) but also the covert meaning relationships (case relations) with the predicative verb which is considered to be central in the clause (Chafe 1970).

In that this paper is also case grammar oriented, it recognises the case relations between various nouns and the predicative verb in the clause (or proposition in Fillmore's terminology). The type of case grammar utilised here introduces both the overt case realisations (or case forms) and covert case relations into the grammar as syntactic features assigned to nouns by verbs (Starosta 1971a, 1971b, 1972a, 1972b, and 1973).

In that this is also a contrastive analysis of Vietnamese and English, the study will present not only the particular construction in Vietnamese but also will point out similarities or differences in parallel constructions in English. However, specific types of drills to teach the patterns will not be suggested because generally drills on case forms and case relations in Chapter I\* will be substitutional, whereas they will be

---

\*This paper is an excerpt from a larger work.

surface-to-surface transformational on the clause and sentence levels in subsequent chapters. The phrase level has been purposely omitted in this study because there are not many similarities on that level in Vietnamese and English. The reader is referred to Liem 1969, chapters III and IV, pp. 75-163 for an analysis of the phrase level in Vietnamese. Particular attention should be paid to the verb phrase on pages 75-108, the noun phrase on pages 109-131, and the cross-level tagmemes and cross-level constructions on pages 160-164 of the 1969 study.

## CLASSIFICATION OF VERBS

The following verb classification is based upon identificational-contrastive features such as the presence or absence of nuclear case form/case relation tagmemes. There are sixteen classes of verbs in the proposed classification. According to their hospitality to case-marked tagmemic slots, verbs in Vietnamese can be classified into sixteen groups as indicated in Chart IV. The sixteen verbs representative of their classes will be listed below with their hospitable case-marked tagmemic slots. It is noted that all the case-marked tagmemic slots, except the slot [+O, +OBJ] hosted by the submissive verbs, can be filled by noun phrases. When they can be filled by other than noun phrases, they will be so indicated. Numbers will refer to the examples found earlier in this study. Learning problems for the English speaking student will be pointed out, if any. A list of sample verbs in each group will be given as examples for the preparation of substitution drills.

### 1. là 'be'

The verbs in this class host [+NM, +OBJ] and [-K] (or caseless) tagmemic slots:

[+[+NM, +OBJ] \_\_\_ [-K]]

|                                                                                                                                 |   |                                                                                                                                 |
|---------------------------------------------------------------------------------------------------------------------------------|---|---------------------------------------------------------------------------------------------------------------------------------|
| $\left. \begin{array}{l} \text{noun phrase} \\ \text{verb phrase} \\ \text{clause} \\ \text{prep. phrase} \end{array} \right\}$ | } | $\left. \begin{array}{l} \text{noun phrase} \\ \text{verb phrase} \\ \text{clause} \\ \text{prep. phrase} \end{array} \right\}$ |
|---------------------------------------------------------------------------------------------------------------------------------|---|---------------------------------------------------------------------------------------------------------------------------------|

Both [+NM, +OBJ] and [-K] slots can be filled by a noun phrase, a verb phrase, an independent clause, or a prepositional phrase:

Ông ấy là bác-sĩ. (2.1.)

*He is a doctor.*

[+NM  
+OBJ]

[-K]

[noun phr.]

[noun phrase]

Đi là chết.

*To go is to die.*

[verb phr.]

[verb phr.]

Ông ấy đi là ông ấy khôn.

*he go be he intelligent (He was wise to have gone.)*

[clause]

[clause]

Từ đây đến đó là hai dặm.

*From here to there is two miles.*

[prep. phrase]

[noun phrase]

The filling of the [+NM, +OBJ] slot hosted by an equational verb such as là 'be' by a verb phrase or an independent clause in Vietnamese is a learning problem on the production level for English speakers who are used to a *To* verb phrase, or a dependent clause introduced by *That* in the same slot in English:

*To go is to die.*

[+NM  
+OBJ]

[-K]

[to verb phr.] [to verb phr.]

*That he went was a good move.*

[that clause]

[noun phr.]

The verb là 'be' seems to be the lone verb of its class. There is another similar verb, làm 'to exercise the profession of' in that it is also an equational verb:

Ông ấy làm bác-sĩ.

*He is a doctor (professionally).*

[+NM  
+OBJ]

[-K]

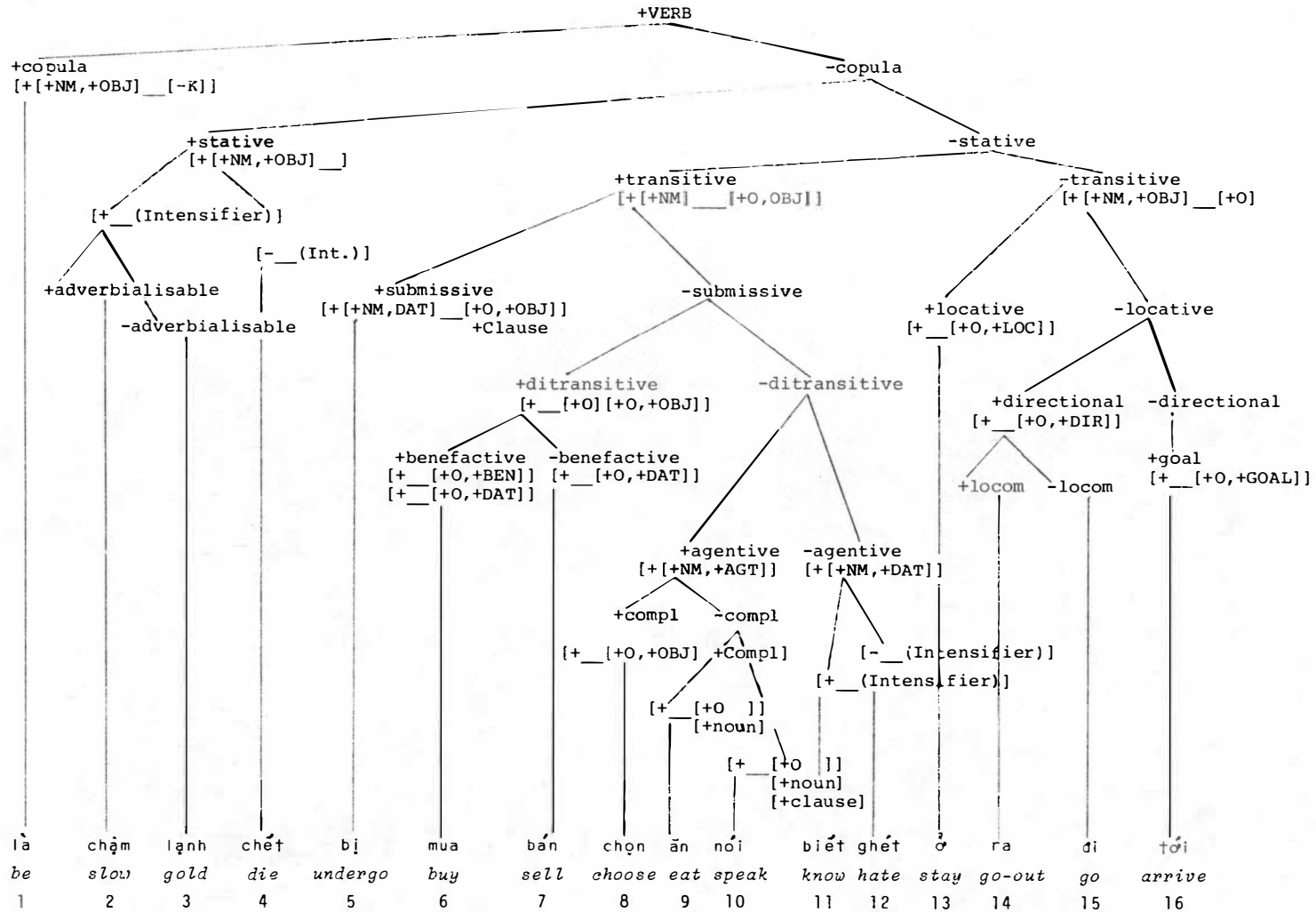
[noun phr.]

[noun phr.]

## 2. chậm 'slow'

The verbs in this class are stative verbs, host a [+NM, +OBJ] slot, can be adverbialised, and can be followed by an Intensifier:

CHART IV  
VERBS



[+[NM, +OBJ] \_\_\_ (Intensifier)]

Ông ấy chậm lắm. (2.2a.)

*he slow very*

*He is very slow.*

[+NM] [Intensifier]  
[+OBJ]

Ông ấy đi chậm.

*He went slowly.*

[+NM] [+MANNER]  
[+OBJ] [adverbialised]

The class of stative verbs presents two learning problems for English speakers. Firstly, they are not introduced by a copula or equational verb like in English, and secondly, they function as adverbs. Note the equivalent English clauses:

*He is very slow.*

[+NM] copula  
[+OBJ]

*He went slowly.*

[+NM] [MANNER]  
[+OBJ] [adverb]

Examples of verbs in the class: *lẹ* 'fast', *giỏi* 'good, well', *dở* 'bad, badly'.

### 3. *lạnh* 'cold'

The verbs in this class are stative verbs, host a [+NM, +OBJ] slot, cannot be adverbialised, and can be followed by an intensifier:

[+[+NM, +OBJ] \_\_\_ (Intensifier)]

Ông ấy lạnh lắm. (2.2b.)

*he cold very*

*He is very cold.*

[+NM] [Intensifier]  
[+OBJ]

Being also stative verbs, these verbs present a learning problem for English speakers because they are not introduced by a copula like their English equivalents are:

*He is very cold.*

[+NM] copula  
[+OBJ]

Examples of verbs in the class: nóng 'hot', vui 'gay', buồn 'sad', đau 'sick'.

#### 4. chết 'to be dead'

The verbs in this class are stative verbs that host a [+NM, +OBJ] slot and that cannot take an intensifier. Incidentally, there are also reasons to consider these verbs intransitive:

[+[+NM, +OBJ] \_\_\_ ]]

Ông ấy chết. (2.2c.)

He died.

[+NM  
+OBJ]

The class of verbs does not present a learning problem since it behaves like its equivalent class of verbs in English.

Examples of verbs in the class: sống 'live', sinh 'be born'.

#### 5. bị 'be adversely affected'

The transitive submissive verbs in the class host [+NM, +DAT] and [+O, +OBJ] slots. The [+O, +OBJ] slot is only filled by a clause:

[+[+NM, +DAT] \_\_\_ [+O, +OBJ]]

Ông ấy bị đi. (3.1a.)

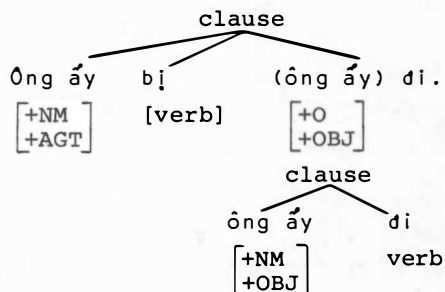
He was forced to go.

[+NM  
+DAT]

[+O  
+OBJ]

[clause]

The phrase structure tree for the clause is as follows, with the Nominative in the embedded clause being deleted obligatorily:



Ông ấy bị họ đánh. (3.1b.)

*he was forced they beat*

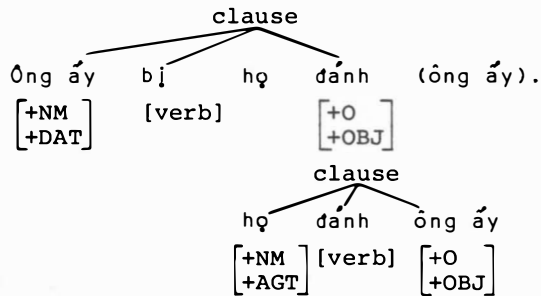
*He was beaten by them.*

[+NM]  
[+DAT]

[+O]  
[+OBJ]

[clause]

The phrase structure tree for the clause is as follows, with the Nominative in the embedded clause being optionally deleted and with the Objective in the embedded clause being obligatorily deleted:



The class of transitive submissive verbs presents a semantic and a structural problem. The semantic problem stems from the fact that the verb utilised always carries an adversative or a benefactive connotation. The structural problem is that of the embedded clause.

The two verbs in the class are: bị 'be adversely affected', and được 'happily experience'. Following are examples with được:

Ông ấy được đi.

*he happily experience go*

*He was allowed to go.*

[+NM]  
[+DAT]

[+O]  
[+OBJ]

Ông ấy được họ khen.

*they praise*

*He was praised by them.*

[+NM]  
[+DAT]

[+O]  
[+OBJ]

It is noted that the obligatorily deleted Objective case form in the embedded clause does not necessarily host an OBJECTIVE case relation as in the above examples. It can also host a BENEFACTIVE case relation:





## 7. bán 'sell'

The class of bi-transitive DATIVE verbs is characterised by the fact that its verbs host [+NM, +AGT], [+O, +DAT], and [+O, +OBJ] slots:

[+[+NM, +AGT] \_\_\_ [+O, +DAT] [+O, +OBJ]]

Ông ấy bán tôi sách. (7)

He sold me books.

|        |        |        |
|--------|--------|--------|
| [+NM]  | [+O]   | [+O]   |
| [+AGT] | [+DAT] | [+OBJ] |

The class of verbs does not present a learning problem since its counterpart in English has the same syntactic features. It is however noted that the DATIVE case relation can be realised by the Dative case form, and yet, is placed before the [+O, +OBJ] tagmeme, a matter that, in English, would depend upon the relative lengths of the two tagmemes:

Ông ấy bán cho tôi sách.

He sold to me books.

|        |        |        |
|--------|--------|--------|
| [+NM]  | [+D]   | [+O]   |
| [+AGT] | [+DAT] | [+OBJ] |

Also, due to phonetic harmony requirements, the shorter of the two post-verbal tagmemes is placed before the longer one:

Ông ấy bán tôi nhiều sách và báo.

He sold me many books and newspapers.

|        |        |        |        |
|--------|--------|--------|--------|
| [+NM]  | [+O]   | [+O]   | [+O]   |
| [+AGT] | [+DAT] | [+OBJ] | [+OBJ] |

Ông ấy bán sách cho Ông Bà Nguyễn-văn-Trương.

He sold books to Mr & Mrs Nguyen-van-Truong.

|        |        |        |
|--------|--------|--------|
| [+NM]  | [+O]   | [+D]   |
| [+AGT] | [+OBJ] | [+DAT] |

Examples of verbs: gởi 'send', chuyển 'transfer', giúp 'help', sang 'transfer'.

## 8. chọn 'choose'

The verbs in this transitive class host [+NM, +AGT] and [+O, +OBJ] tagmemes, and their [+O, +OBJ] tagmeme can be optionally complementalised:

[+[+NM, +AGT] \_\_\_ [+O, +OBJ] (Complement)]

Ông ấy chọn bà ấy làm đại-diện.

He chose her to be representative.

|        |        |               |
|--------|--------|---------------|
| [+NM]  | [+O]   | [+Complement] |
| [+AGT] | [+OBJ] |               |

The relationship between the [+O, +OBJ] tagmeme and its complement is as follows:

bà ậ́y làm đạ̣i-diện.  
*She is representative.*  
 [+NM] copula [-K]  
 [+OBJ]

The class of verbs presents a learning problem because the complement to the OBJECTIVE case must be introduced by a copula in Vietnamese, while the English similar construction does not usually have a copula:

*They elected her representative.*  
 [+NM] [+O] [+Complement]  
 [+AGT] [+OBJ]

Examples of verbs: cử 'to elect', lựa 'choose', bỏ phiếu 'cast vote'.

#### 9. ăn 'eat'

The verbs in this transitive class host [+NM, +AGT] and [+O, +OBJ] tagmemes. The [+O, +OBJ] tagmemic slot can be filled only by a noun phrase and it cannot be complementalised:

[+[+NM, +AGT] \_\_\_ [+O, +OBJ]]  
 Ông ậ́y ăn thậ̣t.  
*He ate meat.*  
 [+NM] [+O]  
 [+AGT] [+OBJ]

The class of verbs does not present any learning problem since there is an equivalent class in English.

Examples of verbs: uống 'drink', hít 'inhale', nhĩ 'secrete', nuốt 'swallow'.

#### 10. nói 'speak'

The verbs in this transitive class host [+NM, +AGT] and [+O, +OBJ] tagmemes. The [+O, +OBJ] slot can be filled by either a noun phrase or a dependent clause introduced by rằng 'that':

[+[+NM, +AGT] \_\_\_ [+O, +OBJ]]  
 [noun phr.]  
 [dep. clause]

Ông ấy nói chuyện này.

*He told this story.*

[+NM  
+AGT]

[+O  
+OBJ]

[noun phr.]

Ông ấy nói rằng Ông Hải đi.

*He said that Mr Hai went.*

[+NM  
+AGT]

[+O  
+OBJ]

[dep. clause]

The class of verbs does not create a learning problem because there exists an equivalent class in English.

Examples of verbs: kể 'recount', khuyên 'advise', bảo 'tell', tin 'inform' (which is different from the same form meaning 'believe').

#### 11. biết 'know'

The verbs in this transitive class host [+NM, +DAT] and [+O, +OBJ] tagmemes. Their [+O, +OBJ] slot can be filled by either a noun phrase or a dependent clause introduced by *rằng* 'that':

[+[NM, +DAT] \_\_\_ [+O, +OBJ]]  
[noun phr.]  
[dep. clause]

Ông ấy biết bà ấy. (3.2a.)

*He knows her.*

[+NM  
+AGT]

[+O  
+OBJ]

[noun phr.]

Ông ấy biết rằng bà ấy đi rồi. (3.2b.)

*He knows that she went already.*

[+NM  
+DAT]

[+O  
+OBJ]

[dep. clause]

The class of verbs does not present a learning problem since there is an equivalent class in English.

Examples of verbs: thấy 'see', nghe 'hear', hiểu 'understand', nghĩ 'think'.

#### 12. ghét 'hate'

The verbs in this transitive class host [+NM, +DAT] and [+O, +OBJ]



Ông ấy sống ở Saigon.  
 He lives in Saigon.  
 [+NM] [+OBJ] [+L] [+LOC]

The verb ở presents a learning problem on the production level in that its LOCATIVE case relation is cast in an Objective case form in Vietnamese whereas in English the same case relation is cast in a Locative case form. It is incidentally noted that Hawaiian English has the same construction as in Vietnamese:

Lucky you live Hawaii.  
 [+NM] [+OBJ] [+O] [+LOC]

#### 14. ra 'go out'

The verbs in this intransitive class are characterised by their being directional and having a [+locomotion] feature. They host [+NM, +OBJ] and [+O, +DIR] tagmemes:

[+[+NM, +OBJ] \_\_\_ [+O, +DIR]]  
 Ông ấy ra Saigon. (2.3b.)  
 He went out to Saigon.  
 [+NM] [+OBJ] [+locom] [+O] [+DIR]

The verbs in this group are also co-verbs. As such, they function as directional prepositions introducing a Directional case relation:

Ông ấy đi ra Saigon.  
 He went out to Saigon.  
 [+NM] [+OBJ] [+Di] [+DIR]

The class of verbs presents a learning problem on the production level in that their DIRECTIONAL case relation is cast in an Objective case form whereas in English the same case relation would be cast in a Directional case form. It is also incidentally noted that such a [+O, +DIR] tagmeme exists in Hawaiian English:

He went Halewa.  
 [+NM] [+OBJ] [+O] [+DIR]

Examples of verbs: vô 'go in', lên 'go up', xuống 'go down', lại 'go back'.

## 15. đi 'go'

The verbs in this intransitive class are characterised by their being directional and having a [-locomotion] feature. They host [+NM, +OBJ] and [+O, +DIR] tagmemes:

[+[+NM, +OBJ] \_\_\_ [+O, +DIR]]

Ông ấy đi Saigon. (2.3c.)

*He went to Saigon.*

[+NM] [-locom] [+O]  
[+OBJ] [+DIR]

The verbs in this group are also co-verbs. As such, they function as a directional preposition introducing a Directional case form:

Ông ấy gửi sách đi Saigon.

*He sent books to Saigon.*

[+NM] [+O] [+Di]  
[+AGT] [+OBJ] [+DIR]

The class of verbs presents a learning problem of the production level in that their DIRECTIONAL case relation is cast in an Objective case form whereas in English the same case relation would be cast in a Directional case form.

## 16. tới 'arrive'

The verbs in this intransitive class are characterised by their having a [+goal] feature. They host [+NM, +OBJ] and [+O, +GOL] tagmemes:

[+[+NM, +OBJ] \_\_\_ [+O, +GOL]]

Ông ấy tới Saigon. (2.3d.)

*He arrived at Saigon.*

[+NM] [+goal] [+O]  
[+OBJ] [+GOL]

The verbs in this group are also co-verbs. As such, they function as a goal preposition introducing a Goal case form:

Ông ấy gửi sách tới Saigon.

*He sent books to Saigon.*

[+NM] [+O] [+GL]  
[+AGT] [+OBJ] [+GOL]

The class of verbs presents a learning problem on the production level in that their GOAL case relation is cast in an Objective case form

whereas in English the same case relation would be cast in a Goal case form, with the rare exception of verbs such as *reach* which host [+NM, +AGT] and [+O, +OBJ] tagmemes:

*He reached his destination.*

|      |  |      |
|------|--|------|
| +NM  |  | +O   |
| +AGT |  | +OBJ |

Examples of verbs: *đến* 'arrive', *về* 'return', *qua* 'go over', *sang* 'go over'.

### 17. *có* 'have'

This verb could be classed in the transitive DATIVE group discussed in 2.2.11. above. Thus, it would host [+NM, +DAT] and [+O, +OBJ] tagmemes:

[+[+NM, +DAT] \_\_\_ [+O, +OBJ]]

*Tôi có sách.*

*I have books.*

|      |  |      |
|------|--|------|
| +NM  |  | +O   |
| +DAT |  | +OBJ |

However, the verb stands out of the group in that it can occur without a Nominative. It means 'there is', 'there are', etc.:

*Có sách ở đây.*

*There are books here.*

|      |  |      |
|------|--|------|
| +O   |  | +L   |
| +OBJ |  | +LOC |

The pattern presents a learning problem for English speakers on the production level. As for the teaching of all the classes of verbs above, substitution drills could be prepared to teach this existential pattern.



## REFERENCES

BECKER, Alton L.

- 1967a *A Generative Description of the English Subject Tagmeme*.  
Ph.D. Dissertation. University of Michigan, Ann Arbor.
- 1967b "Conjoining in a Tagmemic Grammar of English", E.L. Blansitt,  
Jr. (Ed.), *Monograph Series on Languages and Linguistics*,  
No.20: 109-122. Georgetown University Press, Washington,  
D.C.

BREND, Ruth M.

- 1970 "Tagmemic Theory: an Annotated Bibliography", *Journal of  
English Linguistics*, 4:7-47. [Appendix I, 1972, *ibid.*,  
to appear.]
- 1972a *Selected Writings of Kenneth L. Pike*. Mouton, The Hague.  
To appear.
- 1972b *Readings in Tagmemics I*. Bloomington, Ind. To appear.

CHAFE, Wallace L.

- 1970 *Meaning and the Structure of Language*. University of  
Chicago Press, Chicago.

CHOMSKY, Noam A.

- 1965 *Aspects of a Theory of Syntax*. MIT Press, Cambridge, Mass.

CLARK, Marybeth

- 1971a "Passive and Ergative in Vietnamese", *Working Papers in Lin-  
guistics*, 3,8:103-118. University of Hawaii, Honolulu.
- 1971b "Submissive Verbs as Adversatives in Some Asian Languages",  
*Working Papers in Linguistics*, 3,8:119-142. University of  
Hawaii, Honolulu.

forth- *Coverbs and Cases in Vietnamese*. Ph.D. Dissertation, Uni-  
coming versity of Hawaii.

COOK, Walter A.

- 1969 *Introduction to Tagmemic Analysis*. Holt, Rinehart and  
Winston, New York.
- 1970 "Case Grammar: from Roles to Rules", *Languages and Lin-  
guistics Working Papers*, 14-29. Georgetown University Press,  
Washington, D.C.
- 1971a "Case Grammar as Deep Structure in Tagmemic Analysis", *Lan-  
guages and Linguistics Working Papers*, 1-9. Georgetown Uni-  
versity Press, Washington, D.C.
- 1971b "Improvements in Case Grammar", *Languages and Linguistics  
Working Papers*, 10-22. Georgetown University Press,  
Washington, D.C.
- 1972 "A Set of Postulates for Case Grammar Analysis", *Languages  
and Linguistics Working Papers*, 35-49. Georgetown University  
Press, Washington, D.C.

FILLMORE, Charles J.

- 1968 "The Case for Case", in Bach and Harms (Eds.) *Universals in  
Linguistic Theory*. Holt, Rinehart and Winston, New York.
- 1969 "Toward a Theory of Case", in Reibel and Shane (Eds.) *Modern  
Studies in English*. Prentice-Hall, Englewood Cliffs.
- 1971 "Some Problems for Case Grammar", in O'Brien (Ed.) *Monograph  
Series on Languages and Linguistics*, No.24:35-56. Georgetown  
University Press, Washington, D.C.

FRIES, Charles C.

- 1945 *Teaching and Learning English as a Foreign Language*. Uni-  
versity of Michigan Press, Ann Arbor.

FRIES, Peter H.

- 1970 "On Double Function in Tagmemic Analysis", *Anthropological  
Linguistics*, 12:122-135.

HA, Donna

- 1970 "A Severe Case of Vietnamese", University of Hawaii Lin-  
guistics 622 term paper, Honolulu.

LADO, Robert

- 1957 *Linguistics across Cultures*. University of Michigan Press,

Ann Arbor.

LI, Ying-Che

- 1970 *Chinese Grammar*, Ph.D. Dissertation, University of Michigan.  
Published by Seton Hall University Press, South Orange, 1971.

LIEM, Nguyen-Dang

- 1969 *Vietnamese Grammar, a Combined Tagmemic and Transformational Approach*. *Pacific Linguistics*, The Australian National University, Canberra.
- 1971a "Clause Units in English and Southeast Asian Languages in Contrast", in Gerhard Nickel (Ed.) *Papers from the International Symposium on Applied Contrastive Linguistics, Stuttgart, October 11-13, 1971*, 157-174. Cornelsen-Velhagen & Klasing.
- 1971b "Modern Linguistic Theories and Contrastive Linguistics", in Gerhard Nickel and Albert Raasch (Eds.) *Kongressbericht der 3. Jahrestagung der Gesellschaft für angewandte Linguistik*. Julius Groos Verlag, Heidelberg.
- 1972a "Clauses and Cases in Southeast Asian Languages (Burmese, Cambodian, Cantonese, Lao, Thai, and Vietnamese)". Paper presented at the Eleventh International Congress of Linguistics. Bologna and Florence, August 28 - September 2, 1972.
- 1972b "Clauses and Cases in French and English in Contrast", *Third International Congress of Applied Linguistics*. Copenhagen, August 20-26.
- 1973a "Clauses and Cases in Vietnamese", *First International Conference on Austroasiatic Linguistics*. Honolulu, January 2-6.
- 1973b "Cases and Clauses in Southeast Asian Languages (Cambodian, Cantonese, Lao, Thai, and Vietnamese) in Relation to Culture and Thought", *IX International Congress of Anthropological and Ethnological Sciences*. Chicago, August 28 - September 8.

LONGACRE, Robert E.

- 1964 *Grammar Discovery Procedures: a Field Manual*. Mouton, The Hague.

MANLEY, Timothy M.

- 1969 "Verb Phrases in Vietnamese". University of Hawaii Linguistics 650 term paper, Honolulu.

- 1972 *Outline of Sre Structure. Oceanic Linguistics Special Publication No.12.* University of Hawaii Press, Honolulu.
- MILLER, Carolyn P.  
 1966 *Structural Ambiguity within the Vietnamese Relative Clause.* University of North Dakota unpublished M.A. thesis.
- NGUYEN Dinh-Hoa  
 1972 "Passivization in Vietnamese", in Jacqueline M.C. Thomas and Lucien Bernot (Eds.) *Langues et Techniques, Nature et Société*, 179-187. Edition Klincksieck, Paris.
- PIKE, Kenneth L.  
 1954, 1955, 1960 *Language in Relation to a Unified Theory of the Structure of Human Behavior*, volumes I, II, III. Summer Institute of Linguistics, Santa Ana.  
 1966 "A Guide to Publications Related to Tagmemic Theory", in Thomas A. Sebeok (Ed.) *Current Trends in Linguistics III*, 365-394. Mouton, The Hague.  
 1971 "Crucial Questions in the Development of Tagmemics - the Sixties and the Seventies", in O'Brien (Ed.) *Monograph Series on Languages and Linguistics*, No.24:79-98.
- PLATT, Heidi  
 1970 *A Comparative Study of English and German.* Monash University Ph.D. Dissertation, Clayton.
- PLATT, John T.  
 1971 *Grammatical Form and Grammatical Meaning.* North Holland, The Netherlands.
- STAROSTA, Stanley  
 1967 *Sora Syntax, A Generative Approach to a Munda Language.* University of Wisconsin Ph.D. Dissertation, Madison.  
 1971a "Derivation and Case in Sora Verbs", *Indian Linguistics*, 32,3:194-206.  
 1971b "Lexical Derivation in Case Grammar", *Working Papers in Linguistics*, 3,8:83-102. University of Hawaii, Honolulu.  
 1972a "Some Lexical Redundancy Rules for English Nouns", *GLOSSA*, 5,2:167-202.  
 1972b "Case in the Lexicon", paper presented at the Eleventh International Congress of Linguists. Bologna and Florence,

August 28 - September 2, 1972.

- 1973 "The Faces of Case", *Language Sciences*, 25:1-14 (April).
- TAYLOR, Harvey
- 1968 "Vietnamese Syntax, A beginning". University of Hawaii Linguistics 650 term paper, Honolulu.
- 1972 *Case in Japanese*. Seton Hall University Press, South Orange.
- THOMPSON, Laurence C.
- 1965 *A Vietnamese Grammar*. University of Washington Press, Seattle.
- THOMPSON, Laurence C. and David D. Thomas
- 1966 "Vietnam", in Thomas A. Sebeok (Ed.) *Current Trends in Linguistics*, volume II. Mouton, The Hague.
- TRAN Trong Hai
- 1971 "Case and Verbs in Vietnamese". University of Hawaii Linguistics 650 term paper, Honolulu. To appear in *Working Papers in Linguistics*. University of Hawaii, Honolulu.
- YOUNG, Richard E., Alton L. Becker, and Kenneth L. Pike
- 1970 *Rhetoric: Discovery and Change*. Harcourt, Brace and World, New York.

