410
ST92
v.23
1993
cop.2

NOTICE: **Return or renew all Library Materials! The *Minimum* Fee for each Lost Book is $50.00.**

The person charging this material is responsible for its return to the library from which it was withdrawn on or before the **Latest Date** stamped below.

Theft, mutilation, and underlining of books are reasons for disciplinary action and may result in dismissal from the University.
To renew call Telephone Center, 333-8400

UNIVERSITY OF ILLINOIS LIBRARY AT URBANA-CHAMPAIGN

DEC 1 1 2002

MAY 3 0 2003

L161—O-1096

Studies in The Linguistic Sciences

ol. 23, No. 1
Spring 1993

410 STX
S:92
23:1 SPR 1993 COPY 2

v

NIKEN ADISASMITO: Syllable structure and the nature of schwa in Indonesian 1

PAUL AGBEDOR: Verb serialization in Ewe 21

MARTIN J. BAIK and ROSA SHIM: Yes, we have no bananas: English negative tags in cross-linguistic communication 43

ELABBAS BENMAMOUN: The status of agreement and the agreement projection in Arabic 61

MARVIN K. L. CHING: Examining the trustworthiness of the latest *OED* in reflecting current English 73

ABDUL AZIZ DIOP: Language planning across political boundaries: A case study of Pulaar 83

ANDREW TILIMBE KULEMEKA: Bimoraicity in monosyllabic Chichewa ideophones 107

SHARON R. MORRISON: A re-examination of cardinal vowels and auditory equidistance 117

STEVEN SCHÄUFELE: The Vedic clause-initial string and universal grammar 131

Squib

M. LYNNE MURPHY: Discourse markers and sentential syntax 163

Reviews

Anvita Abbi (1992). Reduplication in South Asian languages: An areal, typological, and historical study. (Hans Henrich Hock) 169

Narindar K. Aggarwal (1991). Studies on Nepali language and linguistics: A bibliography. (Mithilesh K. Mishra) 193

John Baldwin & Peter French (1990). Forensic phonetics. (José Ignacio Hualde) 195

Recent Books 197

Contents of volumes 18 - 22 201

UNIVERSITY OF ILLINOIS
LEARNING & LABOR
CHARTERED 1867

Department of Linguistics
University of Illinois

STUDIES IN THE LINGUISTIC SCIENCES

PUBLICATION OF THE DEPARTMENT OF LINGUISTICS
COLLEGE OF LIBERAL ARTS AND SCIENCES
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

EDITOR: Hans Henrich Hock; REVIEW EDITOR: James H. Yoon

EDITORIAL BOARD: Elmer H. Antonsen, Eyamba G. Bokamba, Chin-Chuan Cheng, Jennifer S. Cole, Georgia M. Green, Hans Henrich Hock, Braj B. Kachru, Yamuna Kachru, Chin-W. Kim, Charles W. Kisseberth, Howard Maclay, Jerry L. Morgan, Rajeshwari Pandharipande, James H. Yoon, Ladislav Zgusta, and Alessandro Zucchi

AIM: SLS is intended as a forum for the presentation of the latest original research by the faculty and especially students of the Department of Linguistics, University of Illinois, Urbana-Champaign. Papers by scholars not associated with the University of Illinois are also welcomed.

SPECIAL ISSUES: Since its inception SLS has devoted one issue each year to restricted, specialized topics. A complete list of such special issues is given on the back cover.

EDITORIAL ADDRESS: Editor, Studies in the Linguistic Sciences, Department of Linguistics, University of Illinois, 4088 Foreign Languages Building, 707 South Mathews, Urbana, Illinois 61801; Tel.: 217-333-3563; Fax: 217-333-3466; E-mail: sls@kahane.cogsci.uiuc.edu

BOOKS FOR REVIEW: Review copies of books may be sent to the Editor.

SUBSCRIPTION: Normally, there are two issues per year. Requests for subscriptions should be addressed to SLS Subscriptions, Department of Linguistics, University of Illinois, 4088 Foreign Languages Building, 707 South Mathews, Urbana, Illinois 61801.

UPCOMING ISSUES: Vol. 23:2: Issues in Formal Semantics, Editor: Alessandro Zucchi; Vol. 24:1: Papers in General Linguistics.

Price: $8.00 (per issue)

STUDIES IN THE LINGUISTIC SCIENCES

Papers in General Linguistics

THE LIBRARY OF THE
MAY 2 6 1994
UNIVERSITY OF ILLINOIS

EDITOR

Hans Henrich Hock

EDITORIAL ASSISTANT

Sara L. Michael

VOLUME 23, NUMBER 1
SPRING 1993

DEPARTMENT OF LINGUISTICS, UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS 61801

CONTENTS

Preface v

Niken Adisasmito: Syllable structure and the nature of schwa in Indonesian 1

Paul Agbedor: Verb serialization in Ewe 21

Martin J. Baik and Rosa Shim: Yes, we have no bananas: English negative tags in cross-linguistic communication 43

Elabbas Benmamoun: The status of agreement and the agreement projection in Arabic 61

Marvin K. L. Ching: Examining the trustworthiness of the latest *OED* in reflecting current English 73

Abdul Aziz Diop: Language planning across political boundaries: A case study of Pulaar 83

Andrew Tilimbe Kulemeka: Bimoraicity in monosyllabic Chichewa ideophones 107

Sharon R. Morrison: A re-examination of cardinal vowels and auditory equidistance 117

Steven Schäufele: The Vedic clause-initial string and universal grammar 131

Squib

M. Lynne Murphy: Discourse markers and sentential syntax 163

Reviews

Anvita Abbi (1992). Reduplication in South Asian languages: An areal, typological, and historical study. (Hans Henrich Hock) 169

Narindar K. Aggarwal (1991). Studies on Nepali language and linguistics: A bibliography. (Mithilesh K. Mishra) 193

John Baldwin & Peter French (1990). Forensic phonetics. (José Ignacio Hualde) 195

Recent Books 197

Contents of volumes 18 - 22 201

Preface

Over the past several years the number of submissions from faculty and students not associated with the University of Illinois has greatly increased. While this increase is a welcome sign of the broad recognition that *Studies in the Linguistic Sciences* (*SLS*) has received in the field, it has also brought with it an increase in the time required for the editorial process. I apologize for the resulting delay in this issue of the journal.

To improve communications with outside contributors, this issue adds an e-mail address and a fax number to the complete postal address of the editor. This information is provided on the inside of the front cover.

Manuscripts may be submitted at any time. They should be DOUBLE-SPACED and adhere as closely as possible to the style sheet of *Language*, the journal of the Linguistic Society of America.. Accepted contributions will be returned to authors for any necessary revisions and for final formatting. The final version is to be formatted in accordance with the *SLS* style sheet, which will be mailed along with the acceptance letter, and should be submitted, if possible, both in 'hard copy' and on disk. The disk document should be in the most recent version of Microsoft Word for the Macintosh or IBM, on a 3.5" disk.

As usual, I take great pleasure in acknowledging the help of a number of my colleagues who have refereed submissions for the present issue of *Studies in the Linguistic Sciences*: C.-C. Cheng, Laura Downing, Georgia Green, Braj B. Kachru, Yamuna Kachru, Chin-W. Kim, Rajeshwari Pandharipande, James Yoon, Ladislav Zgusta, and Alessandro Zucchi.

The Department of Linguistics also is grateful for support from the College of Liberal Arts and Sciences toward publishing this issue, and technical support from the Language Learning Laboratory. Last, but not least, I would like to express my appreciation to Beth Creek, Cathy Huffman, and Eileen Sutton from the Department Office, and Sara Michael, my editorial assistant, for their help in preparing this issue.

March 1994 Hans Henrich Hock (Editor)

Studies in the Linguistic Sciences
Volume 23, Number 1, Spring 1993

SYLLABLE STRUCTURE AND THE NATURE OF SCHWA IN INDONESIAN*

Niken Adisasmito

The phonological system of Indonesian allows more than one set of syllabification principles with regards to its treatment of borrowed lexical items. Schwas in borrowed lexical items are derived. In contrast, schwas in the native Indonesian lexical items are present in the underlying representation.

1. Introduction

The treatment of borrowed lexical items in a language has received a lot of discussion, such as by Hyman (1970), Kaye & Nykiel (1979), Silverman (1992), Itô & Mester (1993). In analyzing the phonology of the borrowed lexical items in a language, one needs to take into account the phonological system of the borrowing language. In the process of borrowing, the speaker tries to retain the form of the borrowed lexical items which is closest to that of the source language. At the same time, the (phonological) system of the speaker's language will 'determine' the output forms.

In this paper, I discuss syllable structure and syllabification, as constraints that determine the output of borrowed lexical items in Indonesian. I also discuss the nature of schwa: whether it is part of the underlying representation or whether it is derived, by contrasting native Indonesian lexical items with borrowed ones.

The claims of this paper are (1) that there is more than one set of syllabification principles in Indonesian, and (2) that schwa is underlying in the native Indonesian vocabulary, but inserted in the borrowed vocabulary.

As mentioned in Hardjadibrata 1978, there have been various studies on Indonesian syllable patterns. These studies do not agree on what can constitute a syllable. For instance, according to Hardjadibrata, the 'Committee for Reforms in Indonesian Spelling' (CRIS) claims that the coda of a syllable can consist of a sequence of three consonants, while Kridalaksana lists two consonants as the maximum possible number for a coda. In this paper, I will discuss the possibility for a coda to consist of a sequence of consonants, and the constraints governing the occurrences of such codas.

The data for this paper are based on my intuition as a native speaker of Indonesian. I consulted the dictionary of Echols & Shadily (1989), which has been most helpful in providing the alternations in the written representation of certain lexical items. These orthographic alternations I believe reflect alternations in the pronuncia-

tion of the borrowed lexical items. I further verified my observations regarding these alternations with two native speaker consultants.

2. Theoretical framework

In this paper, the template-based approach to syllabification (Itô 1986) is adopted. According to this approach, the properties of syllable structure fall out from the general prosodic principles. These principles, the settings of which are language-specific, are Maximality, Prosodic Licensing, and Extraprosodicity. The prosodic structure of the syllable is further defined by syllable templates, the Sonority Principle, the Onset Principle, and the Coda Filter.

These principles are sufficient to account for the data under discussion.

3.0 Analysis of the data

In the following sections, I will show the Maximal Syllable Template(s) of Indonesian (section 3.2) and the role of schwa in the syllabification of consonant clusters, based on the data in section 3.1. I further show that schwas are part of the underlying representation in native Indonesian vocabulary (section 3.3). Finally, I propose that there are variations in the syllabification principles among Indonesian speakers; these variations, however, are not mutually contradictory, given certain language-external factors (section 3.4).

3.1 Some relevant facts about Indonesian

Hardjadibrata (1978) argues that there are eleven syllable patterns in Indonesian, as opposed to ten postulated by Kridalaksana, and thirteen by the CRIS. (See Hardjadibrata 1978 for further references.) These differences stem from disagreement on what can constitute a syllable, i.e., whether an onset or a coda can consist of more than one consonant. According to the CRIS, the maximum number of consonants constituting a coda is three. There are, however, only two lexical items (see (1)) with such a coda, found in the Indonesian vocabulary. Ignoring these highly exceptional forms, I show in this section that Hardjadibrata's claim is correct.

(1)	V CCC:	arts (from Dutch)	'physician'
	CV CCC:	kOrps (from Dutch)	'corps'

The syllable patterns in Indonesian are exhibited in (2) - (11) below, involving both native and foreign-origin words.

In (2), it is shown that a syllable may consist of a single vowel, and such a syllable may be word initial or word final.

(2) V pattern

a-nak	'child'
i-kat	'to tie'
o-lah	'to ánalyze'
ba-**u**	'smell'

A syllable may be onset-less, or coda-less, or may consist of a simple onset and a simple coda, as shown in (3) - (5). These syllable patterns can occur at word-initial, word-medial or word-final position. The data in (2) - (5) are representative of the native Indonesian syllable patterns (Grijns 1977).

(3) CV pattern

ku-rus	'skinny'
bi-na-sa	'to die'
sə-ka-raŋ	'now'
la-ma	'long (of time)'

(4) VC pattern

in-dah	'beautiful (for scenery)'
əm-bun	'dew'
ar-ti	'meaning'
ra-**ih**	'to reach'
ha-**us**	'thirsty'
ma-**in**	'to play'

(5) CVC pattern

pas-ti	'certain, exact'
taŋ-kas	'skilled, fast'
pal-su	'fake'
pu-**siŋ**	'headache'
siŋ-gah	'to stop by'

The syllable structure of borrowed lexical items is distinct from that of native ones in that it allows complex onsets and/or complex codas, which may consist of a sequence of two to three consonants (particularly the onsets). A two-consonant onset usually consists of an obstruent (voiced or voiceless) and a liquid (e.g. [tragis]), or a sibilant [s] and a voiceless obstruent (e.g. [spasi], shown in (6) - (9). However, as the alternative forms in (6) - (9) illustrate, there is an alternation in the pronunciation: either the consonants are maintained as a cluster, or a schwa is inserted between them.[1]

(6) CCV pattern

sas-**tra**, sas-tə-ra (Sanskrit)	'literature'
kre-dit, kə-re-dit (Dutch)	'credit'
spe-syal, sə-pe-syal (Dutch)	'special'
gra-tis, gə-ra-tis (Dutch)	'free of charge'
sta-syUn, sə-ta-syUn, sə-ta-si-yUn (Dutch)	'station'
spa-si, sə-pa-si (Dutch)	'space'

(7) CCVC pattern

do-**brak**, do-bə-rak	'to break in'
blun-tas, bə-lun-tas	'a kind of plant'
prak-tis, pə-rak-tis (Dutch)	'practical'
stOp, sə-tOp (Dutch)	'stop'
in-**spɛk**-si, in-sə-pɛk-si (Dutch)	'inspection'

A two-consonant coda may consist of a voiceless obstruent and a sibilant [s], or a liquid and a nasal consonant, or a nasal consonant and a sibilant [s].

(8) (C) CV CC pattern

kOm-**plɛks**, kOm-plɛk, kOm-pə-lɛk (Dutch)	'complex (housing)'
In-**dɛks**, In-dɛk (Dutch)	'index'
hɛlm, hɛ-ləm (Dutch)	'helmet'
flIm, fI-ləm (Dutch)	'film'
in-**tɛrn**, in-tɛ-rən (Dutch)	'internal'
mo-**dɛrn**, mo-dɛ-rən (Dutch)	'modern'
trans-mi-gra-si, tran-sə-mi-gra-si (Dutch)	'transmigration'

(9) V CC pattern

Ons, On	'ounce'
ɛks, ɛk (?)	'ex-'

The alternation of /helm/ and /film/ differs from that of /kompleks/ and /indeks/. In the former case, either the consonants are maintained as a cluster at the coda position, or a schwa is inserted between them. In contrast, the consonant clusters at coda position in /kompleks/ and /indeks/ may be maintained or simplified by deleting the second member of the cluster.

A three-consonant onset consists of a sibilant [s], a voiceless obstruent and a liquid [r], shown in (10) and (11). This kind of onset can occur at word-initial or word-medial position. (The lexical items with this kind of onset are borrowed from Dutch.)

(10) CCCV pattern

stra-tegi, sə-tra-tegi, sə-tə-ra-te-gi	'strategy'
spre, sə-pre, sə-pə-re	'linen'

(11) CCCV C pattern

skrIp-si, sə-krIp-si, sə-kə-rIp-si	'thesis'
strom, sə-trom, sə-tə-rom	'electrical current'
in-**struk**-tur, in-sə-truk-tur, in-sə-tə-ruk-tur	'instructor'

In summary, Indonesian syllable patterns seem to allow up to three consonants in a sequence within a syllable. These consonant clusters show alternations in their pronunciation, i.e. with schwas at word-initial and word-medial position. At word final position, the maximum number of consonants, in general, is two.

3.2 Basic analysis

Based on the syllable patterns and the alternations in the pronunciation of the lexical items, as presented in 3.1, the structure of the syllable in Indonesian is maximally CCCVCC, and minimally V. The latter case indicates that Indonesian syllables require at least a vowel to be well-formed, and that they adopt the 'Relative' Onset Principle.[2] When there is no prevocalic consonant available to be incorporated as an onset, the vowel itself is sufficient to form a

syllable, with or without the presence of a coda. When there is a consonant preceding the vowel, however, the Onset Principle is overridden by the 'Universal Core Syllable Condition' (Itô 1986), shown in (12).

(12) 'Universal Core Syllable Condition'

This condition guarantees that an intervocalic consonant becomes the onset of the following vowel and not the coda of the preceding vowel.

The complex onsets shown in (13), follow the Sonority Principle: The sonority of onsets increases towards the nucleus of the syllable, and the sonority of codas decreases from the nucleus (Selkirk 1982; Steriade 1982). This principle, however, is violated in (14) in that the sonority of the onset decreases towards the nucleus. In the language from which the lexical items in (14) originate, Dutch, complex onsets of this kind are licensed by assigning extraprosodic status to [s] at the edges of the syllable (van der Hulst 1984).

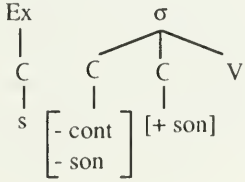
(13)	sas-**tra**	'literature'
	kre-dit	'credit, loan'
	do-**brak**	'to break in'
	blun-tas	'a kind of plant'
(14)	**ska**-la	'scale'
	in-**spɛk**-si	'inspection'
	stri-ka	'iron (from Dutch: strijken 'to iron')
	in-**struk**-tur	'instructor'

The fact that complex onsets are preserved in Indonesian shows that this onset structure is acceptable, at least for some speakers.

The Maximal Syllable Template in Indonesian then, can be characterized as in (15). The condition governing the combinatorial possibility of the consonants is shown in the diagram in (16). Onset sequences of consononants violating the Sonority Principle (as in (14)) are allowed in that the [s] of such structures is extraprosodic; if the sequence of consonants respects the Sonority Principle, then the first segment is an obstruent (voiced or voiceless) and the second a liquid, [l] or [r].

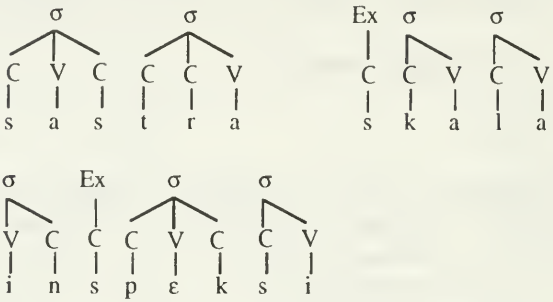
(15) Maximal Syllable 1: CCCVCC

(16) Onset Principle



The Maximal Syllable Template (15), together with the Onset Principle (16), accounts for (13) and (14), as shown in (17).

(17)



The forms in (13) and (14), however, are accepted only by some speakers; for others, the only accepted forms are the alternations shown earlier (repeated in (18)).

(18) **sə-ka**-la
in-sə-pɛk-si
sə-tri-ka
in-sə-truk-tur

For yet other speakers, the complex onsets shown in (13) and (18) are not tolerated either, even though they respect the Sonority Principle. Only the structures in (19) are well-formed for these speakers.

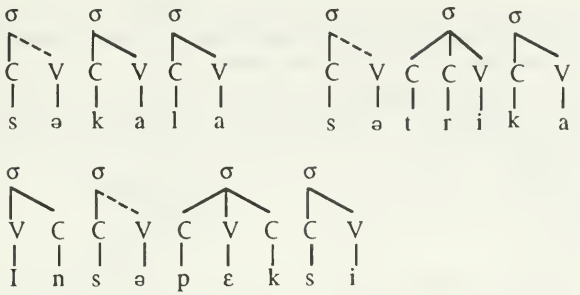
(19) **sə-tə-ri**-ka
in-sə-tə-ruk-tur
gə-ra-tis
pə-rak-tis

This phenomenon suggests that the Maximal Syllable Template previously suggested, viz. CCCVCC,[3] is not observed by all speakers. The Maximal Syllable Template for the forms in (18) is given in (20).

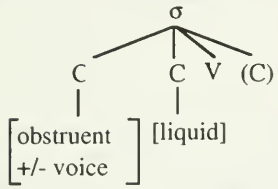
(20) Maximal Syllable 2: CCVC

Comparing the forms in (14) and those in (18), the extraprosodic segments of (14): (as in [skala]), are syllabified in (18) (e.g. [səkala]) in that they become the onset of a syllable followed by an inserted vowel, schwa. The syllabification of some of the forms in (18) is shown in (21). This insertion phenomenon indicates that in this variety of Indonesian extraprosodic elements are not allowed.[4] The Onset Principle (shown in (22)), governing the combinatorial possibility of the consonants thus, determines that if the onset of a syllable includes a sequence of two consonants, then the first segment is an obstruent (voiced or voiceless) and the second a liquid, [l] or [r].

(21)



(22) Onset Principle



The principles of Prosodic Licensing (Itô 1986) require that all phonological units are licensed, either by syllabification or by extraprosody. If a segment at a sylllable edge is not extraprosodic and it is not syllabified, it will be deleted by Stray Erasure. A vowel slot is inserted to syllabify the [s], which otherwise would become a stray segment and would have to be deleted. The melodic material for the vowel is schwa, as in. Yoruba (Abaglo & Archangeli 1989) and Yawelmani (Archangeli 1984).

The data in (19), [gəratis], [pəraktis], etc., show that in this variety of Indonesian the maximal syllable is CVC (23), implying that no tautosyllabic consonant cluster is allowed (24).[5]

(23) Syllable Maximal Template: C V C

(24) Onset Principle

* σ [C C V

To license the unsyllabified segments, a schwa is inserted, as shown in (25).The syllable template shown in (23) is consistent with Grijns's (1977) claim that the native Indonesian syllable pattern consists maximally of a simple onset, a nucleus, and a simple coda.

(25)

The fact that there are various syllable templates existing simultaneously in the language suggests that different speakers have

different parameter settings for the onset condition. In the following paragraph, I show that the Coda Condition also seems to have different parameter settings for different speakers.

Complex codas in borrowed lexical items are maintained by some speakers, but not by others (cf. (8); the relevant information is repeated in (26)).

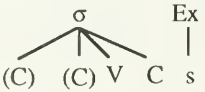
(26)	kOm-**plɛk s**	vs.	kOm-**plɛk**
	hɛl m	vs.	**hɛ-ləm**
	mo-**dɛr n**	vs.	mo-**dɛ-rən**

Kager (1990) claims that in Dutch, a schwa-like vowel, with much shorter duration (relative to underlying schwa),[6] is inserted when consonant clusters are at syllable final position. The pronunciation of /helm/, as he indicates, is [hɛl^m], in which [^] signifies the inserted short schwa-like vowel. The presence of the schwa-like vowel may have been interpreted as a full schwa in the borrowing process of these lexical items — /helm/, /modern/ — into Indonesian. However, there is a clear difference, at least to some speakers, between Dutch [hɛl^m] and Indonesian [hɛləm]. This difference indicates that complex codas are not tolerated in the speech of some speakers. The alternation in the pronunciation of /kompleks/, i.e. [kOmplɛks] and [kOmplɛk] also shows this intolerance of complex codas by some speakers.

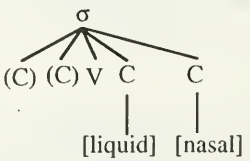
The Coda Condition which allows complex codas follows that of the source language (exhibited in (27)), i.e., the [s] is licensed as extraprosodic, and a nasal consonant following a liquid consonant is incorporated into the syllable forming a complex coda — since the resulting consonant cluster still observes the Sonority Principle (van der Hulst 1984).

(27) Coda Condition 1:[7]

a) extraprosodic [s]

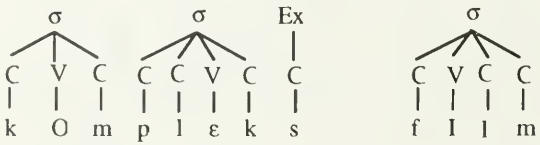


b) liquid-nasal cluster



The effect of the Coda Condition in (27) is exemplified in (28).

(28)

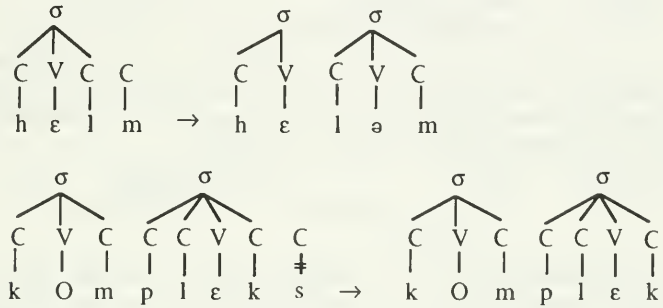


When no complex codas are tolerated, the Coda Condition licenses only a single post-vocalic consonant to be a coda (29). In

accordance with the principles of Prosodic Licensing, the unsyllabified consonant is either syllabified forming a new syllable, or the consonant remains and is deleted by Stray Erasure. When new syllable formation occurs, a schwa is inserted before the nasal and, observing the 'Universal Core Syllable Condition' in (12), the immediately preceding consonant ([l] in /film/, which now becomes intervocalic) is incorporated into the new syllable as its onset. Stray Erasure deletes the extraprosodic [s]. The syllabification is shown in (30).

(29) Coda Condition 2:
* V CC] σ

(30) Syllabification of (28) according to Coda Condition 2:[8]



The conditions set forth in (27) and (29) suggest that Indonesian allows different settings for its Coda Condition in the face of borrowed lexical items, i.e., one setting for complex codas, and another for simple codas, in addition to allowing different settings for its Onset Principle.

In this section, I have discussed the maximal syllable templates, as well as the language-specific parameter settings for the Onset Principles and the Coda Conditions of Indonesian. I also have shown that some schwas are derived in Indonesian, as the result of the application of the syllabification principles to borrowed lexical items.

3.3 The nature of schwa in Indonesian

There have been two opposing claims about the nature of schwa in Indonesian: One is that schwa is derived, due to its predictable distribution (Cohn 1989; Grijns 1976); the other is that it is present in the underlying representation (Lapoliwa 1981). In this section I present evidence that schwa in the native vocabulary is in fact present underlyingly.

The syllabification of borrowed lexical items discussed in section 3.2 supports the argument that the distribution of schwa is predictable: Whenever a consonant cluster is not tolerated, a schwa is inserted to simplify the cluster. This argument seems to also be supported by the native Indonesian vocabulary, as shown by the

example in (31). In these data, the Maximal Syllable Template is CCVC, and the consonant clusters at the onset position, respecting the Sonority Principle, consist of an obstruent (voiced or voiceless) and a liquid ([l] or [r]). An alternative pronunciation is shown in (32), where schwa appears to be inserted in order to break up the consonant cluster, and therefore, appears to be predictable (as argued by Cohn and Grijns). This alternative pronunciation, which is OPTIONAL, is usually determined by language-external factors, namely the formality of context and the rate of speech. For further discussion on language-external factors, see section 3.4.

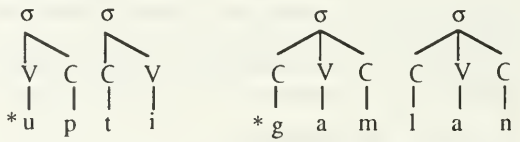
(31)			(32)	
	braŋ-kat	'to go'		bə-raŋ-kat
	cla-ka	'bad luck'		cə-la-ka
	blun-tas	'kind of plant'		bə-lun-tas
	krOn-cOŋ	'type of music genre'		kə-rOn-cOŋ

The predictable occurrence of schwa just shown, however, does not account for the presence of schwas in (33). If the presence of schwas in (33) were predictable, the forms in (34) should be the underlying representations of those in (33).

(33)			(34)	
	u-pə-ti	'tax'		*up-ti
	ga-mə-lan	'Javanese musical instrument'		*gam-lan

The application of the syllabification principles set forth in section 3.2, however, does not result in the surface forms, as demonstrated in (35). The main reason is that in these data all segments are exhaustively syllabified. Thus, there is no motivation for the schwa to be inserted here, since these forms are well-formed by the syllabic template proposed. One could conclude that the schwas in (33) must be underlying, and therefore that NOT all schwas are derived in the language.

(35)



However, one could try to salvage the argument that all schwas are derived by proposing a filter blocking the structures in (35) for other, non-templatic reasons. One may argue that Indonesian has an adjacency restriction on consonants that two adjacent stops are ruled out (hence **upti*), and that a (labial) nasal cannot be adjacent to a liquid [l] (hence **gamlan)*, even when they are not tautosyllabic.

But these filters are not supported by other data, as shown in (36). The data in (36) include forms where adjacent stop consonants and adjacent nasal/liquid sequences are well-formed.

(36) sak-ti 'invincible' jum-lah 'total of addition'
sap-tu 'Saturday' kim-lo 'type of dish'
buk-ti 'evidence' im-lah 'dictation'

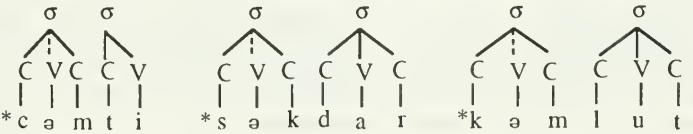
Since the adjacency restrictions which were proposed to salvage the schwa-as-derived argument are themselves not valid, it does not seem possible to salvage this line of reasoning. Since schwa is therefore not derived in (33), it must be underlying.

Forms shown in (37) provide further evidence that some schwas MUST be underlying. If all the schwas in (37) were derived, then the forms in the right column without a schwa should be the underlying representations.

(37) cə-mə-ti ← cmti 'whip'
sə-kə-jap ← skjap 'in an instant'
kə-mə-lUt ← kmlUt 'chaos'

Assuming that these are the correct underlying forms, applying the syllabification principles argued for in 3.2 would result in incorrect surface forms, as demonstrated in (38). Given that the language tolerates codas, there is no reason why, for example, **cəmti* should require further syllabification as *cəməti*.

(38)



Since the correct distribution of schwa cannot be achieved based on the underlying forms proposed in (37) where schwa is not present underlyingly, these forms cannot be correct. Therefore it must be the case that these schwas are present underlyingly, as shown in (39).

(39)

The data in (33) and (37) have shown that schwa cannot be derived in at least these forms. There are three possibilities for the distribution of schwa in the native vocabulary: (1) all schwas are derived; (2) some schwas are underlying, while others are derived; (3) all schwas are underlying. Since the data above have indicated that (1) is not true, the choice is between (2) and (3). On theoretical grounds, it is undesirable that some schwas should be underlying and others derived, since there is no phonotactic evidence that they are distinct.

I propose, therefore, that ALL schwas in the native Indonesian vocabulary are part of the underlying representation. Under this view, the forms in (32): *bəraŋkat*, *cəlaka*, etc., are the underlying representations of the forms in (31): *braŋkat*, *claka*, etc., and not vice versa, as proposed for the schwa-as-derived approach above.

This, in turn, indicates that a schwa may optionally be deleted, provided that the resulting consonant cluster respects the Sonority Principle. This is the case with the forms in (31), repeated in (40), but is not the case in (41), where deletion of schwa results in consonant clusters which violate the Sonority Principle. Therefore the application of schwa deletion is ruled out.

(40) bəraŋkat → braŋkat
bəluntas → bluntas

(41) jəmbatan → *jmbatan 'bridge'
cəndawan → *cndawan 'mushroom'

In conclusion, schwas in the native Indonesian vocabulary are present underlyingly, and are optionally deleted,[9] if and only if the resulting forms respect the Sonority Principle.

3.4 Variations of the UR in Indonesian

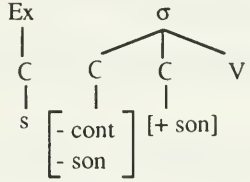
In this section, I propose that Indonesian allows more than one set of syllabification principles to exist simultaneously. I also discuss another dimension of the pronunciation alternations of borrowed lexical items, language-external factors and their interaction with the different sets of syllabification principles.

As discussed in section 3.2, it seems that Indonesian allows different maximal syllable templates, Onset Principles, and Coda Conditions with regards to borrowed lexical items. Based on this discussion, I divide Indonesian into different varieties, which I call VARIETY A, VARIETY B, and VARIETY C.

VARIETY A represents that form of Indonesian which allows complex onsets and complex codas in its Maximal Syllable Template: CCCVCC (shown earlier in (15)). The combinatorial condition governing the occurrences of the consonant clusters is determined by the Onset Principle shown in (16) and the Coda Condition 1 shown in (27). The Onset Principle determines that a sequence of consonants is allowed to occur at the onset position, in that if the first segment is an [s] and if the following segment is less sonorous than [s], then [s] is extraprosodic; and if the sequence of consonants respects the Sonority Principle, then it is a sequence of an obstruent and a liquid. The Coda Condition determines that (1) if [s] is the second member of a consonant cluster at coda position, it is extraprosodic, and that (2) a nasal consonant following a liquid consonant at coda position is tautosyllabic. The syllabification principles of VARIETY A are shown in (42).

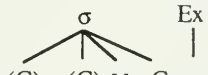
(42) Syllabification principles of VARIETY A

- Maximal Syllable Template: C C C V C C
- Onset Principle:

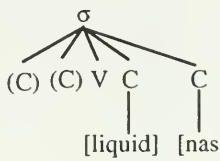


- Coda Condition:

extraprosodic [s]



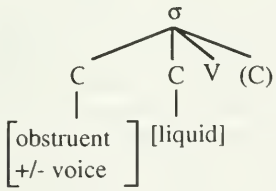
liquid-nasal cluster



VARIETY B allows complex onsets in its Maximal Syllable Template: CCVC (shown earlier in (18)). The consonant sequence allowed to occupy the onset position is an obstruent and a liquid (shown in (22)). The governing coda condition is Coda Condition 2, shown in (29), in which no consonant clusters are allowed. The syllabification principles characterizing VARIETY B are shown in (43).

(43) Syllabification principles of VARIETY B

- Maximal Syllable Template: C C V C
- Onset Principle:



- Coda Condition: * VCC]σ

VARIETY C has the 'simplest' Maximal Syllable Template relative to the other two varieties: CVC. It has Coda Condition 2 governing its syllable structure. The syllabification principles of VARIETY C are shown in (44).

(44) Syllabification principles of VARIETY C

- Maximal Syllable Template: CVC
- Onset Principle: * σ [C C V
- Coda Condition: * V C C] σ

These three varieties are by no means mutually exclusive. Some speakers may choose to use one variety for one kind of situation and another for a different situation. What this means is that depending on the situation, a speaker may apply different syllabification principles with regards to borrowed lexical items.

In the following paragraphs, I discuss the kinds of situations which determine the contexts where speakers apply different syllabification principles.

Grijns (1977) observes that in the Indonesian dialect of Jakarta[10] there are variations in the pronunciation of borrowed lexical items, particularly those involving consonant clusters. The factors determining these variations seem to be the rate of speech, the register used (formal or informal), and to what extent the speaker is exposed to the source language. This last factor is also discussed by Onn (1976) for Johore Malay, where the constraint on the consonant clusters *st-*, *sk-*, *sm-* is violated when the speakers are familiar with the source language, or when the speakers '... aspire to [a] more "sophisticated" type of speech ...' (58).

In my own speech, with regards to the lexical items previously discussed, the less formal the context of the conversation is, the more likely it is that the schwa will be heard. This is also supported by my consultants in that, in our conversation, the schwas in some of the lexical items shown earlier were consistently present, e.g. [fIlləm], [sətOp], [sətrom],[11] [modɛrən], [səkrIpsi], [sətrika], etc.

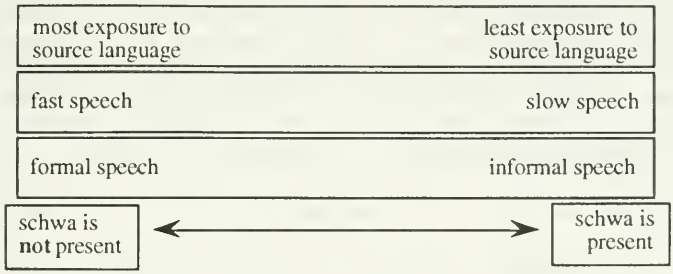
The rate of speech also proves to be significant in the occurrence of schwas in Indonesian, as I observed from my own speech, as well as that of my consultants. The faster the speech, the more likely the schwa is to be dropped. This phenomenon is also observed by Grijns, based on his survey of the *lenong* show which involves reading a dramatic text in very fast and informal speech and in which, speaking loudly seems to be necessary (Grijns 1977). Speaking loudly seems to have the same effect as speaking slowly in deliberate speech, as noted by Grijns, in that the occurrence of consonant clusters is low, i.e., the occurrence of schwa is high.

It seems to be the case for Johore Malay, for the Indonesian dialect of Jakarta, and for Indonesian in general, that the more the speakers are familiar with borrowed lexical items, the more likely the consonant clusters are to be retained.

Based on Grijns' observation of the three factors, the diagram shown in (45) shows the LIKELINESS of the presence of schwa in the pronunciation of borrowed lexical items.

The likeliness of the presence of schwa in this case is to be understood in conjunction with the syllabification principles relevant to each variety. For instance, speakers of VARIETY A — assuming that the source language substantially influences them, in that they are

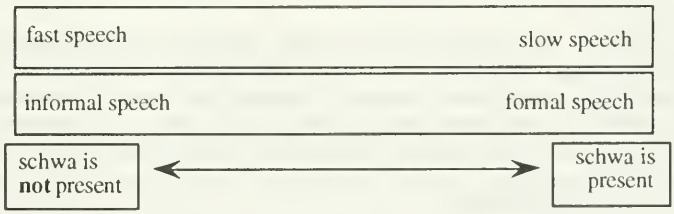
(45) Diagram of [ə] occurrence in borrowed lexical items



able to use the most complex syllable template — may use the alternative form closest to that of the source language during the formal speech. Thus, no schwa is present (e.g. [struktur]). In contrast, in the same formal context, a speaker of VARIETY B or VARIETY C will use the form with a schwa inserted, because the Onset Principle in these two varieties does not permit a segment to be extraprosodic. The result is the surface forms [sətruktur] and [sətəruktur], respectively.

In contrast to the conditions for the occurrence of schwa in BORROWED lexical items, the schwa in the NATIVE vocabulary is more likely to be present in slow speech, and when the context is formal (one obviously has maximum exposure to one's native language), as indicated in diagram (46). Thus, for the forms shown earlier in (31) and (32), [bəraŋkat] is more likely to be used on a formal occasion, as well as in slow speech, and [braŋkat] is used under informal and/or fast-speech conditions.

(46) Diagram of the [ə] occurrence in native lexical items



4.0 Conclusion

I have argued that the phonological system of Indonesian allows more than one set of syllabification principles to exist simultaneously in the borrowing process of lexical items from another language. The different syllabification principles have different maximal syllable templates, and different settings of the Onset Principle and the Coda Condition. I have also pointed out that the varieties characterized by the three sets of syllabification principles do not necessarily identify distinct speech communities, but may be used by a single speaker under different conditions. Schwas are derived in borrowed lexical items, as a means to incorporate these words into Indonesian (i.e. in

the syllabification of consonant clusters). They are, however, present in the underlying representation in native Indonesian vocabulary.

Further investigation is certainly needed to determine which other phonological and morphological constraints borrowed lexical items are subjected to in Indonesian. For instance, the complex affixation system in Indonesian may provide evidence for whether or not borrowed lexical items behave the way native ones do.

NOTES

* I am grateful to Jennifer Cole for her supervision during the initial writing process of this paper. Many thanks also go to Abby Cohn and Laura Downing for their suggestions and comments, some of which are incorporated in this paper, as well as to other professors and colleagues in the department of Linguistics who discussed the ideas with me. I would also like to thank Johannes Kabu and Herry Sutanto for their help in verifying the data. Of course, all errors are solely my responsibility.

[1] Donors of borrowed lexical items in Indonesian are Dutch, English, Arabic/Persian, Chinese, and Portuguese. It is coincidental that the majority of the data in this paper are of Dutch origin. For discussion on borrowed lexical items in Indonesian, see Spitzbardt 1970, Lowenberg 1983, Verhaar 1984, de Vries 1988, etc.

[2] In contrast, a language may choose to require the 'Strict Onset Principle', in that a well-formed syllable must have an onset (Itô 1989).

[3] At this particular point, the Coda Condition in Indonesian is not yet discussed.

[4] It seems to be a strain for some speakers to pronounce word-medial consonant clusters, as in *instruksi*, *transmigrasi*, etc. Sometimes the [s] is completely deleted, resulting in *intruksi*, *tranmigrasi*, etc. A lot of speakers are quite familiar with these two words due to the socio-political context. The [s] deletion in this case may indicate a tendency in which [s] in the middle of consonant clusters: *-nstr-*, *-nsm-*, is not syllabifiable.

[5] Even though schwa insertion is obligatory for the speakers who accepts only the forms in (19), it is never obligatorily absent for other speakers. Its absence, however, may be preferable in formal situations. See section 3.4. for discussion on language-external factors.

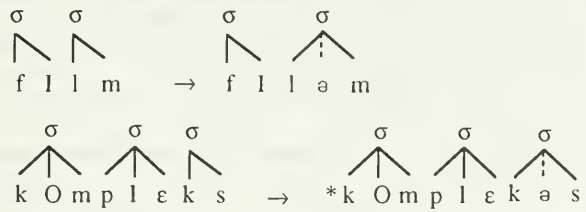
[6] The occurrence of full schwas in the final syllable of Dutch words is shown in the following (Kager 1990):

kolOnə	'column'	rItmə	'rhythm'
Ortnər	'file'	kataloxəs	'catalogs'

[7] It may be argued that the Maximal Syllable Template in (18) allows complex codas: CCVCC, governed by this Coda Condition. While this may be true in Indonesian, further research needs to be done. For this paper, I assume that Coda Condition 1 governs the Maximal Syllable Template in (15): CCCVCC, and not in (20): CCVC.

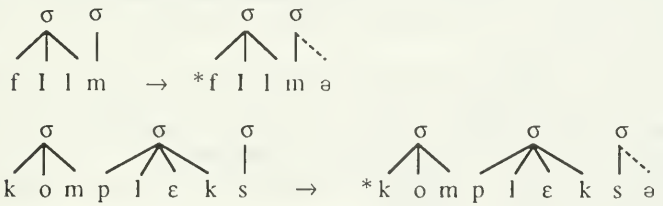
[8] The syllabification of /film/ and /kompleks/ seems to be problematic. Taking into account the directionality of, say, CCVC-syllable template mapping, if the syllabification is R→L, the resulting surface form is (i), in which case the surface form of /kompleks/ is ill-formed.

(i)



If the syllabification is L→R, as in (ii), the surface forms of both /film/ and /kompleks/ are incorrect.

(ii)



A possibility to account for this phenomenon is to assume that [s] in /kompleks/ is not present in the underlying representation of some speakers, as suggested by Cohn (personal communication). This suggests that the underlying representation may vary among different speakers, as observed for Hindi by Ohala (1974).

[9] Cohn (personal communication) has suggested that while this is true historically, there has been reanalysis synchronically: many of the optional deletion cases have been reanalyzed as insertion cases, although she points out that this needs to be tested through psycholinguistic experiments.

[10] This particular dialect of Indonesian is influenced by Dutch, Chinese, Javanese, Sundanese, and various regional languages of Indonesia. This is partly due to Jakarta being the administrative center.

[11] All consultants felt that it would be more natural to pronounce [sətOp] or [sətrom] in general, even though they could pronounce those two words as [stOp] and [strom], when asked. This may show that monosyllabic words are avoided, when possible, in Indonesian. Grijns notes that, according to M. Zain's *Kamus moderen*

Bahasa Indonesia, 93% of Indonesian words are bi- and trisyllabic, and only 7% are monosyllabic or longer than three syllables.

REFERENCES

ABAGLO, P., & D. Archangeli. 1989. Language particular underspecification: Gengbe /e/ and Yoruba /i/. Linguistic Inquiry 20.457-480.

ARCHANGELI, Diana. 1984. Underspecification in Yawelmani phonology and morphology. MIT Ph.D. dissertation in Linguistics.

COHN, Abigail. 1989. Stress in Indonesian and bracketing paradoxes. Natural Language and Linguistic Theory 7.167-216.

DE VRIES, J. W. 1988. Dutch loanwords in Indonesian. International Journal of Society and Languages 73.121-136.

ECHOLS, John, & Hassan Shadily. 1989. Kamus Indonesia Inggris. Jakarta: PT Gramedia.

GRIJNS, C. D. 1977. Jakartan speech and Takdir Alisjahbana's plea for the simple Indonesian word form. Studies in Austronesian linguistics, ed. by S. A. Wurm, 1-34. Ann Arbor: University of Michigan.

HARDJADIBRATA, R. 1978. Consonant clusters in Indonesian. Second International Conference on Austronesian Linguistics: Proceedings, ed. by S. A. Wurm, 165-180. Canberra: The Australian National University.

HYMAN, Larry. 1970. The role of borrowing in the justification of phonological grammars. Studies in African Linguistics 1:1.1-48.

ITO, Junko. 1986. Syllable theory in prosodic phonology. University of Massachusetts Ph.D. dissertation in Linguistics.

-----. 1989. A prosodic theory of epenthesis. Natural Language and Linguistic Theory 7.217-259.

-----, & Armin Mester. 1993. Japanese phonology constraint domains and structure preservation. To appear in: A handbook of phonological theory, ed. by John Goldsmith. Blackwell: Handbooks in Linguistics Series.

KAGER, René. 1990. Dutch schwa in moraic phonology. Papers from the 26th Regional Meeting of the Chicago Linguistic Society 2: The Parasession on the Syllable in Phonetics and Phonology, 241-255.

KAYE, Jonathan, & Barbara Nykiel. 1979. Loan words and abstract phonotactic constraints. The Canadian Journal of Linguistics 24:2.71-93.

LAPOLIWA, Hans. 1981. A generative approach to the phonology of Bahasa Indonesia. (Pacific Linguistics: Series D, No. 34.) Canberra.

LOWENBERG, Peter H. 1983. Lexical modernization in Bahasa Indonesia: functional allocation and variation in borrowing. Studies in the Linguistic Sciences 13:2.73-86.

OHALA, Manjari. 1974. The abstractness controversy: experimental input from Hindi. Language 50.225-235.

ONN, Farid Mohamed. 1976. Aspects of Malay phonology and morphology. University of Illinois at Urbana-Champaign Ph.D. dissertation in Linguistics.

SELKIRK, E. 1982. The syllable. The structure of phonological representations, ed. by H. van der Hulst & N. Smith, 337-383. Dordrecht: Foris Publications.

SILVERMAN, Daniel. 1992. Multiple scansions in loanword phonology: evidence from Cantonese. Phonology 9.289-328.

SPITZBARDT, H. 1970. Sanskrit loan words in the Bahasa Indonesia. Beiträge zur Linguistik und Informationsverarbeitung 9.62-79.

STERIADE, Donca. 1982. Greek prosodies and the nature of syllabification. MIT Ph.D. dissertation in Linguistics.

VAN DER HULST, Harry. 1984. Syllable structure and stress in Dutch. Cinnaminson, N.J.: Foris Publications.

VERHAAR, John W. M. (ed.). 1984. Towards a description of contemporary Indonesian: preliminary studies, part II, in NUSA: Linguistic Studies of Indonesian and Other Languages in Indonesian 19.

Studies in the Linguistic Sciences
Volume 23, Number 1, Spring 1993

VERB SERIALIZATION IN EWE

Paul Agbedor
Department of Linguistics, University of Victoria

This paper examines serial verbal constructions in Ewe, a Kwa language of West Africa, in the light of the recent treatment of such phenomena in Yoruba by Baker (1989). Some of the structural characteristics of serial verbal constructions (SVCs) are examined. The treatment of SVCs in Yoruba is discussed. Some problems with the framework suggested by Baker are pointed out. The main area of concern in this paper is the object sharing phenomenon, which Baker assumes to be obligatory for 'true' SVCs. In this paper, I argue that what Baker calls 'covert coordination' constructions are, in fact, 'true' SVCs in Ewe. As a result, the object sharing phenomenon is not always obligatory. The result is that the syntactic framework proposed for Yoruba would not work for Ewe. An alternative proposal is made to address the discrepancy. The idea is to have a framework that can account for both the object-sharing or 'argument-sharing' and non-argument-sharing types of SVCs.

1.0 Introduction

Serial verbal constructions (SVCs) are a phenomenon commonly found in the Kwa languages of West Africa and Caribbean Creoles. They are also reported in Chinese (Li & Thompson 1973) and in Burmese (Matisoff 1973).[1] In this paper, I will examine SVCs in Ewe (a Kwa language of Ghana) in the light of the principles of Government and Binding Theory. The paper will specifically examine the syntactic framework proposed by Baker (1989), and the problems with that framework will be outlined. I will argue, for instance, that the object-sharing phenomenon, which Baker (1989) suggests is obligatory for 'true' SVCs, is not found in all cases (at least for Ewe). I will also argue that certain SVC structures in Ewe pose a problem for Baker's model, with respect to the Projection Principle, and I will suggest an alternative framework for SVCs. The paper is organized as follows: In the next section, I discuss the major characteristics of SVCs that distinguish them from other structures in Ewe. In section 2.0, I discuss Baker's proposal in some detail, noting the problems with his framework in section 4.0. I show that Baker's proposal does not account for all the SVC types in Ewe, and in section 5.0, I make an alternative proposal.

2.0 Characteristics of SVCs

One of the early linguists who hinted at the notion of SVCs was Westermann, in his 1930 Ewe grammar:

> A peculiarity of Ewe is that we often find a row of verbs one after the other. The chief features of this are that all the verbs stand next to each other without being connected, that all have the same tense or mood, and that in the event of their having a common subject and object, these stand with the first, the others remaining bare (1930:126).

Baker (1989) describes SVCs as a construction in which a sequence of verbs appears in what seems to be a single clause (513). According to him, there is usually one tense/aspect specification for the whole chain of verbs. The verbs in a SVC are also believed to have a single structural subject and they share logical arguments. The following are some examples of SVCs.

(1) Yoruba: Aje sunkun lo ile
Aje weep go home
'Aje wept on his way home.' (Awoyale 1989)

(2) Haitian: Emil pran liv la bay Mari
E. take book DET give M.
'Emil gave the book to Mary.' (Dechaine 1988)

(3) Sranan: Kofi naki Amba kill
Kofi hit Amba kill
'Kofi struck Amba dead.' (Baker 1989)

(4) Akan: Kofi tɔɔ bayire dii
Kofi bought yam ate
'Kofi bought yam and ate.' (Campbell 1991)

(5) Ewe: Kofi ɖa nu ɖu
Kofi cook thing eat
'Kofi cooked and ate.'

One powerful test that has been developed over the years to distinguish SVCs from coordinate and purposive constructions is WH-extraction. If the NP argument of a verb in a SVC can be extracted by WH-movement, it follows that the structure cannot be a coordination or an embedded purpose or result clause (see Ross 1967). This follows from the *Coordinate Structure Constraint*, which restrains extraction from either conjunct of a coordinate construction. Consider the following examples in Ewe.

(6) (a) Kofi ɖa nu ɖu
Kofi cook thing eat
'Kofi cooked and ate.'
(b) Nuka Kofi ɖa ɖu?
Thing-which K. cook eat
'What did Kofi cook and eat?'

(7) (a) Kofi šle agbale na Ama
Kofi buy book give Ama
'Kofi bought a book for Ama.'
(b) Nuka Kofi šle na Ama?
Thing-which K buy give Ama
'What did Kofi buy for Ama?'

In (6) and (7), the (b) examples are Wh-extractions of the NP arguments. In (7), there is an extra argument because of the presence of the 3-place predicate verb *na* (give). All the three NP arguments can undergo WH-extraction, showing that they are not coordinate structures.

The verbs in a SVC form a complex VP with a single-event interpretation. One test that proves this for Ewe is negation. In Ewe, the negative marker is a discontinuous element, *me ... o*, and the negated constituent lies between the two elements. The first element is considered to be the head, and is always attached to the verb. If the verbs in a SVC were to have multiple-event interpretation, we would expect the negative marker to be attached to each verb that represents a separate event. This is found not to be the case for Ewe SVCs, showing that they represent single events. Consider the ungrammaticality of the (c) and (d) examples in (8).

(8) (a) Mešle agbale na Ama
1sg. buy book give Ama
'I bought a book for Ama'
(b) Nye mešle agbale na Ama o
1sg. NEG-buy book give Ama NEG.
'I did not buy a book for Ama'
(c) *Nye mešle agbale me-na Ama o
1sg. NEG-buy book NEG-give Ama NEG.
(d) *Mešle agbale me-na Ama o
1sg. buy book NEG-give Ama NEG.

3. Syntactic representation of SVCs

One of the most challenging phenomena in SVC analysis, according to Baker (1989), is the notion of object sharing. Stewart (1963) in analyzing Twi SVCs suggested that SVCs formed out of a sequence of two transitive verbs show an object deletion under identity (i.e. one of the objects is deleted (normally that of V2) because it is identical to the object of V1)). So for example, the Ewe example in (9) below will be derived from (10) by the traditional generative transformational rule of *Equi-NP Deletion*.

(9) Kofi šo Ama wu
Kofi beat Ama kill
'Kofi beat Ama to death.'

(10) Kofi šo Ama wu Ama
Kofi beat NP kill NP

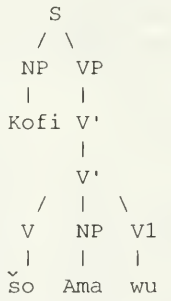
With the demise of transformations as they prevailed in early generative transformational grammar, one option that readily comes to mind is to posit a D-structure in which the two verbs in the SVC take an object NP either to their right or left, depending on the type of language.

(11) [$_{VP}$ [$_{V1}$ beat [$_{V2}$ kill [$_{NP}$ Ama]]]].

In the above structure, only the V2 directly theta-marks the NP. The V1 does not case-mark the NP because the adjacency condition is violated. One alternative is to move the NP to a position between the two verbs. In this case the V1 directly theta-marks the NP. But how can we account for the theta-marking and case-marking properties of the V2, which is a transitive verb, and selects an NP object to which it may assign case?

Baker (1989) suggests a framework for analyzing SVCs in Yoruba. He proposes that the NP that comes between the two verbs in the SVC is a shared object, in that it occupies a position which is theta-marked by both verbs (or their projections). Thus under Baker's analysis, (11) above will be assigned the structure in (12).

(12)



Baker assumes that SVCs are double-headed — that the serial verbs jointly constitute a complex predicate. In the above structure, therefore, the VP is double-headed and the NP it contains is governed by both verbs. From the structure in (12) above, the theta-marking of the NP within the VP by the V1 is straightforward, but the notion of V2 also theta-marking the same NP might seem doubtful. To account for this, Baker invokes the standard conditions on theta role assignment from Chomsky (1986), which are stated as follows.

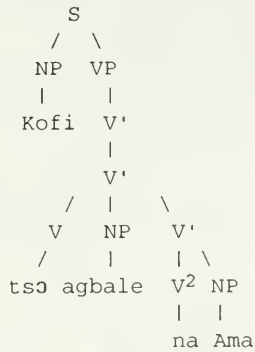
(13) α may theta-mark β iff:
(a) α and β are structural sisters.
(b) a projection of α is a structural sister of β

Condition (a) allows for theta-marking of the NP by V1 while condition (b) allows for theta-marking of the NP by V2, whose projection is a structural sister to the NP. Under Baker's analysis, theta-marking of the external argument is achieved by invoking Williams's (1984) notion that the external (argument) role of the verb

percolates to its maximal projection. Since VP in the structure in (12) is the maximal projection of both V1 and V2, the external theta role of both verbs percolate up to it, where they are assigned to the subject by condition (13b) (Baker 1989:520). So the lexical theta role assignment properties of both verbs are satisfied and the Projection Principle is obeyed. In the rest of this section I examine how Baker's framework can handle the various types of SVCs in Ewe.

The example in (12) above involves two transitive verbs with a shared NP. Each of the verbs is a 2-place predicate. There are examples in which one of the verbs is triadic (i.e. a 3-place predicate). This type can fit into Baker's framework. We only have to expand the last V' into a V and NP as in (14).

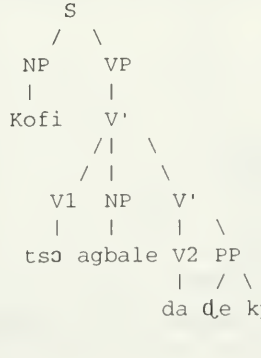
(14)



In (14), the shared NP is *agbale* 'book'; and the V2, a triadic *na* 'give', takes an additional argument *Ama*. V1 theta-marks the NP *agbale* by (13a) and V2 theta-marks the same NP by (13b). The V2 theta-marks its additional argument by (13a).

There is a type of SVC which differs from the one in (14) in that the V2 takes a PP complement instead of the NP in (14). This type is also successfully represented in the framework under discussion, as shown in (15).

(15)



Now consider the sentence in (16) below:

(16) Adela da tu wu xevi
Hunter shoot gun kill bird
'The hunter shot and killed a bird'

In the SVCs represented so far, the two verbs have a common NP object which is theta-marked by both verbs. The verbs assign the same theta role to the NP. But in (16), the NP 'gun' between the two verbs receives two different theta roles from the two verbs THEME. from V1 and INSTRUMENT from V2, under the standard conditions of theta-role assignment outlined in (13). Despite this assignment of different theta roles to the NP object, the sentence fits well into the syntactic frame we are considering. It is possible for the NP to receive two different theta roles, provided it is the same structural position that is involved (cf. Baker 1989:521).

One characteristic of SVCs that comes out clearly at this stage is that they share at least one argument, and that this argument is not always the grammatical object of both verbs.

A class of verbs in Ewe and Yoruba (Baker 1989) which raises some questions about their role in SVCs, is the so-called 'bimorphemic' verbs which are made up of a bound verb and a noun complement referred to as bound verb complement (BVC). For example:

(17) ɖa nu
cook thing
'cook'

(18) no tsi
drink water
'drink'

(19) dzi ha
sing song
'sing'

(20) kpa ha
compose song
'compose'

The controversy is whether the BVCs are syntactic objects of the verbs or whether they compound lexically with the verb root to form true intransitives. I suggest that the BVCs are syntactic objects to the verbs concerned in (17) - (20) above. Consider the following sentences.

(21) E-ɖa nu ɖu
3sg-cook thing eat
He cooked and ate'

(22) E-kpa ha dzi
3sg.-compose song sing
'He composed a song and sang'

(23) E-ku tsi no
3sg.–fetch water drink
'He fetched water and drank'

In the above examples (all grammatical and acceptable), the V2s have no object after them. Like most of the SVCs we have examined already, these examples have NPs which are theta-marked by both

verbs. If the verbs in (17) - (20) are true intransitives (as Baker assumes for Yoruba), then they should not share an argument, and we expect forms like (24) and (25).

(24) *E-ɖa nu: ɖu nu.
He-cook thing eat thing

(25) *E-ku tsi no tsi
He-fetched water drink water

But these forms are not grammatical in Ewe. These ungrammatical sentences in (24) and (25) are quite different from the sentence in (16), where we have two verbs and two NP arguments. In (16), the two verbs share an argument, 'gun', but while this argument is the grammatical object of the first verb, it is not the object of the second. The second verb has its own object, *xevi*. In the ungrammatical (24) and (25), the two verbs share the same object, and since the first occurrence of the object is theta-marked by both verbs, there is no need for an overt NP object for the second verb. Moreover, other nouns can be substituted for the BVCs. Consider (26) and (27).

(26) E-ɖa te ɖu
3sg.-cook yam eat
'He cooked yam and ate'

(27) E-no aha mu
3sg.-drink wine be drunk
'He got drunk by drinking wine'

In the above examples, the nouns act as complements to the so-called bound verbs in grammatical SVCs. If other NPs can be complements of the bound verbs in question, then the BVCs in (17) - (20) are syntactic objects of the bound verbs. I, therefore suggest that the verb roots in (17) - (20) are real transitive verbs that subcategorize for NP objects, just like any other transitive verb in the language. It seems the syntactic properties of the types of verbs found in (17) - (20) in Ewe are similar to the corresponding examples in Yoruba, in which *momi* and *jeun* are both bimorphemic verbs made up of two morphemes as follows:

(28) (a) mu + omi ----> momi
drink water 'drink'
(b) je + oun ------> jeun
eat something 'eat'

The Yoruba sentence in (29) is ungrammatical, just like the Ewe (25):

(29) *Mo bu omi mumi
I pour water drink water
'I poured water and drank'

There are, however, some verb-noun pairs which can be said to form true intransitives in Ewe. These are shown in (30) - (33).

(30) ku dzi
kill heart
'annoy'

(31) šu du
run race
'run'

(32) ve dɔme
hurt stomach
'annoy'

(33) tsi megbe
remain behind
'be late'

This class of verb-complement pairs consists of more or less idiomatic expressions that behave like single verbs and true intransitives, in that no NP can occur in place of the BVCs in these examples. As intransitives they cannot take direct objects. They can, however, take other verbs in SVCs.

(34) Devi-a šu du dzo
child-the run race go
'The child ran away.'

(35) Nufiala tsi megbe va suku
Teacher remain behind come school
'The teacher was late to school.'

In these examples, the complements (BVCs) of the first verbs (V1) are not shared by the second verbs (V2), because the V1 in each case is made up of the 'bound' verb root and the complement (BVC) to become an intransitive verb.

Another class of verbs in some African languages (including Ewe and Yoruba (cf. Awoyale 1987)) is the class of morphologically complex transitive verbs which have been treated as some kind of serial verbs (Bamgboṣe 1982) or as a distinct class of their own and referred to as 'splitting verbs' (Awolobuyi 1969; quoted in Awoyale 1988:21). Ewe has such examples as:

(36) xɔ se
receive hear
'believe'

(37) ɖe fia
remove show
'introduce'

(38) bia se
ask hear
'inquire'

(39) ɖɔ kpɔ
taste see
'taste'

The term 'splitting' is applied here to refer to the fact that the two verbs forming the complex can be 'split' by an intervening NP object. Those who hold this view regard the verb pairs as single lexical items. Others like Bamgboṣe (1982)[2] regard them as relatively 'frozen serial collocations'. I hold the latter view that these verb pairs form serial verb strings. They behave just like other serial verbs. If they are, then it follows that the two verbal elements forming the pair are syntactic lexical heads forming a complex VP. To show that they are lexical heads in a complex VP, we subject these pairs in SVCs to the tense/aspect test. Since TENSE or ASPECT is a functional element which is always attached to the verb in Ewe, I assume that

this can serve as an indicator of the headedness of the verbs in the SVC.

(40) (a) Kofi ɖa nu ɖu
K cook thing eat
'Kofi cooked and ate.'
(b) Kofi a-ɖa nu a-ɖu
K FUT-cook thing FUT-eat
'Kofi will cook and eat.'

(41) (a) Kofi xɔ nya la se
K receive word the hear
'Kofi believed the message.'
(b) Kofi a-xɔ nya la a-se
K FUT-receive word the FUT-hear
Kofi will believe the message.'

As the (b) examples in (40) - (41) show, each of the verbs receives a FUTURE marker, proving that it forms a double-headed predicate. Example (40) involves the normal transitive verb while (41) involves the 'splitting' verbs under discussion. These 'splitting' verbs, though they are not single syntactic units, form single semantic units. They are a kind of fixed collocations, because the two verbs forming the pair in each case have a fixed semantic interpretation.

4.0 The problem

The examples of Ewe SVCs examined so far seem to be adequately accounted for by Baker's model. All involve transitive verbs Moreover, note that the V1 takes only one argument. Now let us consider cases in which the V1 takes an extra argument.

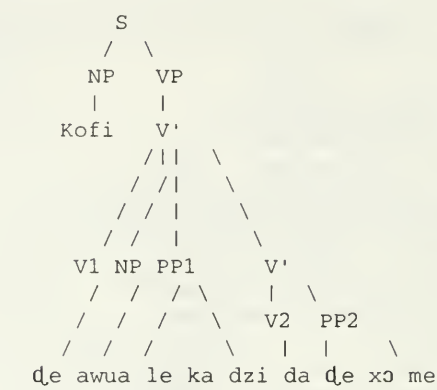
(42) Kofi ɖe awua le ka dzi da ɖe xɔ me
Kofi / remove / shirt / on / rope / top / put / LOC / room / in
Kofi removed the shirt from the line and put it in the room.'

(43) Kofi fɔ agbaleawo le xɔa me da ɖe gota
Kofi / collect / book-the-PL / in / room / put / LOC / outside
'Kofi collected the books from the room and put them outside.'

In the two examples in (42) and (43), the V1 in each case assigns an additional theta role to the PP. Applying Baker's model to these examples, we would expect that the V2 would theta-mark the PP argument of V1.

In the representation in (44), a projection of V2 is a sister to both the NP and the PP1. Note that Baker (1989) claims that the sharing of the NP by the two verbs is obligatory (527). That is, the two verbs should theta-mark the NP between them. Since the PP1 in (44) above is an argument of V1 and a structural sister to the projection of V2, we would expect that V2 should theta-mark the PP too; but it does not. This constitutes a violation of the Projection

(44)



Principle and the standard conditions on theta-role assignment adopted by Baker. At the same time, the two sentences in (42) and (43) are not coordinate structures because they pass the Wh-extraction test. The following examples show that it is possible to extract an NP from any of the conjuncts.

(42') Nuka Kofi ɖe le ka dzi da ɖe xɔ me?
Thing-what / Kofi / remove / from / line / put / in / room
'What did Kofi remove from the line and put in the room?'

Now let us consider examples of SVCs involving a transitive verb and an intransitive or two intransitives.

(45) Kofi no tsi ku
K drink water die
'Kofi died by drinking water.'

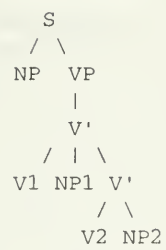
(46) Kofi tutu ɖevia dze anyi
K push child-the fall down
'Kofi pushed the child and fell down.'

(47) Xevia dzò dzó
Bird fly go
'The bird flew away.'

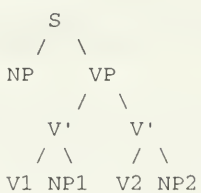
In all the examples in (45) - (47), the V2 is intransitive and object-sharing does not apply. To account for these examples, Baker suggests that where the two verbs theta-mark the NP between them, the structure in (48a) is projected. The structure in (48b) which Baker terms 'covert coordination' is projected where only the first verb theta-marks the NP.

According to Baker, the Projection Principle forces the V2 to theta-mark the NP1 in (48a) above, but this requirement would not hold for (48b). He suggests that the V2 would not be able to theta-mark the NP1 in the configuration in (48b) because the NP is not sister to V2 or any of its projections. He views the so-called 'overt coordinations' as a sequence of distinct events, whereas the true SVC is

(48) (a)



(b)



perceived as a single event. This suggests that 'covert coordination' types are not true SVCs.

This position, however, is not acceptable (at least for Ewe). The Ewe examples in (45) - (47) above all pass the WH-extraction and negation tests for SVCs:

(49) (a) Kofi no tsi ku
K drink water die
'Kofi died by drinking water.'

(b) Nuka Kofi no ku
What K drink die?
'What did Kofi drink and die?'

(c) Kofi me no tsi ku o
K NEG drink water die NEG
'Kofi did not die by drinking water.'

(50) (a) Kofi šo ɖevia si
K beat child-the flee
'Kofi beat the child and fled.'

(b) Ameka Kofi šo si?
Person-which K beat flee
'Who did Kofi beat and flee?'

(c) Kofi me šo ɖevia si o
K NEG beat child-the flee NEG
'Kofi did not beat the child and flee.'

Since the sentences in (45) - (47) also pass the Wh-extraction test, it follows that they are not coordinate structures, but true SVCs.

One other problem has to do with subject-sharing by the verbs in the SVC. As noted earlier, SVCs have the characteristic of having the same subject. While this is true for Ewe, Baker (1989) reports something different for Yoruba. The following example is taken from Baker (1989:529).

(51) Olu ti omo na a subu
Olu push child the fall
'Olu pushed the child down.'

This example is accounted for by Baker's model as far as object-sharing is concerned (i.e. the V2 theta-marks the NP between it and the V1). But instead of the two verbs sharing the subject 'Olu', the object of the first verb becomes the subject of V2. The verb 'fall'

takes 'child' as NP within its second V-bar projection as its only argument (it is unaccusative).

(52)

(Baker, 1989:530)

Baker suggests that this structure has two consequences for the intransitive V2:

(a) V2 must be lexically capable of assigning an internal theta role;

(b) its theta role must be assigned to the object of V1 rather than to the subject of V1.

From this example, Baker predicts that only an unaccusative type of intransitive verb can follow a transitive or unaccusative in a true SVC (i.e. V2 must always be unaccusative if it is intransitive). This prediction is proved wrong by example (50a) above, where the V2 'flee' is unergative and has only an external argument.

In the Ewe example in (46) which is similar to the Yoruba example in (51), the V2 does not theta-mark the NP object of V1, and the sentence has only one interpretation (i.e. Kofi pushed the child and Kofi fell down, not the boy, as in the Yoruba example). In the Ewe example, therefore, the subject-sharing phenomenon is preserved. For an Ewe equivalent to the Yoruba example in (51) (with the interpretation given by Baker), a pronominal third person singular coindexed with the object NP will have to precede the V2, as in the example below.[3]

(53) Kofi tutu ɖevia wò dze anyi
K push child-the 3sg. fall down
'Kofi pushed the child down.'

So we see examples in which no object-sharing takes place but the sentences are true SVCs. I, therefore, propose that all SVCs, whether object-sharing or non-object-sharing, should be regarded as true SVCs. This claim rejects the distinction drawn between true SVCs and 'covert coordination'.

One other problematic type of SVC for Baker is the one involving an unergative and an unaccusative. An example is (47), repeated below.

(54) Xevia dzò dzó
bird-the fly go
'The bird flew away.'

As Baker observed in a note, this particular type of SVC raises a problem for the theta-criterion. 'Bird' receives theta-roles in two different positions: one as the subject of 'fly' and one as the object of 'go'. It has been observed that the same structural position can receive more than one theta-role. But the case under examination involves two different structural positions, external argument position of 'fly' and internal argument position of 'go'. In a structure involving only an unaccusative, the base-generated object which is in a non-case assigning position has to move to subject position to receive NOMINATIVE case in order to satisfy the case-filter.

(55) [bird [v go t]]

But when the verb 'go' is combined with the unergative 'fly', a problem arises: The subject position is filled by the external argument of the unergative. How then do we account for the theta-role assignment by the unaccusative verb?

5.0 An alternative analysis

In the preceding section, I have tried to unearth various problems posed by some Ewe SVCs for the framework of Baker 1989. In this section, I will make an alternative proposal in an attempt to address the problems raised in the preceding section.

The main issue so far is the argument sharing of the verbs in the SVC and the projection of this phenomenon in the tree. The standard conditions on theta-marking in (13) were used to account for this phenomenon. Now, it has been shown that certain SVCs in Ewe do not display object or argument sharing (see examples (45) - (47)). There are two possible ways to account for these non-argument-sharing SVCs: one is to revise the conditions on theta-marking in (13); the other is to modify the syntactic projection of SVCs. Whatever option is chosen affects the other in a way. To modify the projection would lead us back to the structure in (48b) in which the V2 is not able to theta-mark the NP1. This would account for the non-argument-sharing SVCs, but may not account for the argument-sharing ones. Moreover, the structure in (48b) looks more like a coordinate structure, since there are two V1 heads at the same level. I therefore would resort to the first option (i.e., revise the conditions on theta-marking) and modify the projection proposed by Baker a little.

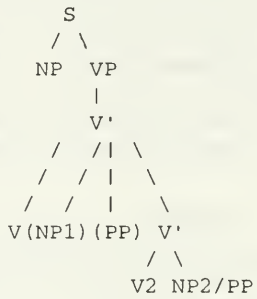
I adopt a rather strict version of the conditions for theta-marking by doing away with condition (b) in (13). The revised condition is stated as follows:

(13') α may theta-mark β iff α is a structural sister of β.

The above condition would prevent V2 from theta-marking NP1 in the tree below. However, while these conditions can account for non-

argument-sharing examples like (45) - (47), it would be difficult to account for SVCs like (6a) or (9), which are of the argument sharing type. To account for these types as well, I propose that the NP2 projection in (56) below is either empty or overt, depending on the type of SVC. NP1 and PP are made optional in (56) by putting them in parenthesis. This is to make it possible to use the same framework for verbs that do not select an NP or PP.

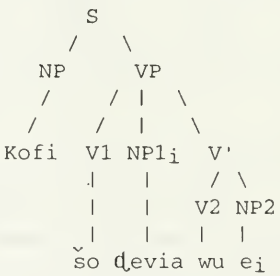
(56)



In argument-sharing types, NP2 would be null and coindexed with the NP1 object of V1. We noted earlier that in SVCs where the V1 takes an additional (PP) argument, the V2 is not able to theta-mark this argument thus violating the Projection Principle under Baker's proposal. This suggests that the object (argument) sharing phenomenon should be projected in a different way than Baker suggested. Applying a typical example like (57a), we shall have the structure in (57b).

(57) (a) Kofi šo ɖevia wu
K. beat child the kill
'Kofi beat the child to death.'

(b)



In the above tree, V1 theta-marks only NP1 by the revised conditions on theta-marking. V2, being a transitive verb, selects an object NP, which is identical to that of V1. NP2 is therefore rendered null and coindexed with NP1. Note that in this particular SVC, NP2 cannot be overt. It can only be overt in coordinate structures, in which there must be a pronominal subject for V2 and a coordinating conjunction as in:

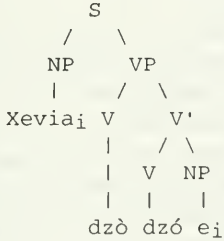
(57a') Kofi šo ɖevia eye wò wu-i
Kofi beat child-the and 3sg. kill-3sg(OBJ)
'Kofi beat the child and he killed him.'

My proposal, I believe, would avoid the three major problems of Baker's proposal, namely:

(i) the inability of V2 to theta-mark a PP argument of V1 even though the argument position satisfies the condition for theta-marking proposed by Baker;
(ii) non-sharing of NP between V1 and V2 in some SVCs;
(iii) the problem of the 'fly go' type.

In my proposal, the V2 would not have to theta-mark the object of V1, since V2 would have its own object position projected. This solves the second problem automatically; that is, the verbs do not have to share NP objects. Concerning the 'fly go' type, my proposal does not involve movement of the internal argument of V2 to subject position as is normally proposed for unaccusatives, since that position would already be filled by the external argument of V1.

(54')



So the internal argument position of V2 is projected as an empty category co-referential with the subject of the sentence. In (57a), the object of V2 is understood. It is, therefore natural to assume that there is an empty category in the object position of V2. The question then arises as to what type of empty category the null object is. It is assumed that there are four types of empty categories, on the basis of the two binary features [± anaphor] and [± pronominal] (cf. Chomsky 1982).

(58) (a) PRO: [+ anaphor, + pronominal]
(b) pro: [- anaphor, + pronominal]
(c) WH-trace: [- anaphor, - pronominal]
(d) NP-trace: [+ anaphor, - pronominal]

PRO is ruled out as a probable candidate, since it is said to be ungoverned at S-structure. The object position in (57b) is governed by the verb, since that position must be case-marked.

NP-trace is also not a possible candidate because NP-movement involves movement from a theta-position to a theta-bar position. The null object in (57b) is in a theta position, and its antecedent is also in a theta position. Therefore, it cannot be NP-trace. However, the null object under discussion seems to share a property with NP-traces. An

NP-trace is subject to Principle A of the Binding Theory which says that an anaphor must be bound in its governing category. The governing category for the null object in (62b) is the entire clause, and the null object is coindexed with an NP within its governing category. So in a way, the null object has something in common with an NP-trace, i.e., they both occur in A-positions. But they differ in their case-marking properties.[4]

WH-trace is also ruled out because it must be bound by an antecedent in an A-bar position. In the examples under discussion, there is no A-bar binder for the trace. This leaves us with *pro*. Supposing we assume at this stage that the null object is *pro*. This raises two questions:

(a) What are the conditions that formally license the *pro*?
(b) How is the content of *pro* determined or recovered?

To answer these questions, let us look at the proposal by Rizzi (1986) that *pro* is subject to two requirements, formulated in what is termed the 'pro-drop parameter'.

(59) The *pro*-drop parameter:
(a) *pro* is governed by X^0
(b) Let X be the licensing head of an occurrence of *pro*; then *pro* has the grammatical specification of the features on X coindexed with it.

Condition (a) can be satisfied in (57b); the object position is licensed by the verb *wu* 'kill'. The problem is with condition (b). In Italian, a typical *pro*-drop language, the content of *pro* in subject position is recoverable from the rich morphology of the verb (i.e. from strong agreement features). But for *pro* in object position, Rizzi (1986) suggests a distinction between English and Italian in the way null objects are licensed. He claims that an occurrence of *pro* in a verb-governed position is allowed in Italian but not in English. He argues that in Italian, the understood object is syntactically 'active' in that it can act as a controller, as a binder, and as a subject of predication for adjunct and small clauses, whereas the null object in English appears to be syntactically 'inert' in the same environment (502). Compare the following sentences from English and Italian.

(60) (a) This leads people [PRO to conclude what follows]
(b) *This leads [PRO to conclude what follows]

(61) (a) Questo conduce la gente alla seguente conclusione
(b) Questo conduce — alla seguente conclusione
(Rizzi, 1986:503)

In (60b), we find that we cannot delete the object controller, whereas in (61b) we can. This suggests that in object-control structures in English, the object NP controller must be overtly represented. Ewe follows English in this respect. Consider the following.

(62) (a) Esia nana amewo susuna be nuwuwua ɖo
This make-HAB people think-HAB that end-the arrive
'This makes people think that the end is near.'
(b) *Esia nana — susuna be nuwuwua ɖo

In (62b) the object cannot be deleted. Rizzi (1986) also points out that in Italian, argument small clauses selected by causative verbs can take null subjects having the same interpretive and formal properties as the null objects.

(63) (a) Questa musica rende [— allegri]
This music renders — happy [+pl]
(b) Certe medicine rendono [— piu intelligenti/calmi]
Certain medicines render — more intelligent/calm [+pl]

The English glosses in (63) are ungrammatical. The missing null object in the Italian examples must be present in the English glosses to be grammatical in English. Ewe behaves just like English in this respect too.

(64) (a) Atike aɖe-wo wɔa [ame drɔzii]
medicine INDE.-pl make person weak
'Some drugs make people weak.'
(b) *Atike aɖewo wɔa — drɔzii
'Some drugs make weak.'

(64b) is ungrammatical because of the missing small clause subject.

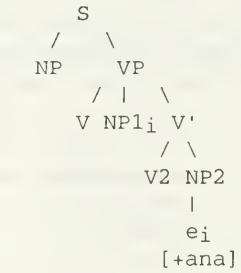
Evidence adduced so far points to the fact that the null object being proposed for Ewe SVC may not be *pro*. In fact the discussion so far suggests that there is no structural NP position. But as has been pointed out, the failure of Baker's model to satisfy the Projection Principle suggest that the object sharing phenomenon should be projected in a different way. One plausible way is to project an empty NP object for V2, an NP which will be co-indexed with the NP object of V1. Moreover, *pro* as a pronominal must be free in its governing category (i.e. it is subject to Principle B of the Binding Theory). The null object being proposed here is quite different, in that it is co-indexed with an NP within its governing category. It will also be shown that this null object is bound by the NP object of V1. Also, there is no evidence that Ewe allows object *pro* independently of SVCs. Therefore, the null object in question cannot be *pro*.

Raposo (1986) proposed for European Portuguese that the empty category in object position is a variable.[5] According to Principle C of the binding theory (Chomsky 1981), a variable, like other referring expressions, cannot be coreferential with a c-commanding nominal occurring in an argument position. This is because variables, like other referring expressions cannot be A-bound. Pronominals, on the other hand, are not subject to Principle C and can, therefore, be coreferential with a c-commanding argument (as long as these arguments do not occur in the governing category (GC) of the

pronominal). This rules out the null object being a variable and brings us back to a point mentioned earlier.[6]

We noted earlier that the null object being proposed shares a characteristic with NP-traces, but that it cannot be NP-trace because it is in a case-marked and a theta position. The shared characteristic is that the null object is bound in its governing category (i.e., it has as its antecedent an argument in its GC). This position satisfies the condition for an anaphor. I, therefore, propose that the null object in the Ewe SVC structure in (57b) is an 'empty anaphor' (cf. Saxon 1989, 1990; Chung 1989). We will have a base-generated NP 'empty anaphor' as the object of V2, and this would be coindexed with the NP object of V1. This gives us the structure in (65).

(65)



Since this empty category is an anaphor, it must obey Principle A of the binding theory. We shall now explore the conditions on this binding principle and see how far the structure in (65) fits into it.

(66) **Binding Theory:**.
Principle A
An anaphor must be bound in its governing category.

(67) *A-binding*
α binds β iff
(i) α is in an A-position;
(ii) α c-commands β;
(iii) α and β are co-indexed.

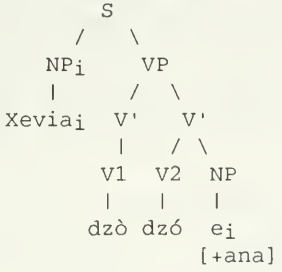
(66) involves two notions, 'binding' and 'governing category'. (67) outlines the conditions for binding. The first is that α must be in an A-position. In (65), the NP binder is in an A-position. The second condition is that α must c-command β. Here, I will adopt the revised version of c-command which is also known as m-command, and which is stated in Sells 1985:39 as follows:

(68) *C-command (revised definition)*
α c-commands β iff every maximal projection dominating α also dominates β

Under this interpretation of c-command, the governing category of the empty anaphor being proposed will be the entire clause. This empty anaphor will be bound by the NP which is within the GC, thus satisfying Principle A in (66).

Now let us consider the other problematic type of SVC (the one involving unergative and unaccusative verbs). I propose the same base-generated 'empty anaphor' as above for the argument of V2.

(69)



As noted earlier, the only argument of the V2 is internal, and if this verb occurs alone in a structure, then the D-structure argument moves to subject position to receive NOMINATIVE case. But here it is combined with an unergative, whose only argument is external. Therefore, when these two verbs combine, the subject position is already filled by the external argument of the V1 and, therefore, the object of V2 cannot move there to receive case. So it must be empty, and since it is coreferential with the subject, it must be an anaphor. Its GC is the entire clause. The subject 'bird' m-commands the empty anaphor and they are co-indexed.

One question that needs to be addressed is whether the empty anaphor, *ana*, needs case. I assume that *ana*, being the object of a transitive verb, must be case-marked. The position is governed and theta-marked and, therefore, nothing prevents the empty anaphor from receiving case. Moreover, WH-traces and *pro* are case-marked. So we can say that the empty anaphor is case-marked. The case on the empty anaphor will make the theta position visible and allow the predicate to assign its theta role.[7]

6.0 Conclusion

I have been discussing Ewe serial verbal constructions within a framework suggested by Baker (1989). I noted that the Ewe data pose certain problems for Baker's framework, especially with respect to the Projection Principle and the notion of object-sharing. Baker's distinction between a true and non-true SVCs has been rejected (at least for Ewe). An alternative proposal has been made for handling SVCs in Ewe. It was found necessary to modify the conditions on theta-assignment and the syntactic projection of SVCs to make room for both object-sharing and non-object-sharing types. The notion of 'empty anaphor' was introduced as the null object of V2 in the SVC. Object-sharing is achieved by coindexation of the empty anaphor with the object of V1. This proposal has to be tested against other SVC languages to determine its universality. Moreover, the notion of

empty anaphor or 'Little *ana*' (Saxon 1989, 1990) is quite new and needs to be tested by further research.

Despite the extensive work on SVCs in the past twenty years, there are still more questions than answers. The issue of what constitutes a true SVC is not clear. Also pertinent is the question of what constitutes a main verb in SVCs. Should there even be a main verb and a subordinate one? What principles determine the order of verbs in SVCs? These and other questions need to be addressed in future research.

NOTES

[1] Reported in Awoyale 1988.

[2] Cited in Awoyale 1988.

[3] There seems to be a pragmatic issue here. The pronominal third person that distinguishes the two sentences in (51) and (53) is always coreferential with the immediately preceding NP. It is the subject of the embedded clause and, being a pronominal, it should be free in its governing category. The antecedent of this pronominal is the object of the matrix verb, which is outside the GC of the pronominal. The V2 *dze anyi* seems to be a fixed collocation forming an intransitive verb. The issue as to whether it is an unaccusative is not clear, and I do not intend going into that here.

[4] This issue will be taken up again later in this section.

[5] Reported in Cole 1987:597.

[6] It may be possible to say that the null object is a variable left by an empty operator. But assuming that an operator moves only into [Spec, CP], the issue will be where to locate the operator in the tree in (57b) for example. Carstens (1988) is reported to have taken on that issue (see Baker 1989).

[7] Case assignment for the empty anaphor in (69) is, however, problematic. It may be possible to propose that the empty anaphor in (69) is case-marked, with the case realized on its antecedent. This needs to be further explored.

REFERENCES

AWOYALE, Y. 1985. On the semantic interpretation of serial verb constructions. West African languages in education: Papers from the 12th West African Linguistics Conference, ed. by K. Williams, 144-174. Port Harcourt.

-----. 1987. Perspectives on verb serialization. Niger Congo syntax and semantics, ed. by V. Manfredi, 1.3-35. Boston: Boston University Press.

BAKER, M. 1989. Object sharing and projection in serial verb constructions. Linguistic Inquiry 20.513-553.

BAMGBOṢE, A. 1982. Issues in the analysis of serial verb constructions. Journal of West African Languages 12.2-21.
CAMPBELL, R. 1991. Argument-sharing serial verbs in Akan. University of Pennsylvania, MS.
CARSTENS, V. 1988. Yoruba serial verbs. Paper presented at the 2nd Niger-Congo Syntax and Semantics Workshop, MIT, Cambridge, MA, April 1988.
CHOMSKY, N. 1982. Some concepts and consequences of the theory of government and binding. Cambridge, MA: MIT Press.
-----. 1986. Knowledge of language: Its nature, origin, and use. New York: Praeger.
CHUNG, S. 1989. On the notion 'null anaphor' in Chamorro. The null subject parameter, ed. by O. Jaeggli & K. Safir, 143-184. Dordrecht/Boston/London: Kluwer Academic Publishers.
COLE, P. 1987. Null objects in universal grammar. Linguistic Inquiry 18.597-612.
DECHAINE, R. 1988. Towards a typology of serial constructions in Haitian. Niger Congo syntax and semantics, ed. by V. Manfredi, 1.49-64. Boston: Boston University Press.
EKUNDAYO, S.A., & F.N. Akinnaso. 1983. Yoruba serial verb string commutability constraints. Lingua 60.115-133.
HAEGEMAN, L. 1991. Introduction to government and binding theory. Oxford: Basil Blackwell.
LANIRAN, Y., & O. Sonalya. 1988. The lexical nature of Yoruba serialization. Niger Congo syntax and semantics, ed. by V. Manfredi, 1.37-48. Boston: Boston University Press.
LARSON, R. K. 1991. Some issues in verb serialization. Serial verbs: grammatical, comparative, and cognitive approaches, ed. by C. Lefebvre. Amsterdam/Philadelphia: Benjamins.
LEFEBVRE, C. 1991. *Take* serial verb constructions in Fon. Serial verbs: grammatical, comparative, and cognitive approaches, ed. by C. Lefebvre. Amsterdam/Philadelphia: Benjamins.
LI, C. N., & S. Thompson. 1973. Serial verb constructions in Mandarin Chinese: subordination or coordination? Chicago: Chicago Linguistic Society.
MATISOFF, J. 1973. The grammar of Lahu. Berkeley: University of California Press.
MUYSKEN, P. 1988. Parameters for serial verbs. Niger Congo syntax and semantics, ed. by V. Manfredi, 1.65-75. Boston: Boston University Press.
OYELARAN, O. 1982. On the scope of the serial verb construction in Yoruba. Studies in African Linguistics 13:2.109-146.
RIZZI, L. 1986. Null objects in Italian and the theory of *pro*. Linguistic Inquiry 17:3.501-557.
ROSS, J. R. 1967. Constraints on variables in syntax. MIT Ph.D. dissertation in Linguistics.
SAXON, L. 1989. Agreement in Dogrib: Inflection or cliticization? Theoretical perspectives on Native American languages, ed. by D.

B. Gerts & K. Michelson, 149-162. Albany, NY: State University of New York Press.

----- 1990. Reflexive agreement binding. Proceedings of the 16th Annual Meeting of the Berkeley Linguistic Society, Special Session on Amerindian Linguistics. Berkeley: Berkeley Linguistic Society.

SCHACHTER, P. 1974. A non-transformational account of serial verbs. Studies in African Linguistics, Supplement 5.253-270.

SEBBA, M. 1987. The syntax of serial verbs. Amsterdam: Benjamins.

SELLS, P. 1985. Lectures in comparative syntactic theory. Stanford: Center for the Study of Language and Information.

STEWART, J. 1963. Some restrictions on objects in Twi. Journal of West African Languages 1.145-149.

WESTERMANN, D. 1930. A study of the Ewe language. London: Oxford University Press.

WILLIAMS, E. 1984. Grammatical relations. Linguistic Inquiry 15.639-673.

Studies in the Linguistics Sciences
Volume 23, Number 1, Spring 1993

YES, WE HAVE NO BANANAS: ENGLISH NEGATIVE TAGS IN CROSS-LINGUISTIC COMMUNICATION*

Martin J. Baik and Rosa J. Shim

A study was conducted with Korean-English bilinguals in both Korean and English. Two groups of bilinguals, proficient and non-proficient in English, were asked to answer a series of negative tag questions. Results indicate the following: First, the proficient group was generally more accurate than the non-proficient group in answering English questions. Second, Korean-English bilinguals were not aware of the constraints in English positive disagreement strategies; thus they do not feel that positive disagreement is any more difficult than negative agreement. Third, learning the English system does not have any effect on the native language system; thus they are able to code-switch freely without difficulty.

1. Introduction

As human beings we depend largely on language to communicate with other people in society. In the ever-increasingly complex world that we are living in, cross-linguistic communication is a common feature. On the campus of the University of Illinois, for example, there is a large body of foreign students who come from diverse linguistic backgrounds. The question that arises in such a situation is whether or not we truly understand each other by what we mean rather than by what we say.

In the profession of language teaching and learning, awareness of cross-linguistic interaction has brought about various types of studies that show contrasts between the learner's language and the target language. Contrastive Analysis (Dulay & Burt, 1972, 1974, 1975; Dulay et al. 1982), Error Analysis (Corder, 1971; Richards, 1971), Interlanguage Hypothesis (Selinker, 1972), and Markedness Hypothesis (Eckman, 1977) can be considered major developments in the recent past. Unfortunately, most of the efforts in these fields have come from the English Language Teaching (ELT) profession, and there is an unmistakable bias in these studies that treat the systems in English as the norm and the systems in other languages as either deficient, or as a departure from the norm.

In this study we try to break away from such a bias. Our concern is in describing the use of one simple but major system of communication, that of expressing agreement or disagreement to negative opposite-polarity tag questions in English, and a comparable construction in Korean.

It has been noted by Pope (1976) that in English, Yes/No answers correspond with the positivity or negativity of the proposition that follows the answers while there are other languages in the world in which Yes/No answers signal agreement/ disagreement with the proposition in the question. The former system is referred to as the positive-negative system and the latter is called the agreement-disagreement system. Korean is one of the languages that can been described as having an agreement-disagreement system.

The two systems do not differ in answering simple questions with positive propositions. In most cases, comparable structures of questions exist in both English and Korean and the same answers are given in both languages. Thus for question (1) and question (2), the answers are not different in the two languages.

(1) Are you a student?
Yes. (I am a student) / No. (I am not a student)

(2) hakseng- ip- nika?
student BE QM

Ye. (hakseng- ipnita) / Aniyo (haksengi- anipnita)
Yes (student BE) / No (student not BE)

However, problems arise when negative questions are asked in the two different systems. In the following section of this paper, we will give a brief description of the two systems of answering negative tags and some of the pragmatic conditions that are related to the use of negative opposite-polarity tags in English and a comparable construction in Korean. Then, we will discuss some of the practical problems that arise from the differences in the two systems. Following this discussion we will outline the present study and present the results. Finally, we will discuss the results in relation to previously mentioned problems and suggest further areas of research.

2. Negative tag questions: Answering systems and pragmatic concerns

When Yes/No answers are given to negative opposite-polarity tags in English, the prescriptive English grammar that we impose on non-native speakers of English mandates that answers must match the positivity or negativity of the proposition that follows the answers. For example, when question (3) is asked, the answers can be *Yes, I did* in disagreement or *No* in agreement with the questioner's assumption that the answerer did not go to the library. Since the positive proposition *I did* is in disagreement with the negative assumption in the question, Pope (1976) calls this POSITIVE DISAGREEMENT. Similarly, negative answers are labelled NEGATIVE AGREEMENT since the negative proposition in the answer is in agreement with the assumption of the question.

(3) You didn't go to the library, did you?
Yes, I did. / No. (I didn't.)

In addition to Pope's simple description of answers to negative tags in English, Houck (1991) discusses the pragmatic implications in asking such questions. In the case of falling-intonation opposite-polarity tag questions, she claims that the questioner has to assign a strong probability to the truth of the proposition and at the same time assume that the answerer would also assign a strong probability to the truth of the proposition.

The implication of Houck's contention is that the necessary pragmatic context for falling-intonation opposite-polarity tag questions are met only when the question is asked in confirmation of the truth of the proposition in the question. If this is in fact true, then the expected answer to a negative falling-intonation opposite-polarity tag question is always *No*.

However, it does not necessarily follow that we only hear the answer *No* to negative tag questions in real life. Regardless of the expectations, we sometimes do need to answer such questions with a disagreeing *Yes*, and when we do, it is not sufficient to simple say *Yes*. Rather, we normally repeat the tag in the answer in order to achieve successful communication. The point noted by Pope which explains the need to give fuller answers in positive disagreement is our tendency to use the word *Yes* in agreement. This is what we do in answers to positive questions. However, this tendency is not fulfilled when answering negative questions since the answer *Yes* here is used to disagree rather than agree. This explains the need to say something more.

Whether the reasons lie in our tendency to use *Yes* in agreement rather than in disagreement, or in the fact that pragmatic conditions are not fulfilled in positive disagreement, Pope agrees with Bellugi's (1967) observations that the use of the word *Yes* as a positive disagreement strategy is the last category of negation to be acquired by children.

In Korean there is no parallel construction to the English tags. However, it is possible to ask negative questions which have similar pragmatic constraints as English negative tags. In this sense, there is a comparable construction in Korean that matches question (3):

(4) tosekwan -e an -ka ss -ci -yo?
library to not go PAST QM HON
(You) didn't go to the library, did you?

In question (4) the question marker (QM) *-ci* has the function of asking confirmation of the truth of the proposition in the question. The morpheme *ss* marks the past tense of the verb *ka* 'go', and *-yo* is an honorific morpheme. Thus the pragmatic constraints on question (4) are similar to that of question (3).

It is interesting to note, however, that the expected answer to this question is exactly the opposite of English. Although we have already seen that *Ye* means *Yes* and *Aniyo* means *No*, the answer *Ye*

is given to confirm the assumptions of the negative question. At the same time, the answer *Ye* signals that the negative proposition that is given in the question is in fact true. When the negative answer *Aniyo* is given to this question, it is used to contradict the assumptions of the question. In addition to disagreeing with the assumptions, *Aniyo*, in this case, signals that the negative proposition of the question is false. In other words, this signals positive disagreement since the true proposition that naturally follows the answer is a positive one.

A further interesting observation is that, unlike in English, this does not cause problems in giving simple answers in Korean. The reasons is that the answers *Ye/Aniyo* signal clear agreement/ disagreement with the proposition in the question, and these answers do not have a direct relationship with the proposition of the extended answer. Thus misunderstanding is unlikely to occur, and Koreans do not feel the need to supply extended statements to positive disagreement answers.

3. Problems in cross-linguistic communication

The problems are predictable when we have two systems of interaction that are identical on the one hand and antithetical on the other. Once Korean-English bilinguals establish the similarity in the English system and the Korean system of answering positive questions, they will invariably transfer that knowledge when answering negative questions. The consequences of this transfer are that there is a possibility for a mismatch between one's meaning and one's utterance when a Korean answers negative questions in English.

In the real world, this type of transfer can cause serious problems for the non-native speaker if native speakers are not aware of differences in answering systems in different languages. Gumperz (1982) reports of a Philippino doctor who was charged with perjury when in fact he was simply misunderstood by the FBI agents who were not aware that the doctor was using a different system of linguistic interaction. The case was eventually dismissed when the defense attorneys called in a linguist to closely examine the doctor's speech. The doctor was shown to be using English in a way that was not familiar to monolingual English speakers, and one of the major areas of conflict was in the system of answering negative questions. The doctor's native language, Tagalog, is a language that has the agreement-disagreement system, and occasionally he made mistakes when giving answers to negative questions, leading to apparent contradictions in his testimony. The following is an example of his speech at a Navy hearing (Gumperz 176):

(5) Q: It's the testimony by Lieutenant Commander Gilbert that you did not attend the briefing.
(6) A: Yes.
(7) Q: You did attend it?
(8) A: No.

As the above example clearly shows, the Philippino doctor is misunderstood by the questioner when he gives answer (6). The questioner takes this answer to be positive disagreement even though the answer is not elaborated on with an extended propositional statement. The point is that the doctor was being consistent in his use of the agreement-disagreement strategy, and this was interpreted by monolingual native speakers of English to be a contradiction in his testimony.

4. Aims of the study

The study presented in this paper was conducted for two main purposes: First it was to investigate the uses (or misuses) of Yes/No answers to English negative tag questions by two groups of Korean-English bilinguals, those that are proficient and those that are non-proficient in English. We have established that there is a source of miscommunication which is a direct result of antithetical rules that are applied to answering negative questions in English and Korean. Cowan (1983) gives the generalization that it is precisely in such circumstances that there is the greatest possibility of errors. Nevertheless, we do not believe that this is a permanent condition. Human cognitive abilities enable us to encompass many conflicting systems of thought. It should not be impossible to attain a level of proficiency in English that allows proficient non-native English speakers to answer negative tags to near-native accuracy. Thus it is expected that Korean-English bilinguals that have achieved the level of proficiency in English that is deemed appropriate for graduate level study will be more accurate in answering negative tag questions in English than the non-proficient, intermediate level English learners.

The second question asked in this study stems from the different answering strategies in the two languages when giving answers of positive disagreement. How many Korean-English bilinguals are aware that they not only need to give a completely different answer in English but also that they need to elaborate on their answers of positive disagreement? Will they feel more constrained in giving a simple Yes/No answer when the answer is in positive disagreement? In other words, will they make more mistakes in positive disagreement situations than in negative agreement situations?

The third question that we wanted to look at in this study was whether the learning and using of a different system of answering negative questions had any effect on the bilingual's first language. In other words, as a Korean-English bilingual becomes more proficient in the English answering system, would this have a negative effect on the Korean answering system? A related questions is whether Korean-English bilinguals in general will have difficulty maintaining the answering system of their native language if for some reason they need to code-switch to Korean from an English speaking situation. In these circumstances, will they become less accurate in giving answers to Korean negative questions?

5. Methodology of the study

In order to answer the above research questions, two groups of Korean-English bilinguals, non-proficient, intermediate-level English learners that were enrolled in the Intensive English Institute at the University of Illinois, and proficient English speakers that were graduate students at the University of Illinois, were asked 16 negative tag questions in English and Korean. There were 17 subjects in each group. The mean length of stay of the students in the US was six months, and three years and ten months, respectively.

Both English and Korean questions were masked as asking opinions on four commercial products: jeans, diamonds, cars, and cola. Advertisements were taken from magazines on Levis, diamonds, Toyota, and Coca Cola. Similar advertisements were chosen for Korean questions. The Korean questionnaire was a direct translation of the English questionnaire.

In order not to give the subjects the impression that they were asked only negative questions, an equal number of positive questions were asked. Four possibilities were given for answers to English questions: *Yes*, *No*, *Don't know*, and *Don't understand the question*. For Korean questions, the possibility that the subjects would not understand the question was remote, hence the fourth option was not given.

We had to be sure that the answers to the negative questions were either correct or incorrect depending on the simple *Yes/No* answers that the subjects gave. Thus all of the negative questions were very simple in that there was no ambiguity in the truth or falsity of the proposition in these questions. Discussions with the subjects after the test confirmed this fact. The positive questions, however, asked for the subjects' genuine opinion so as to disguise the purposes of the test to some extent. The questionnaire is given in full detail in Appendix A.

Both the English questions and the Korean questions were audio-taped with approximately equal limit of time (7 seconds) to answer for each question. All of the English negative questions were asked with falling-intonation tags. Negative questions in Korean were asked with normal rising intonation which implies the same pragmatic conditions as the English falling-intonation opposite-polarity tag questions.

When all the subjects arrived at the testing site, they were asked to read the cover sheet for the English questionnaire. Then they were given photocopies of the four advertisements which they could examine at leisure. They were allowed to discuss the product or ask questions about the product before they started answering questions on them. The subjects were told that they will be asked to answer questions about the four products. When everyone was ready, the tape containing the questions was played back. Subjects

were not allowed to speak during the test. Upon completion of the English questionnaire, the same process was followed for the Korean questionnaire.

At the end of the testing sessions, completed questionnaires were collected, and the subjects were informally interviewed regarding the difficulty of the questions asked, whether or not they guessed what the purposes of the tests were, if they felt that they had to say something more than a simple Yes or No, and various other questions concerning the products in the questionnaires. A full list of questions asked at the informal interview is given in Appendix B.

6. Results

The accuracy rate was calculated in percentages of number of correct answers over number of answers given. Since the subjects had the option of saying 'Don't know' when they in fact did not know the answer to a question, the items for which a subject gave 'Don't know' as the answer was not included in the analysis. Results from the proficient speakers are given in Table 1. and results from the non-proficient speakers are given in Table 2. PD refers to POSITIVE DISAGREEMENT and NA refers to NEGATIVE AGREEMENT in each table.

n = 15	English			Korean		
Length - Stay (Year:Month)	Overall	PD	NA	Overall	PD	NA
10:03	100.00%	100.00%	100.00%	93.33%	100.00%	85.71%
1:04	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
8:07	100.00%	100.00%	100.00%	93.33%	88.89%	85.71%
3:04	86.66%	75.00%	100.00%	100.00%	100.00%	100.00%
2:06	93.75%	88.89%	100.00%	93.75%	88.89%	100.00%
4:07	80.00%	62.50%	100.00%	81.25%	100.00%	57.14%
4:03	93.75%	88.89%	100.00%	86.67%	100.00%	71.43%
5:05	87.50%	100.00%	71.43%	100.00%	100.00%	100.00%
6:03	93.33%	87.50%	100.00%	100.00%	100.00%	100.00%
1:02	100.00%	100.00%	100.00%	75.00%	66.67%	85.71%
2:00	86.66%	100.00%	66.67%	100.00%	100.00%	100.00%
1:04	87.50%	100.00%	71.43%	87.50%	77.78%	100.00%
4:09	93.33%	87.50%	100.00%	83.33%	66.67%	100.00%
5:04	93.75%	88.89%	100.00%	93.33%	87.50%	100.00%
1:03	92.86%	100.00%	85.71%	100.00%	100.00%	100.00%

Table 1. Accuracy rate of proficient Korean-English bilinguals

Out of 17 possible subjects for each group, data from two subjects in the proficient group were excluded from the analysis. Accuracy rates for the two subjects excluded from the proficient group are given in Table 3.

n = 16	English			Korean		
Length - Stay (Year:Month)	Overall	PD	NA	Overall	PD	NA
0:03	92.86%	87.50%	100.00%	93.75%	100.00%	85.71%
0:03	100.00%	100.00%	100.00%	93.33%	100.00%	85.71%
0:04	18.75%	11.11%	28.57%	100.00%	100.00%	100.00%
0:05	12.50%	11.11%	14.29%	87.50%	88.89%	85.71%
0:10	0.00%	0.00%	0.00%	100.00%	100.00%	100.00%
0:10	37.50%	22.22%	57.14%	68.75%	77.78%	57.14%
0:07	12.50%	0.00%	28.57%	86.67%	87.50%	85.71%
0:11	93.75%	100.00%	85.71%	100.00%	100.00%	100.00%
0:09	33.33%	37.50%	28.57%	86.67%	75.00%	100.00%
0:03	12.50%	0.00%	28.57%	100.00%	100.00%	100.00%
0:03	7.14%	0.00%	16.67%	86.67%	87.50%	85.71%
0:11	86.67%	87.50%	85.71%	92.86%	85.71%	100.00%
0:05	6.67%	0.00%	16.67%	93.75%	88.89%	100.00%
0:04	86.67%	88.89%	83.33%	93.75%	100.00%	85.71%
0:04	100.00%	100.00%	100.00%	93.75%	88.89%	100.00%
0:05	7.69%	0.00%	14.29%	81.25%	66.67%	100.00%

Table 2. Accuracy rate of non-proficient Korean-English bilinguals

		English			Korean		
Length - Stay (year:month)		Overall	PD	NA	Overall	PD	NA
4:01	I-1	93.75%	88.89%	100.00%	25.00%	33.33%	14.29%
1:05	I-2	6.25%	0.00%	14.29%	100.00%	100.00%	100.00%

Table 3. Idiosyncratic responses from the proficient group

As Table 3 illustrates, the first subject (henceforth I-1) had stayed in America for four years and one month. His responses showed that he had almost perfect accuracy rates for English questions while his accuracy rates for Korean questions were well below chance level. On the other hand, the second subject (henceforth I-2) who had lived in America for a year and five months had very low accuracy rates for English questions while there were absolutely no errors in his answers for all the Korean questions. In order to avoid the risk of having these data distort the general picture of the results, we will discuss them separately.

As for the non-proficient group, responses of one subject were taken out of the analysis. The reason was that he admitted to giving the answer 'Don't know' whenever he was not sure which answer was correct. He understood the questions and knew the answers, but he answered 'Don't know' whenever he did not have enough time to figure out which one of them conveyed his meaning. He was very conscious of the differences in the two answering systems and was

afraid of making mistakes. Thus we felt that his accuracy rates were not representative of his unmonitored use of Yes/No answers.

Thus analysis of data was carried out with data from 15 subjects in the proficient group and data from 16 subjects in the non-proficient group. The means and standard deviations of all the different columns in Tables 1 and 2 are given in Table 4.

		English			Korean		
		Overall	PD	NA	Overall	PD	NA
Proficient	Mean	**92.61%**	**91.94%**	**93.02%**	92.50%	91.76%	92.38%
	s.d.	6.01%	11.15%	12.58%	8.09%	12.25%	13.08%
Non-Proficient	Mean	**44.28%**	**40.36%**	**49.26%**	91.17%	90.43%	91.96%
	s.d.	40.46%	44.19%	36.80%	8.25%	10.49%	11.63%

Table 4. Results of mean and standard deviation

As the above table illustrates, the major differences were in the accuracy rates for answers given to English questions by the proficient group (in bold) and the non-proficient group (in bold and underlined). Tests of variance between the two groups revealed differences in the variance for the two groups that were statistically significant ($p < .001$) in all three categories (Overall, PD, NA). Furthermore, t-tests[1] revealed that the accuracy rate of the proficient group was significantly higher than that of the non-proficient group in all three categories (overall: $p < .001$; PD: $p < .001$; NA $p < .001$).

Difference of accuracy rates within the non-proficient group for English (bold, underlined) and Korean (underlined) questions were also statistically significant ($p < .001$ for all three categories). Tests of variance also revealed that the variance for the two sample data sets were significantly different ($p < .001$ for all three categories).

Within-group t-tests for differences between PD category and NA category did not reveal any statistically significant differences for any group or language. Further, differences between accuracy rates for English and Korean questions within the proficient group were not statistically significant, and no significant difference was found for the accuracy rate for Korean questions by the two groups.

One other analysis that was carried out was the analysis of correlation. Since we had access to the length of stay in the US for each subject, correlational analyses were done so as to find out if there was any relationship between length of stay in the US and the subjects' accuracy rate for English questions. No significant correlation was found. Further correlational analyses were done to see if the subjects' length of stay had an effect on the accuracy rate for Korean questions, and no significant correlation was found.

7. Discussions and implications

Our first research question was whether proficient bilinguals would be better in giving accurate Yes/No answers to negative tag questions in English than the non-proficient bilinguals. Our hypothesis was that they would be, and the results of the study support this hypothesis. In addition to the fact that proficient bilinguals attain a very high level of accuracy, the results also show that there is no relationship between length of stay in the US and the subjects' accuracy in answering negative tag questions. Although the subjects' length of stay in US ranged from a year and three months to ten years and three months, the subjects' level of accuracy did not reflect this fact. This was also true of the non-proficient speakers: There was no correlation between length of stay and accuracy in answering negative tag questions.

The second question asked in this study was whether Korean-English bilinguals understood the constraint in giving answers of positive disagreement in English. Results of the study and results from interviews with the subjects indicate that Korean-English bilinguals are not aware of such a constraint. Neither of the groups showed any difference in their accuracy to positive disagreement (PD) answers and negative agreement (NA) answers. When asked if any of them felt that they needed to elaborate on their answers, the responses were that sometimes they would, but that they neither felt nor knew that it was necessary to do so.

The third aim of the study was to find out if learning and using the English system of answering negative questions would have any effect in the Korean answering system. Although we could not find perfect accuracy rates in either of the groups, the results did show that there is no difference in the performance of proficient and non-proficient groups in answering Korean questions. Therefore, whatever the reasons may be for the subjects in this study to have a less than perfect accuracy rate in giving answers in their native language, it does not seem to be due to the level of mastery of a different system. In other words, gaining proficiency in answering English negative questions does not seem to contribute toward a decline in their level of accuracy for Korean questions.

This brings us to the discussion of the two idiosyncratic responses in the proficient group. They are repeated here for convenience sake.

Length - Stay (year:month)		English			Korean		
		Overall	PD	NA	Overall	PD	NA
4:01	1-1	93.75%	88.89%	100.00%	25.00%	33.33%	14.29%
1:05	1-2	6.25%	0.00%	14.29%	100.00%	100.00%	100.00%

For I-1, it is apparent that there has been counter-transference from the English system to the Korean system. Although his respons-

es do not fit into the general picture obtained from this study, we cannot rule out the possibility that using the English positive-negative system may sometimes result in mental restructuring of a Korean-English bilingual's answering system that could cause communicative problems in the bilingual's native language. In the case of I-2, the results are completely the opposite. He showed 100% accuracy for the Korean questions while he was able to answer only one question correctly in English. Keeping in mind that his general level of proficiency in English placed him in the proficiency group, it is rather strange that he did not even match the performance of the non-proficiency group. An interesting observation from the cases is that both subjects seem to have more difficulties in the PD category than in the NA category. While this could be mere coincidence, it could also be a reflection of the acquisition sequence of the two categories.

A tentative interpretation that can be obtained from the above idiosyncracies is that there are some individuals who fail to maintain two opposite systems of interactions at the same time. In one case, the consequence was in the restructuring of thè whole system to fit the system of the second language. In the other case it resulted in maintaining the native language system at the expense of disregarding the system of the second language to the extent that the acquisition rate for the system of the second language was considerably slowed down. It is possible that I-2 is still in the process of acquiring the English system. It may be taking him longer to acquire the different system, and once he does acquire it, he could follow the pattern shown by I-1 and lose his proficiency in the Korean system.

According to the results of this study, we can say that most proficient bilinguals would simply say 'Yes' for positive disagreement in the majority of instances and be correct to a certain degree. As far as prescriptive grammar is concerned, this is the correct answer. However, prescriptive grammar does not mention the fact that one needs to elaborate when the answer is a positive disagreement. When non-native speakers fail to give any elaboration, monolingual native speakers may feel that the response of the non-native speaker was rude and curt. Although propositional meaning may be conveyed by the simple 'Yes' for positive disagreement, it unfortunately conveys an additional message that is certainly not part of the answerer's intent.

A second point to be noted is the fact that non-proficient bilinguals would say 'Yes' to negative tag questions, not as positive disagreement, but as negative agreement more than 50 percent of the time. The fact that they do answer the questions correctly some of the times adds to the difficulty of interpreting their answers. The consequences in cross-linguistic communication are predictable. Unless the native-speaker is aware of the communicative situation (that there is a non-native speaker involved who may use an opposite

answering system for negative questions), the native-speaker will make judgments about the non-native speaker similar to that of the the FBI agents' accusations of perjury against the Philippino doctor.

The issue that we would like to raise at this point concerns the accuracy rate of both groups in Korean questions. We have established that level of proficiency could not be a factor in their less than perfect performance. One relatively uninteresting reason behind this result may be that the Korean test was administered right after the English test: It could simply be a testing effect.

If the above factor does not account for the inaccuracies observed, a second straightforward cause could be the fact that all the subjects in this study were bilinguals living in a second language environment, regardless of their level of proficiency. The simple fact that they are exposed to a different system could have resulted in an error range in their native system.

A third explanation, which we advocate is that the level of accuracy obtained in this study is simply a reflection of real language use by monolingual Koreans. In other words, we believe that even monolingual native-speakers of a language can sometimes give an answer that is different from what is prescribed in the grammar. It is not a reflection of their competence but a reflection of what is said in everyday communication.

In a language such as Hindi, Bhatia (1974) reports the co-existence of three different systems of answering negative questions. The questions is: How do people understand what they mean? It is apparent from Bhatia's account of the phenomenon that the answers are rarely misunderstood because of various pragmatic presuppositions, implications, and expectations (PIEs) governing the use of negative questions and their answers in Hindi. This does not mean that Hindi speakers do not have a preferred norm. Bhatia reports that the positive-negative system is the predominant strategy used by Hindi speakers. Nevertheless, there are instances when Hindi speakers would use the agreement-disagreement system and successfully convey what they mean.

Likewise, Americans do sometimes give answers in agreement-disagreement while communicating with other Americans, and they usually figure out the correct meaning from the context of situation. Koreans, too, may give answers following the positive-negative strategy while communicating with other Korean speakers. Although further studies need to be done to confirm our belief, observations of language use in real-life contexts lead us to believe that this is what is reflected in our data.

8. Summary and conclusions

In summary, the results of this study have indicated the following:

1. The level of accuracy in answering negative questions by non-native speakers of English corresponds to their general level of proficiency.
2. Non-native speakers of English who are accustomed to the agreement-disagreement system in their native language simply apply the opposite rule in answering negative questions in order to be accurate, but they are generally unaware of the pragmatic constraint on positive disagreement answers.
3. The attainment of proficiency in a second language with a different system of answering negative questions does not adversely affect the bilingual's performance in his/her native language.

Based on these results we have argued that proficiency in the answering system of the second language in terms of following the prescriptive rules does not necessarily guarantee successful communication. Further we have pointed out that non-proficient bilinguals will have predictable difficulties in getting their meanings across when the native-speaker interactant is not aware of the cross-linguistic variables. Third, we postulated that competence in one language system may not mean 100 percent accuracy in performance.

These discussions lead us to ponder upon the arguments by Kachru (1985) and Nelson (1985) that the difficulty in understanding institutionalized non-native varieties of English may be the result of attitudinal baggage rather than true unintelligibility. This may also be true of language learner varieties, especially in the case of answering negative questions where the source of the problem may be in intolerance stemming from prescriptivism.

9. Implications for further research

The discussion from the present study is limited to the performance of bilinguals that are aware of two systems of answering negative questions that are completely opposite to each other. We have found that the bilinguals do not perform at 100 percent accuracy rate in their native language. We have determined that this could not be caused by gaining proficiency in the second language. As we have already discussed, it is important to find out if native speakers of a language actually follow the system of the specific language a hundred percent of the time. It may be possible that the 90+% accuracy obtained in this study is representative of the native monolingual speakers' performance. However, this is only a speculation at this point. Further studies are needed to confirm this observation.

One of the limitations of this study lies in the scope of the negative questions asked. We concentrated only on opposite-polarity,

falling-intonation negative questions. The reason for limiting the scope of this study was that the pragmatic constraints for these questions have already been well defined. We feel that similar research needs to be carried out for simple negative questions without tags. The pragmatic conditions differ, and change in intonation plays an important role in conveying the presuppositions, implications, and expectations (PIEs) of the questioner. Undoubtedly these different PIEs could certainly have an effect on the strategies of answering if a speaker uses the agreement-disagreement system.

The challenge in second language acquisition is not only in acquiring the syntactic structure of the language, but also in acquiring communicative competence in the language. We believe that the knowledge of implicit PIEs in negative questions is a significant part of such communicative competence.

NOTES

*This paper was presented at the 7th Annual International Conference on Pragmatics and Language Learning, University of Illinois at Urbana-Champaign, April 1-3, 1993.

[1] Since tests of variance led us to reject the null hypothesis that the group variances were the same, the t-tests conducted were also tests comparing two means from samples of different variances.

REFERENCES

BELLUGI, U. 1967. The acquisition of the system of negation in children's speech. Harvard University Ph.D. dissertation.

BHATIA, T. K. 1974. The coexisting answering systems and the role of presuppositions, implications, and expectations in Hindi simplex Yes/No questions. Papers from the Tenth Regional Meeting Chicago Linguistic Society (CLS), 47-61.

CORDER, S. P. 1971. Idiosyncratic dialects and error analysis. International Review of Applied Linguistics 9.147-159.

COWAN, R. 1983. Toward a psychological theory of interference in second language learning. Second language learning: Contrastive analysis, error analysis and related aspects, ed. by B. W. Robinett and J. Schachter, 109-119.

DULAY, H. C., & M. Burt. 1972. Goofing: an indication of children's second language learning strategies. Language Learning (Urbana, IL) 22.235-252.

-----. 1974. Errors and strategies in child second language acquisition. TESOL Quarterly 8.129-136.

-----. 1975. Creative construction in second language learning. Papers in second language acquisition: Proceedings of the Sixth Annual Conference on Applied Linguistics, ed. by H. D. Brown, 65-80. Ann Arbor, Michigan: University of Michigan Press.

DULAY, H. C., M. Burt, & S. Krashen. 1982. Language two. New York: Oxford University Press.
ECKMAN, F. R. 1977. Markedness and the CA hypothesis. Language Learning (Urbana, IL) 27.315-330.
GUMPERZ, J. J. 1982. Fact and inference in courtroom testimony. Language and social identity, ed. by J. J. Gumperz, 163-194. Cambridge: Cambridge University Press.
HOUCK, N. 1991. Tag questions: a necessary pragmatic context. Pragmatics and Language Learning 2.29-38.
KACHRU, Y. 1985. Discourse analysis, non-native Englishes, and second language acquisition research. World Englishes 4:2.223-232.
NELSON, C. L. 1985. My language, your culture: whose communicative competence? World Englishes 4:2.243-250.
POPE, E. N. 1976. Questions and answers in English. The Hague: Mouton.
RICHARDS, J. C. 1971. A non-contrastive approach to error analysis. English Language Teaching 25.204-219.
SELINKER, L. 1972. Interlanguage. International Review of Applied Linguistics 10.210-231.

APPENDICES

Appendix A: English questionnaire

Thank you very much for participating in this little survey. This is a survey to find out the level of information that is available to the public on four widely known commercial products: the Levis jeans, the Coca-Cola drink, diamonds, and Toyota cars.

Please fill in the information on this page. In the following pages, you will find four pages of advertisement on the above four products. You will be allowed to study them at your leisure. Then you will be asked to answer several questions on each of the products.

There are two types of questions. The first type of question asks for a *Yes*, *No*, *Don't know* ,or *Don't understand question* answer. There will be a time limit on each of the questions (i.e., you will be given approximately seven seconds to answer each question).

Example question:

Please circle one of the answers.

Have you heard of the Levis jeans?

Yes No Don't know Don't understand question

In order to ensure your anonymity in the answers that you will give in this survey, we ask that you do not identify yourself on any part of the survey itself. Instead, we encourage you to leave your name and address with us on a separate information sheet so that we can inform you about the purposes of the survey after we have collected

all the data. Since the data will be collected from many different people on many different occasions, we are sorry that we cannot give out that information at this time.

1. Length of Stay in the United States: _____ years _____ months

2. Familiarity with products in survey:

	Very familiar	Slightly familiar	Unfamiliar
Levis Jeans	3	2	1
Toyota Cars	3	2	1
Diamonds	3	2	1
Coca-Cola	3	2	1

Questions on Levis Jeans

1. Is the Levis Jeans a popular brand in Korea?

Yes No Don't know Don't understand question

2. Jeans are not popular among young people, are they?
3. Do you think jeans are durable?
4. People do not usually wear jeans on formal occasions, do they?
5. Do women like to wear jeans as much as men?
6. Elderly people in Korea do not usually wear jeans, do they?
7. Do you think Levis jeans are expensive?
8. Office workers in Korea do not wear jeans to work, do they?

Questions on Toyota cars

1. Toyota is not made in Korea, is it?
2. Do you like Toyota cars?
3. Corolla is not a model of Toyota, is it?
4. Are Toyota cars generally more expensive than BMW cars?
5. Toyota cars are generally not cheaper than Hyundai cars, are they?
6. Have you seen this advertisement before?
7. Toyota is not a Japanese brand, is it?
8. Do you think the new aero-dynamic model from Toyota is attractive?

Questions on Diamonds

1. Is the diamond the most expensive jewel?
2. The pearl is not as expensive as the diamond, is it?
3. The diamond is supposed to be forever, isn't it?
4. Most people do not use the diamond as wedding gifts, do they?
5. Will the diamond break if it falls on rocks?
6. Women generally do not like diamonds, do they?
7. Is the size of the diamond an important factor to the price?
8. The clarity of the diamond is not important when choosing a diamond, is it?

Questions on Coca-Cola

1. Coca-Cola is not made from fruit juices, is it?
2. Is Pepsi-Cola the major competition of Coca-Cola?

3. There is no Diet form of Coca-Cola, is there?
4. Do you know that there is an RC brand cola?
5. Coca-Cola is not as famous as RC-Cola, is it?
6. Do you think it is not healthy to drink too much Coca-Cola?
7. Coca-Cola company does not make caffein-free products, do they?
8. Did you know that a can of Coca-Cola contains twelve teaspoons of sugar?

Appendix B: Interview questions

0. Do you have any general comments about the test? Any questions?
1. Were the questions difficult to answer?
2. Was the source of difficulty in the question itself or in the lack of information regarding the products?
3. Did any of you answer 'Don't know' when in fact you knew the answer but did not know how to answer the question?
4. We will give you the answers that we feel are correct. Please tell us if you disagree. (Go through the questions and discuss their answers.)
5. Did any of you guess at what the purposes of this test could be?
6. Did you feel more constrained in giving answers of positive disagreement, either in English or Korean? (Explain the meaning of positive disagreement.)
7. What is the strategy that you use in answering English negative questions?
8. Do you give propositional statements in addition to Yes/No answers?
9. Are you aware that you are required to give an extended answer when the your answer is 'Yes' in positive disagreement?
10. When you answer Korean negative questions, do you ever find yourself using the English system, rather than the Korean system?
11. Have you ever had the experience of miscommunication either in English or Korean because you were using a different system of answering than that of the questioner?

Studies in the Linguistic Sciences
Volume 23, Number 1, Spring 1993

THE STATUS OF AGREEMENT AND THE AGREEMENT PROJECTION IN ARABIC

Elabbas Benmamoun

This paper deals with two main issues that revolve around the problem of subject agreement. The first issue is whether there is one single pattern of subject agreement. The second issue is whether there are any motivations for the claim that agreement heads an independent syntactic projection. As far as the first issue is concerned, I show that Arabic evidence supports the claim that there are two independent agreement relations involving the subject and the verb, namely person and number. With regard to the second issue, I argue that there is no evidence that agreement is involved in nominative Case assignment in Arabic. This weakens the claim for an independent agreement projection, since Case has been the main syntactic reason for the postulation of such a projection.

1. Introduction

Pollock (1989) suggests an agreement projection located right above VP and below TP (and NegP). This raises two questions that have been left relatively open in most of the recent treatments of functional categories in general and of subject agreement in particular (Ouhalla 1988, Mahajan 1989, Chomsky 1991, Belletti 1991, among many others).

The first question concerns the internal structure of the agreement morpheme (its feature composition). The underlying assumption is that the node Agr contains Phi features of person, number, and gender which are realized by an affix under an agreement relation with the subject. I will present data that show that there are two agreement patterns in Arabic, one in number and gender and the other in person and gender. A verb, for example, may show both patterns or just one of them.

The second question has to do with whether there are any conceptual and empirical reasons for postulating an independent agreement projection beyond the need to have a place holder to derive the right word order. Again the facts from Arabic do not support a positive answer to this question. The main focus in this paper will be on Arabic and particularly on Standard Arabic and Moroccan Arabic.

2. Agreement features

Consider the following sentences from Standard Arabic:

(1) a. daxal-a ṭ-ṭullaab-u
enter-3SM the-students-M-Nom
'The students entered.'
b. daxal-at ṭ-ṭaalibaat-u
entered-3SF the-students-F-Nom
'The students entered.'

(2) a. ṭ-ṭullab-u daxal-uu
the-students-F-Nom entered-3PM
'The students entered.'
b. ṭ-ṭaalibaat-u daxal-na
the-students-F-Nom entered-3PF
'The students entered.'

(3) a. kaan-a ṭ-ṭulaab-u ya-drus-uun
was-3SM the-students-M-Nom Imp-3M-study-PM
'The students were studying.'
b. kaan-at ṭ-ṭaalibaat-u ya-drus-na
was-3SF the-students-F-Nom Imp-3F-study-PF
'The students were studying.'

As is well known, when the verb precedes the subject, agreement is in gender (and person) only, as shown in (1). When the subject precedes the main verb as in (2), agreement is in person and number, in addition to gender. Moreover, in (3) the subject agrees with the auxiliary that precedes it in person and gender only.[1]

The main question is how to account for the subject-verb agreement feature alternation. If we confine our attention to (1) and (2) we may account for the facts by positing two agreement affixes. One affix contains gender and person and is realized under the VS order as in (1). Another agreement affix contains number, person, and gender and is realized under SV order as in (2).[2]

Another way to implement the same basic idea would be to posit one abstract agreement affix and allow it to be either partially specified (gender and person) or fully specified (number, gender and person) depending on the structural relation that obtains between the verb and the subject. According to this account the Phi features are encoded on one affix, with full or partial specification contingent on the relative order of subject and verb.

(4) a. V S O ---> V+AGR (P.G)
b. S V O ---> V+AGR (N.P.G)

However, this account cannot carry over to the embedded verb in (3), where the imperfective form of the lexical verb carries two agreement affixes. The prefix carries person and gender, and the suffix carries number and gender. Moreover, the suffix shows up only when the subject precedes the verb (3). Thus, (3) shows that we

are not dealing with one agreement affix but rather with two affixes that happen to overlap in the feature gender. In other words, number and person agreement do not necessarily belong to the same morpheme, but rather constitute two independent morphemes. A plausible representation could be as shown in (5).

(5) Agr [N G] Agr [P G]

Further support for this conclusion comes from copular constructions. In Standard Arabic (and in fact in all the dialects that I know of, including Moroccan Arabic), the element corresponding to the verb *be* in the present tense is realized as a pronominal. I follow Eid (1991) and refer to this element as a pronoun copula. The use of this term is for purely descriptive purposes. For detailed analyses of this element in Arabic the reader is referred to the important work of Eid (1991), who, as far as I know, was the first to point out the interesting agreement patterns in (6). The examples are from Moroccan Arabic (disregarding phonological details).

(6) a. Omar **huwa** mul ḍ-ḍar
Omar 3MS owner the-house
'Omar is the owner of the house.'
b. Salmaa **hiya** mulat ḍ-ḍar
Salmaa 3FS owner-F the-house
'Salmaa is the owner of the house.'
c. Omar wa Sami **hum** mmalin ḍ-ḍar
Omar and Sami 3MP owners the-house
'Omar and Sami are the owners of the house.'

Confining our attention to the agreement features on the pronominal copula, notice that in (6) the copula apparently agrees with the subject in person, gender, and number. However, it is easy to show that agreement in person in the above examples is not with the subject but is a default feature on the copula. In (7), where the subject is a first or second person pronoun, the copula agrees with the subject only in gender and number; the person marking remains the same as in (6), viz. third person. There is no agreement in person as shown by the ungrammaticality of (8):

(7) a. ʔanaa **huwa** mul ḍ-ḍar
I-1S 3MS owner the-house
'I am the owner of the house.'
b. ʔanti **hiya** mulat ḍ-ḍar
You-2FS 3FS owner-F the-house
'You are the owner of the house.'

(8) a. *ʔanaa **ʔanaa** mul ḍ-ḍar
I-1S 1S owner the-house
'I am the owner of the house.'

b. *?anti ?**anti** mulat ḍ-ḍar
You-2FS 2FS owner-F the-house
'You are the owner of the house.'

The same agreement facts obtain with participial predicates, which agree with the subject in number and gender only:

(9) a. kaan-a ?al-?aṭfaal-u **naa?im-iin**
was-3SM the-children-Nom sleep-PM-Acc
'The children are sleeping.'
b. kaan-a ?al-ṭifl-u **naa?im-an**
was-3SM the-child-SM-Nom sleep-SM-Acc
'The child is sleeping.'

To sum up, the facts from the imperfective form of the verb, copular constructions, and participles strongly support the conclusion that there are two independent subject-verb agreement relations in Arabic, encoded by different features on different affixes; one number and gender affix which shows up under the SV order only, and another person and gender affix which shows up under both VS and SV orders.[3, 4]

The recognition of two agreement patterns in Arabic puts the debate over the agreement relation in general and the agreement projection in particular in a different perspective. Under an analysis where the derivation of inflectional morphology is derived in the syntax by the process of head movement that adjoins the lexical head to the inflection, one may need to posit more than one agreement projection. This may not be a negative move after all as long as we can conceptually and empirically ground the postulation of an agreement projection. The question that arises then is whether we can conceptually ground the agreement projection. This question has been addressed by Iatridou (1991) in the context of verb movement and adverb placement, and by Rouveret (1991) in the context of his analysis of agreement in Welsh. In the next section I address this question from a different perspective, namely its role in Case assignment.

3. Agreement as a projection

The second important issue that arises in the context of any discussion of agreement is the problem of projection. This issue has not been subjected to much debate, and it seems that this has to do, at least partially, with the still unclear role and status of agreement. Whether there is an agreement projection or not raises many questions concerning the necessary conditions for projecting a functional category. If we compare agreement on one hand and tense and affixal negation on the other, we realize that there is good evidence that tense and negation need to be syntactically accessible as independent projections rather than as features on a lexical head (see Benmamoun 1992 for discussion). As far as tense and negation

are concerned the evidence, both syntactic and semantic, for projection seems to be strong.

Consider negation for example. It is semantically relevant because it contributes to the truth value of a proposition (10):

(10) a. Omar qra le-ktaab
Omar read the-book.
b. Omar ma-qra-s le-ktaab
Omar Neg-read-Neg[5] the-book
'Omar did not read the book.'

It is also syntactically active. For example, it licenses negative polarity items (11) and interacts with quantifiers to give a narrow scope reading for the quantifier in (12):

(11) a. ma-zaa Hatta wahed
Neg-came any one
'No one came.'
b. ma-qra Hatta ktaab
Neg-read any book
'He did not read any book.'

(12) a. ma-qra-s bazaaf d-le-ktuba
Neg-read-Neg. many of-the-books
'He did not read many books.'

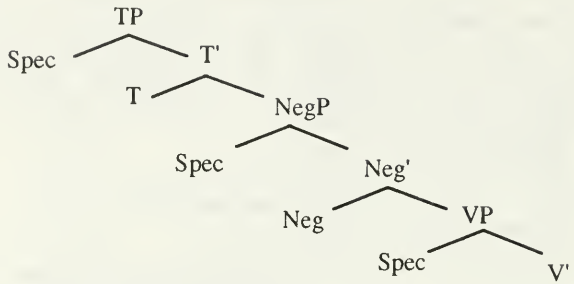
All these facts suggest that the negative is syntactically visible and accessible. As to whether it heads an independent projection, the morphological evidence of Standard Arabic shows that the negative carries tense (Benmamoun 1992):

(13) lam yaktub
Neg-Past write-3SM
'He did not write'

Sentences such as (13) can be accounted for by considering the negative as head of a projection between TP and VP. As such it blocks verb movement to Tense (due to minimality by acting as a potential antecedent in the sense of Aoun & Li 1989 and Rizzi 1990, see Benmamoun 1992 for details). Consequently, the tense inflection is hosted by the negative head (14). Notice the fact that negation hosts temporal information supports its status as head of a projection.

Turning to agreement, it seems to behave differently from the other heads. Its semantic function, if any, is not entirely clear. Syntactically, one possible motivation for an agreement projection can be found in Chomsky 1993, where subject-verb agreement, for example, is taken as an expression of Case assignment (also Borer 1986). However, given the theory that tense and agreement each head their own projection, the problem arises about how to derive the generalization that tense seems to play a role, given the fact that

(14)



nominative Case usually (but not exclusively) obtains in tensed clauses. According to Chomsky (1993) this can be derived by raising Tense to Agr (Agr-S). This insures that the Case relation is given its full expression under Spec-head relation mediated via the Agr head. This amounts to saying that nominative Case requires that the subject be in Spec-head agreement relation with both tense and agreement. The requirement obtains by combining agreement and tense.[6]

However, there are three contexts where this does not seem to be the case. First, the requirement is obviously not fulfilled in languages where the combination of tense and agreement fails. This is the case of negative sentences in Arabic where the negative carries tense and the verb carries agreement as in (13). Second, in copular constructions, the subject gets nominative Case even in contexts where there is no agreement that would mediate Case-assignment:

(15) ṭ-ṭaalib-u fii l-bayt-i
the-student-Nom in the-house-Gen
'The student is at home.'

We could obviously posit an agreement projection headed by an empty head. However, such a move makes one wonder why such an abstract head does not induce processes and repair strategies akin to Do-support in English when there is no eligible host, as is clearly the case in copular constructions in the present tense in Arabic.

Third, there are cases where the NP is assigned Case by one head while it agrees with another head. In the context of Q-float (Benmamoun 1993) the NP agrees with Q, though it is assigned Case by the external head (Joseph Aoun p.c.):

(16) ra?aytu l-?awlaad-a kulla-hum
(I) saw the-children-Acc all-Agr
'I saw all the children.'

Thus, nominative Case assignment can take place in the absence of an agreement feature on the Case-assigning head.[7]

Nevertheless, there is a solid generalization that an intimate relation exists between tense and agreement in languages like

English, where absence of tense entails absence of agreement. This is also reflected in contexts of sentential negation and questions, where Do-support affects tense and agreement.

I would like to suggest that the agreement feature may be part and parcel of the morphological template of heads such as tense, verbs, negation, questions.[8] Thus, in English, tense is specified for agreement while lexical categories such as verbs and adjectives are not. On the other hand, in Arabic, tense is not specified for agreement but only lexical categories and some functional categories such as the negative *laysa*.[9] This implies that in contexts where verb movement does not take place we correctly predict the verb to carry agreement in Arabic but not in English. Conversely, in such contexts, we correctly expect the categories that support tense to carry agreement in English but not in Arabic. This is indeed the case with tensed negatives in Standard Arabic where the tensed negative does not carry agreement (Benmamoun 1992).

Consequently, the agreement relation that obtains in English between the subject and the verb is indeed a reflection of nominative Case assignment but only because the agreement feature is part of the lexical specification of tense in that language.[10]

If this conclusion is correct, together with the arguments that Case is not mediated by an agreement morpheme (head of an agreement projection), we can conclude that agreement does not seem to correspond to a syntactic projection that would be treated on a par with other projections such as tense and negation.[11]

NOTES

[1] For detailed studies of agreement in Standard Arabic see Ayoub 1981, Fassi Fehri 1988, and Mohammad 1990, and the references cited there.

[2] It should be mentioned that not all analyses assume that there is person agreement with the postverbal subject. Thus, for example, in (i) it is assumed that the suffix *-at* carries gender only. This assumption can be problematic given (ii), where the same suffix is used in the context of a null pronominal.

(i) ʕaad-at Nadia
return Nadia

(ii) ʕaad-at
returned
'She returned.'

Under the theory of identification of null thematic pronominal elements, the person feature has to be present for the null pronominal to be identified. The crucial role of person agreement is shown by the fact that the subject cannot be dropped in the context of a predicate that does not contain the person feature as shown by

(iii) (see Kenstowicz 1989; see also Borer 1984, 1986 and Shlonsky 1989 for similar facts in Hebrew). This entails that the agreement suffix in (ii) contains the person feature to identify pro:

(iii) *(Omar) naaʔim
Omar sleeping
'Omar is sleeping.'

If the suffix *-at* contains the person feature in (ii) as required by the identification condition on null pronominals (in Arabic at least), and since the same affix is used in (i), then it is safe to conclude that the latter carries person agreement as well. See also Ayoub 1981 where it is assumed that under VSO the verb has both person and gender agreement.

[3] The suffix cannot be considered a resumptive pronoun related to a topicalized or left-dislocated NP. The main reason against this view is that the suffix lacks the main pronominal feature, which is person. Moreover, in (3) there is no representation in which the NP subject is left-dislocated or topicalized. Also, the suffix does not constitute together with the prefix a discontinuous morpheme. If we assume that it is a discontinuous morpheme, we still have to explain why the second part of this morpheme does not show up under VS. This is apart from the problems of representation and derivation that arise if a discontinuous morpheme is posited.

[4] Shlonsky (1989) independently argues for separating agreement features. He proposes that the features person, number, and gender each head an independent projection.

[5] I gloss *-s* as Neg. though I remain agnostic about its syntactic status. See Benmamoun 1992.

[6] This essentially goes back to the LGB-type analysis where nominative Case is assigned by INFL which contains both tense and agreement.

[7] Another variant of the same facts involves multiple agreement in the same clause. In Arabic both the auxiliary and the verb agree with the subject as illustrated in (i) from Moroccan Arabic:

(i) a. lewlaad kaanu taylaʕbu
the children be-Past-P play-P
'The children were playing.'
b. kaanu lewlaad taylaʕbu
be-Past-P the children play-P
'The children were playing.'

For an analysis of subject agreement in Moroccan Arabic and its comparison with Standard Arabic see Benmamoun 1992. See also Aoun, Benmamoun, & Sportiche forthcoming.

[8] This is not a new idea. It is essentially the content of the traditional analysis of the agreement relation as concord between two lexical elements (adjective and noun for example). The current

implementation of this idea consists of restricting the agreement relation to take place under only one of the three grammatical relations within the X' schema, namely Spec-head (Chomsky 1993).

[9] A more radical proposal is made by Rouveret (1991:7) who argues that agreement morphology can only attach to a functional head: 'La morphologie d'accord ne peut être affixé qu'à une tête fonctionelle.'

[10] In her GLOW 1991 abstract, Ritter accounts for the fact that in Hebrew, in the present tense the verb agrees with the subject in number and gender only, while in the past and future it agrees in person, number, and gender, by suggesting to specify agreement as [gender, number] in the present (her Num) and as [person, number, and gender] in the past and future (her D).

[11] Notice that even if subject agreement heads its own projection, the facts from Standard Arabic suggest that it should be located below negation; otherwise verb movement to agreement will be blocked for the same reason that verb movement to tense in the context of the negative *laa* is blocked in Standard Arabic (see Benmamoun 1992). This is another problem that arises whenever agreement is projected as the head of its own syntactic projection. One then is forced to locate it in different positions relative to the other functional projections depending on its morpho-phonological distribution in the language in question (in a kind of mirror principle fashion). This may well be the case, but what is needed are the theoretical principles that underly these 'parametric' choices. The mirror principle of Baker (1988), that is implicit or explicit in most analyses of inflectional and functional categories, cannot be used in a principled way as long as the syntactic role of agreement as a projection is not well defined. The mirror principle deals primarily with the isomorphism between the relative order of morphemes in a phonological string and their (relatively) well defined syntactic scope (e.g, passives, causatives, reciprocals ...). It is far from clear that the same applies to, say, the relative ordering of tense and agreement, or agreement and negation.

REFERENCES

AOUN, J., & A. Li. 1989. Scope and constituency. Linguistic Inquiry 20.141-172.

-----, E. Benmamoun, & D. Sportiche. Forthcoming. Agreement and conjunction in some Arabic dialects. Linguistic Inquiry 25.

AYOUB, G. 1981. Structure de la phrase en Arabe Standard. Université de Paris VII Ph.D. dissertation in Linguistics.

BAHLOUL, M., & W. Harbert. 1993. Agreement asymmetries in Arabic. Proceedings of the Eleventh West Coast Conference on Formal Linguistics, ed. by J. Mead, 15-31. Stanford, CA: Center for the Study of Language and Information.

BAKER, M. 1988 Incorporation: A theory of grammatical function changing. Chicago: University of Chicago Press.

BELLETTI, A. 1990. Generalized verb movement. Torino: Rosenberg & Sellier.

BENMAMOUN, E. 1992. Functional and inflectional categories: Problems of projection, representation, and derivation. University of Southern California Ph.D. dissertation in Linguistics.

-----. 1993. Null pronominals in the context of NP and QP. Proceedings of the Eleventh West Coast Conference on Formal Linguistics, ed. by J. Mead, 32-43. Stanford, CA: Center for the Study of Language and Information.

BORER, H. 1984. Parametric syntax. Dordrecht: Foris.

-----. 1986. I-Subjects. Linguistic Inquiry 17.375-416.

CHOMSKY, N. 1991. Some notes on the economy of derivation and representation. Principles and parameters in comparative grammar, ed. by R. Freidin, 417-454. Cambridge, MA: MIT Press.

-----. 1993. A minimalist program for linguistic theory. The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger, ed. by K. Hale & S. J. Keyser. Cambridge, MA: MIT Press.

EID, M. 1991 Verbless sentences in Arabic and Hebrew. Perspectives on Arabic linguistics III, ed. by B. Comrie & M. Eid, 31-62. Amsterdam: John Benjamins.

-----. 1992. Pronouns, questions, and agreement. Perspectives on Arabic linguistics IV, ed. by M. Eid, J. McCarthy, & E. Broselow. Amsterdam: John Benjamins.

FASSI FEHRI, A. 1988. Agreement in Arabic, binding and coreference. Agreement in natural languages: Approaches, theories, descriptions, ed. by M. Barlow & C. Ferguson, 107-158. Stanford, CA: Center for the Study of Language and Information.

-----. 1989. Generalized IP structure, case, inflection, and VS word order. MIT Working Papers in Linguistics, 10.75-111.

IATRIDOU, S. 1990. About Agr(P). Linguistic Inquiry 21.551-577.

JAEGGLI, O., & K. Safir. 1989. The null subject parameter and parametric theory. The null subject parameter, ed. by O. Jaeggli & K. Safir, 1-44. Dordrecht: Kluwer.

KENSTOWICZ, M. 1989. The null subject parameter in modern Arabic dialects. The null subject parameter, ed. by O. Jaeggli & K. Safir, 263-275. Dordrecht: Kluwer.

MAHAJAN, A. 1989. Agreement and agreement phrases. MIT Working Papers in Linguistics 10.217-252.

MOHAMMAD, M. 1990. The sentence structure of Arabic. University of Southern California Ph.D. dissertation in Linguistics.

OUHALLA, J. 1988. The syntax of head movement: A study of Berber. University College, London, Ph.D. dissertation in Linguistics.

POLLOCK, J-Y. 1989. Verb movement, UG, and the structure of IP. Linguistic Inquiry 20.365-424.

RITTER, E. 1991. Evidence for number as a nominal head. Generative Linguistics of the Old World (GLOW). Abstract.

RIZZI, L. 1990. Relativized minimality. Cambridge, MA: MIT Press.
ROUVERET, A. 1991. La nature des prépositions conjuguées. Université de Paris VIII, MS.
SHLONSKY, U. 1989. The hierarchical representation of subject-verb agreement. University of Haifa, MS.

Studies in the Linguistic Sciences
Volume 23, Number 1, Spring 1993

EXAMINING THE TRUSTWORTHINESS OF THE LATEST *OED* IN REFLECTING CURRENT ENGLISH

Marvin K. L. Ching
Department of English, Memphis State University

Although the second edition of the *Oxford English Dictionary (OED)* is still the paragon of lexicographical achievement, it suffers from seven serious flaws which must be rectified in future editions if it is to continue its supremacy. These deficiencies, though named by other writers with brief examples, are illustrated with a multiplicity of examples in this study to allow readers to understand the magnitude of the task of revision and to decide for themselves the pragmatic possibility or impossibility of achieving the *OED*'s goals of accuracy and comprehensiveness in future editions. It is easy to accede to the need for revision according to principles, but perhaps the size of the data nullifies the possibility of accomplishment, unless an innovative way is found to collect, organize, and revise the data.

1.0 Introduction

Through an investigation of contemporary American dictionaries, *The Oxford Dictionary of New Words* (1991), as well as words and expressions I have encountered in print and in other media and from actual speech, this study examines the second edition, the 1989 version, of the venerable *OED* for its dependability in reflecting current English. The *OED*, though still venerable, is lacking in: 1) the inclusiveness of terms cited, 2) the updating of the newest meanings of old terms, 3) the updating of quotations, 4) the clarity of definitions, 5) social sensitivity to changes in world view when defining terms, 6) usage labels or usage notes, and 7) a systematic method of listing terms and defining them for reader convenience.

Although both Harvey (1991:84-86) and Algeo (1990:131-150) have cited a number of these same points, readers cannot fully grasp the magnitude — and thus the probability or improbability of accomplishment — of the task for future editions of the *OED* without more specific, detailed examples. Moreover, although it is easy to accede to the areas for revision Harvey and Algeo have named, with their brief citation of specific examples of shortcomings, readers need to see a multitude of data to question the present method of employing paid and volunteer workers to accomplish the mammoth task of maintaining the *OED*, as Algeo has eloquently dubbed the *OED*, as the

'emperor' and 'the crowning achievement' of lexicography (131). But the explosion of lexical terms together with the explosion of publications in English are like the modern world's warp-speed changes predicted by Toffler's 1971 *Future shock*. Thus, after examining the strengths and liabilities of the new second edition of the *OED*, this study proposes an additional method for acquiring data to meet the demands of the times, if the *OED* is to remain the comprehensive and exemplary dictionary it has been.

Many (though not exclusively all) of the specific terms examined are those used by the general populace, which this writer has come across. They cannot be dismissed merely as 'soft terms,' because the *OED* has shifted not only from the literary world to terms in science, business, and medicine, but also to North American popular terms and slang (e.g., *brain-dead*, *nose job*, *acid rain*, *crack*, *asset stripping*, *barf*, and *drunk tank*) in its second edition (Gray 1989:95, 98). But in examining the specific terms used, readers can judge for themselves whether the *OED* should limit its aims in comprehensiveness, leaving certain terms for specialized dictionaries, or whether the missing gaps should be filled.

Details and examples of problems in scientific, business, and medical terminology will be left to future investigators.

A cross-comparison of American dictionaries reveals the adequacy or inadequacy of the *OED* in the seven areas named at the beginning. The comparison uses three general standard American desk dictionaries and the unabridged *Random House II* (2nd ed., 1987). They, together with their abbreviated names for this study, are as follows, and will be referred to as 'the four dictionaries':

Webster's Ninth New Collegiate Dictionary (1991) or *W*^{9th};
The American Heritage Dictionary (2nd college ed., 1991) or *AHD*;
Webster's New World Dictionary (3rd college ed., 1988) or *WNWD*;
The unabridged *Random House Dictionary of the English Language* (2nd ed., 1987) or *RHII*.

At times, reference shall also be made to these other works:

The Oxford Dictionary of New Words (1991) or *OD of NW*;
Webster's Third New International Dictionary (1971) or *WIII*;
Dictionary of American Slang (1975) or *D of AS*;
New Dictionary of American Slang (1986) or *ND of AS*; and
Webster's Dictionary of English Usage (1989) or *WD of EU*.

The *OD of NW* is also used because it is a wonderful source for 2,000 high-profile current words and phrases, only a fourth of which have been included in the new words and senses added to the *OED*. These dictionaries were not consulted in Algeo's study (1990:143).

2.0 Strengths

No doubt the *OED* is still a treasure. Its quotations indicate the context for the first recorded appearance of a term and its growth and spread to other contexts. For example, we find that 1952 is the year for the first appearance of *Ms.* when the National Office of Management Association in Philadelphia in *The simplified letter* directed that the term be used for all women addressees. After the following year with another quotation showing its use by the same business association, eight quotations in the 1970s appear, indicating its use in the Women's Movement. Another social trend can be observed: Apparently, with the growth of crime, *neighborhood* was compounded with *watch*, producing *neighborhood watch*, which appeared in print in 1972.

The *OED* is also important in dating the first use of a term in metaphorical contexts: e.g., *headhunter* 'a person recruiting a top executive officer' (1961); and *litmus test* 'a touchstone or decisive test' (with *a cultural litmus test* in 1957). The *OED* can thus trace new uses of old terms, correlating them with new trends in society.

Moreover, the quotations also include interesting conjectures on the etymology of a word, when it is uncertain. For example, although the *OED* states that the etymology of *jazz* is unknown, the numerous quotations surmising its origin include New Orleans creole and the West Coast of Africa, among other guesses.

The *OED* is also up-to-date in citing words reflecting new phenomena, sometimes the only dictionary with these words, which are not entered in the four general dictionaries cited at the beginning. For example, the *OED* alone cites *Scientology*, *head voice* (a musical term for the register between chest voice and falsetto), *Teflon-coated* (the others just cite *Teflon*), and *Rambo* with its derivative forms like *Ramboesque*, *Ramboism*, and *Rambo-like*. Moreover, unlike the three general standard desk dictionaries, the *OED*, with *RHII*, cites the slang meaning of *Oreo*, coined by African-Americans to indicate Blacks with White loyalties. The *OED* (like the four dictionaries) also includes the slang word *jawboning*. And the *OED* gives the dialectal meaning of *viewing* (a Southeastern Pennsylvania term for looking at a deceased body on display before the funeral according to Carver 1987:265), defined only by *WNWD* of the four dictionaries but cited by *RHII* as an example for the general sense of *viewing*.

The *OED* cites such computer terms as *lap-top* (omitted by *AHD* and *WNWD*); *window* (omitted by *W*^{9th} and *AHD*); and *hack*, *hacking*, and *hacker*. Among items from popular culture, the *OED* cites *junk food* (omitted by *W*^{9th}); *break dancing* (omitted by *AHD*); more compounds beginning with the word *women's* than *RHII* (which, in turn, contains more *women's* compounds than the general desk dictionaries) — namely the *OED* citations of *women's liberation*, *women's movement*, *women's rights*, and *women's studies*; *televangelist* (omit-

ted by *AHD* and *WNWD*); *passive smoking* and *passive smoker* (omitted by the three desk dictionaries, but *passive smoking* cited in *RHII*); and *access* as a transitive verb (in line with the American dictionaries); and treats *massaging the data or books* (omitted by *AHD* and *WNWD*).

Despite these achievements, however, the second edition of the *OED* lacks a number of terms used in current life, probably because of a heavy reliance, as the *OED* acknowledges in its prefatory materials (xxii), upon two *Barnhart Dictionaries of New English*, the *Barnhart Quarterly Companion*, and the record of new vocabulary words in book form published by Merriam-Webster. Although entirely new information dealing with an additional 5,000 words, combinations, and senses were included and integrated in the second edition of the *OED* chiefly for the first third of the alphabet where the Supplement was 20 or more years old (xii), many revisions are needed for a future third edition.

3.0 Problems

3.1 Terms not included

Some of the words or phrases not presently included are now basic to American life and can safely be predicted to endure, because of present trends of life. They often appear in some of the other dictionaries cited at the beginning of this paper, though the *OED* omits these items. In the following discussion whenever other dictionaries include these common terms that the *OED* neglects, the dictionary or dictionaries are listed in parentheses. The *OED* has not included *living will* (*RHII*, *AHD*, *WNWD*, *OD of NW*); *caregiver* (*RHII*, though a sharper definition is needed); *inclusive language* (*OD of NW* under *inclusive*); and *stress management*, though the *OED* cites these related health phrases: *stress disease*, *stress-free*, and *stress test* (the last item in *RHII*); and has also omitted *sexual harassment* (*RHII*), though citing these compounds having to do with sex: *sexual athlete*, *sexual revolution*, and *sex discrimination*.

Also missing are *blended family* (*RHII*); *Moral Majority* (*RHII*, *WNWD*, *OD of NW*); *New Age* (*WNWD*, *OD of NW*); *holiday blues/Christmas blues/yuletide blues*; *singles-bar* (*AHD*, *WNWD*, *RHII*); and *junk bond* (*W*^{9th}, *WNWD*, *RHII*, *OD of NW*). *Organ donor* is also omitted (its meaning in the *OED* under *donor* alone).

Words for biological classification of mother are also absent from the *OED*: namely, *birth mother* (but the *OED* lists *birth parent*); *donor mother* (but one of the definitions for *donor* in the *OED* pertains only to artificial insemination from the semen donated by a male); and *biological mother/father/or parent* (*RHII* under *biological parent*).

Words also missing are *rap*, a noun or verb pertaining to music (*RHII* as a verb and labeled *slang* and also under *rap music*; *WNWD* as a sense of the verb *rap*; *OD of NW* as a noun); *tabloid television*;

gay bashing; and *chaos theory*, a term recently applied to literary criticism (*OD of NW* only in its mathematical and physics sense; *OED* including only the word *chaology* with no definition but with two quotations, the last being 1775).

Unlike the terms above, the following words or phrases are rather new to the American public: *African American* (the *OED*, like the four dictionaries, lists *Afro-American*); the African-American use of *dis* or *diss*, an abbreviated verb clipped from a longer verb, perhaps from *disrespect*, *dismiss*, *discount*, or *distance* (*WNWD*, though it is unclear whether the African-American use is meant, because it is not labeled as an Americanism nor as slang; *OD of NW*); and *friendly fire*, pertaining to destruction from one's own military forces by mistake — a term widely used by the military in the Gulf War (*RHII*; *OD of NW*, which indicates that it was a euphemism used since the Vietnam War).

3.2 Updated definitions needed

Another area needing rectification is the updating of newer senses of old words. At times, the newer metaphorical meanings of old words and phrases have not been cited. The *OED*, for example, does not list the sense of *lightning rod* referring to a person bearing the criticism and punishment for others (*RHII*); and the meaning of *greening* as in C. Reich's *The Greening of America* (1970) (*WNWD*).

Certain senses of words in specific contexts are also lacking; e.g., *linkage* in terms of political connection between two different issues so that progress in one area must occur for progress on another issue to take place (*RHII*, *AHD*, *OD of NW*); *myth* as an explanation of the world, its phenomena, and the place of human beings in it, as used by Lakoff & Johnson (1980:185-194) in what they call the myths of *objectivism*, *subjectivism*, and the *experientialist synthesis*; *myth* also used by more liberal religious groups to refer not only to other people's but also to their own beliefs of the world, their destiny, and the purpose of life (*AHD*) — not understood as a fictitious, legendary, or prehistoric explanation of the creation of the world; and *ownership*, not meaning legal possession of something, but possession in the sense of understanding something and making it a part of oneself, such as in the *ownership of an idea or a poem.*

Other old words for which new senses are not given include *mosey*, 'to move along leisurely or aimlessly' with the *OED* only listing the original meaning of 'to hurry along' (but both meanings appear in all four dictionaries); the slang meaning of *cowboy* as a verb, meaning 'to murder recklessly and openly' (*ND of AS*); *futon*, an updated meaning of a Japanese mattress with a slatted wooden base which can be converted into a sofa for day use in the US (*OD of NW*); *street people* (*RHII*, *WNWD*), to mean homeless people who live on the streets (not as the *OED*'s older meaning of people who live on the streets, especially to protest against the values of society); and *home*

boy, without the new extended definition of someone who is as intimate and dear to one as a person from one's own hometown (*W9th*, with only the older literal meaning; *OD of NW*), though presenting *home boy*, *home girl*, and *home people* as slang expressions in a 1967 *American Speech* quotation about its use at Southern African-American colleges. Moreover, shortened forms, such as *homey* and *homes*, and the synonym *homeslice* are only listed by *OD of NW*.

3.3 Updated quotations needed

In informing readers of how old words are used in new ways, the *OED*'s updating of quotations would help.

For many words and senses of words, the latest quotation cited is a nineteenth-century example. For example, the *OED* gives four quotations for *humankind* (cited in the four dictionaries), the first recorded in 1645; the last in 1860. Without quotations in the twentieth century, a reader would not know that this word has currently replaced *mankind* in some circles, especially feminist or liberal religious groups. Similarly, the reader would not know, without current quotations, that *working class*, the British euphemism for *lower class*, is now used in American reports and in dialect studies (the four dictionaries). And for *downsize* or *downsizing*, the *OED*'s numerous quotations are largely about automobiles, with one quotation about a remark being *downsized*. Quotations need to be cited on downsizing many other matters — style of living — clothes, shoes, homes — and portfolio stocks, etc. (an updating is also needed in the four dictionaries).

3.4 Sharpened definitions needed

Besides updating the newest meanings of old terms and the updating of quotations for new uses of old words, the *OED* should sharpen its definitions of some of its words for clarity.

Although the *OED* defines *carpet-bagger* adequately in its denotative sense as a person interfering in politics with 'no permanent or genuine connexion' it does not match *RHII*'s description of an 'opportunistic or exploitive outsider' or *WNWD*'s mention of the contempt and resentment felt by the local populace. Another example is the *OED*'s definition of *sisterhood* as a relationship of women with a common aim or common characteristics and calling, a term used in feminist descriptions, with quotations which have derogatory or ominous implications of the word: 'sisterhood's demonology' and '... a portent of what the sisterhood is now brewing up.' The definition should include positive meanings, such as a feeling of oneness and closeness in mutual understanding and support among each other, followed by illustrative quotations of the positive meaning (only *WNWD* gives this positive view).

The *OED* also needs to refine these definitions: *recycling* to include the reuse of discarded materials, such as bottles, cans, and paper (the four dictionaries, *OD of NW*) — not only the recycling of organic wastes or the recycling of industrial processes; and *demoralize* to mean not only the lowering of morale especially because of armed force but also because of any trying adversity (the four dictionaries).

Another problem with the *OED*'s definitions is that the definitions concerning learned or technical matters would not be understood by the novice, but only by persons already familiar with the field. For example, although *deconstruction* and *deconstruct* are listed in the *OED*, the definition as using the method of Jacques Derrida in critical analysis of language and literature is arcane to most people, unlike the excellent definitions by *RHII* and *WNWD*. And definitions for new grammatical terminology are entirely inaccessible for the uninitiated: e.g., *deep structure* and *surface structure* (also incomprehensible for the novice in the four dictionaries); *case grammar* (*RHII*, but also opaque for the novice); and *homorganic* (*RHII*).

3.5 Social sensitivity needed

Another problem with the definition of terms is that a reader socially sensitive to changes in world view would be displeased with language once accepted, but now considered offensive.

The *OED*'s second edition has tried to be more socially sensitive but falls short, for example, in its definition of *canoe*. Its first edition states that it is a 'rude craft' used by 'uncivilized nations' or 'uncivilized people' and 'savages'. In contrast, the second editions describes a canoe as a 'roughly-made craft' used by 'primitive societies'. But who is to say that these crafts are 'roughly made'? Canoes that traveled all the way from Samoa to Hawaii in rough waters, for example, cannot be just 'roughly made'. Nor is 'primitive societies' acceptable. Substitutes now acceptable for *primitive* are *pre-literate*, *pre-industrial*, or *nontechnological*.

3.6 Usage problems

Not only is the *OED*'s usage outdated as shown above, but some of its usage LABELS are also outdated, and there is an absence of usage notes or labels where they should appear.

Some of the usage labels are incorrect. For example, the *OED* indicates that *blurb* is slang, whereas the four dictionaries have no usage label for this term. The *OED* also labels *misspeak* as obsolete, although I have heard it used by my department chair and a young engineering professor. In comparison, *W9th* and *AHD* do not cite *misspeak* at all, but *RHII*, *WNWD*, and *WIII* cite this word with no status label indicating obsolescence. Another word labeled by the *OED* as obsolete is *disinvite*, although columnist George Will in one of his columns (1991:A8) spoke of Linda Chavez being disinvited or of

certain groups disinviting her because of political correctness, using the passive voice of the verb twice and its present participle once. In comparison, only *RHII* of the four dictionaries lists this term, and it bears no status label; *WIII* lists this term, but marks it as obsolete.

Moreover, usage labels or usage notes do not appear where they should. For example, there are no usage notes for *mankind* and *brotherhood* to indicate that they are old fashioned and offensive to some people if they refer generically to both males and females. In comparison, *W9th*, *AHD*, and *WNWD* also have no usage notes for *mankind*. But *RHII* does say, 'See *-man*,' which states, with extensive examples, that compounds with *-man* have declined in recent years, though it does not cite *mankind* specifically. One would have to look at *WD of EU*, which cites alternatives to the word if it is offensive to the audience: alternatives such as *humankind*, *human beings*, *humans*, and *people*. Similarly, for *brotherhood*, cited in the *OED* in the phrases *brotherhood of man* and *universal brotherhood*, no usage note appears. In comparison, there is also no usage note or label for *brotherhood* in the four dictionaries nor in *WD of EU*.

3.7 A systematic method needed in listing and defining terms for reader convenience

Another problem when looking up words is the lack of a systematic method in listing and in defining terms.

At times, related terms are not listed. For example, the *OED* lists *AIDS* as an entry, and in its definition, cites *HIV* as the cause, but there is no separate entry for *HIV*. In comparison, the only one of the four dictionaries which lacks a separate entry for *HIV* is *AHD*. Similarly, the *OED* lists *pro-life* and *pro-lifer* as one entry with the request to look under *pro-*, and in a 1979 quotation for *pro-life* mentions *pro-choicer*, but there is no separate entry containing both *pro-choice* and *pro-choicer*, unlike *W9th*, *RHII*, and *WNWD*, which list *pro-choicer* under *pro-choice*. *AHD* lists *pro-choice* with no mention of *pro-choicer*.

Another problem in listing terms is that the older form of a word may be used rather than its more familiar, current form: e.g., the *OED* does not cite the clipped form *blush*, meaning 'rouge', as an entry at all, but lists only its older, fuller form, *blusher* (this is also true of *W9th*, *AHD*, and *WNWD*; *RHII* lists *blush* as a noun, but asks the reader to look under *blusher*); and the *OED* does not cite *bag lady* (cited by the four dictionaries), today's familiar clipped form, but lists the compound as *shopping-bag lady* under *shopping* as an attributive.

Another difficulty is that some phrases are not entered as compounds in bold print, as the *OED* often does before citing them in the illustrative quotations. Thus, the reader must plough through the illustrative quotations to find *primal scream* (*RHII*) and related compounds under the psychological subject heading for *primal: primal*

screamers and *primal therapy* (*RHII*, *WNWD*). (*W*^{9th} and *WNWD* list, in addition, *primal scream therapy*). Nor is the popular sense of *primal scream* given, for example, a 'loud, primitive scream venting one's pent-up frustrations for catharsis', as used on college campuses on an agreed-upon hour, such as midnight, during examinations; only the technical psychotherapeutic definition is implied in the quotations on *primal therapy* (like *W*^{9th}, *RHII*, *WNWD*).

Moreover, *highlighter*, meaning 'a marking pen highlighting words or passages', is confusingly listed immediately under the second sense of the verb *highlight*, meaning 'to tint or bleach portions of the hair', with 'hence *highlighter*', rather than under the verb's first sense, 'to bring into prominence ... or draw attention to'.

Also for improved clarity a parallel method of defining related terms is needed: e.g., *deep structure* and *surface structure*. For *surface structure*, the term is defined immediately when it is listed under the attributive combinations of *surface*; then one may look at its use in a quotation cited from Chomsky's *Current issues in linguistic theory*. For *deep structure*, on the other hand, no definition immediately follows its listing as a compound of *deep*. Instead, there is an admonition to 'see quot. 1965, a quotation from Chomsky's *Aspects of a theory of syntax*.'

One thorny matter which also needs resolution is the way compounds are listed, either under the first word or as a separate entry, a situation which confuses the reader about where to look first for the compound. For example, *camp meeting* is an entirely separate entry, not listed under *camp*. On the other hand, *human rights* is listed under the word *human* in italicized bold print in the section on compounds with *human*, and *surrogate mother* is found only in quotations for *surrogate* without *surrogate mother* listed in bold letters and with no separate section listing compounds for *surrogate* before the quotations. Some logical rationale needs to be devised for the citing of compounds.

4.0 Conclusion

Fortunately, the *OED* editors in the prefatory materials acknowledge that full revision and updating are long-term goals and that the great accomplishment thus far has been a merger of the first edition with the supplements of the last two decades using computer technology (xi). Moreover, the editors acknowledge many of the problems cited in this study (lv-lvi). Besides, the critical examination in this study is not meant to be a disparagement of the *OED* but only a lover's quarrel with this venerable treasure of information, to which we have been fondly attached and from which we have grown to expect high standards nearing perfection, if not perfection itself. Yet the task for revision is mammoth, as we have seen, and more so, because the *OED* now seeks to include not only the common literary and colloquial words, as in the first edition, but also more of the

'main technical vocabulary and a large measure of dialectal usage and slang' (vii). In addition, the *OED* seeks not only to expand the coverage outside of the United Kingdom, but also especially North American English, which the editors say, 'is the greatest source of linguistic change' (lvi), along with the English of other parts of the world so that 'the *Oxford English Dictionary* may continue to be an accurate and comprehensive register of the whole vocabulary of English' (lvi). A tall order!

If this venerable dictionary is to retain its scholarly reputation, it must use other procedures to glean the data necessary to make emendations for the seven areas named above.

REFERENCES

ALGEO, John. 1990. The emperor's new clothes: The second edition of the society's dictionary. Transactions of the Philological Society 88:2.131-150.

THE AMERICAN HERITAGE DICTIONARY. 1991. 2nd college ed. Boston: Houghton.

CARVER, Craig. 1987. American regional dialects: A word geography. Ann Arbor: University of Michigan Press.

CHAPMAN, Robert. L. 1986. New dictionary of American slang. New York: Harper & Row.

GRAY, Paul. 1989. A scholarly Everest gets bigger. Time, 17 March, pp. 95, 98.

HARVEY, A. D. 1991. Road to rot: The new 'Oxford English dictionary'. Contemporary Review, August, pp. 84-86.

LAKOFF, George, & Mark Johnson. 1980. Metaphors we live by. Chicago: University of Chicago Press.

THE RANDOM HOUSE DICTIONARY OF THE ENGLISH LANGUAGE. 1987. 2nd ed. unabr. New York: Random House.

TOFFLER, Alvin. 1971. Future shock. New York: Bantam Books.

TULLOCH, Sara. 1991. The Oxford dictionary of new words: A popular guide to words in the news. New York: Oxford University Press.

WEBSTER'S DICTIONARY OF ENGLISH USAGE. 1989. Springfield, MA: Merriam-Webster.

WEBSTER'S NINTH NEW COLLEGIATE DICTIONARY. 1991. Springfield, MA: Merriam-Webster.

WEBSTER'S NEW WORLD DICTIONARY. 1988. 3rd college ed. New York: Simon and Schuster.

WEBSTER'S THIRD NEW INTERNATIONAL DICTIONARY OF THE ENGLISH LANGUAGE. 1971. Unabr. Springfield, MA: G. & C. Merriam.

WENTWORTH, Harold, & Stuart Berg Flexner, comp. and ed. 1975. Dictionary of American slang. 2nd supp. ed. New York: Crowell.

WILL, George. 1991. 'PC'? Never heard of it. Commercial Appeal [Memphis, TN] 22 October, p. A8.

Studies in the Linguistic Sciences
Volume 23, Number 1, Spring 1993

LANGUAGE PLANNING ACROSS POLITICAL BOUNDARIES: A CASE STUDY OF PULAAR*

Abdul Aziz Diop

Much of the previous literature on language planning and language policy (LP/LP) focused attention primarily on the discussion of the associated issues within national boundaries. In this study I discuss LP/LP at the supra-national level. I examine factors that can impede language planning (LP) across political boundaries in Africa, focusing on the case of Pulaar (Fula). I make suggestions towards the development of standard Pulaar and consider some of the problems underlying such an enterprise — problems such as those caused by language loyalty, (language) nationalism, power conflict, and language maintenance. I suggest ways of dealing with (some of) these problems. At the end of the paper I indicate directions for future research in the area of LP/LP for Pulaar.

1. Introduction

1.1. Preliminaries

Language planning, whether carried out consciously or not, is a secular enterprise. Yet it remains one of the areas where human beings have probably been the least successful. Whenever an institution or a body of institutions chooses a language (or variety thereof) among others to serve as the medium of communication (in all of its aspects) between speakers of different languages or dialects, we have a case of language planning (Whiteley 1972). However, as Tollefson (1991) mentions, there has been 'a failure to relate language planning and policy to broader sociopolitical changes, and ideologies and political values passing for theoretical frameworks prevail'.

Language planning has become, through the ages, more of the domain of politics than of linguistics. Haugen (1966a:26) stated that 'language planning, like politics of which it is a part, is the art of the possible'. This statement is certainly true of LP but unlike the politician's goal the language planner's goal is substantially the same as that which the people have unconsciously accepted as their own. Thus, LP, unlike politics, is not the art of preventing people from meddling in their own affairs.

Much of the previous literature on LP/LP (Fishman & Gupta 1968, Ferguson 1966, Jernudd & Rubin1971, Haugen 1966a, Whitely 1966, etc.) focused on LP at the national level, i.e. within one country or state; and that was pretty much the general paradigm. Of the few

studies about LP/LP at the regional level one can cite the recent study by Berns (1992) which focuses on the use and status of English within the European Council (EC) member countries. Berns points to the fact that '... the special status of English as a language of wider communication among Europeans is recognized in its use, with French as the language of Council publications and documents'. This raises, however, a number of concerns on the part of Europeans who speak languages other than English. I discuss those concerns as they relate to Fula in section 4.2 below.

The lack of studies on LP/LP across national boundaries or at the regional level seems to suggest that LP/LP must be confined to national boundaries. If this is correct, the question must be raised why this is the case. What factors, for example, impede LP at the regional level? Can such difficulties be overcome and what benefit would accrue to the nations concerned? To answer these and other questions we must, first of all, understand the problem and opportunities presented by the language under consideration here, Pulaar.

1.2. The problem

Pulaar is a West Atlantic language (Greenberg 1955) spoken in approximately seventeen to twenty countries in Africa. Though the figures may vary, the most recent ones available indicate that there are 10-15 million native speakers (Crystal 1987:438) speaking many different varieties of the language, some of which are almost mutually unintelligible. While certain (neighboring) countries, for example Senegal, Guinea, Mali, and Mauritania, have had some common politico-cultural ties, such ties do not necessarily exist between any two (or more) given countries where Fula is spoken (e.g. between Mali and Ethiopia). The popularity of Fula offers numerous opportunities as well as challenges. One such opportunity (or challenge depending on how one looks at it) is the possibility of serving as a lingua franca or language for wider communication at a regional level much like Kiswahili in East Africa or English within EC countries in Europe. A discussion of this problem in terms of LP is called for both for practical and theoretical purposes. There are classic problems that LP/LP have always faced. Four of those will be mentioned in this paper because they surely would constitute a major problem for Pulaar planners. Three of them are sociopolitical: (language) nationalism, language loyalty, and (international) power conflict; and the fourth one is linguistic and has to do with problems of maintaining the variety that is chosen as standard and making sure that it would not undergo further atomization. Before discussing these issues let me first outline the objectives of this paper.

1.3. Objectives

The main objective of this paper is to argue for the necessity of a cross-national language planning and language policy for Pulaar in

Africa, a language policy that consists first and foremost in standardizing the language at the regional level. In fact the grounds for a similar proposal were established as far back as 1965 following the creation of 'la Société Linguistique d'Afrique Occidentale', one of the objectives of which was 'to study the Peul language and culture'. Another initiative with respect to Pulaar was taken in 1986 following meetings in Yaounde (Cameroon) and Niamey (Niger). One of the goals outlined in those meetings was 'la promotion des langues à vocation régionale (languages for wider communication): Peul, Manding' (CONFEMEN 1986).

A secondary goal of this paper is to examine the feasibility and desirability of undertaking LP/LP at a supranational level. The discussion of these objectives will lead to the exploration of other issues of both practical and theoretical importance.

The paper will to some extent draw on the theoretical model of language standardization of Albanian discussed by Byron (1976). The similarity between Albanian and the case under study is that in both cases we are dealing with selection (of a standard) among alternates of the same language. However, while in the Albanian case we are basically faced with two dialects, Tosk and Geg, in the case under study we may be dealing with many more dialects.

The Albanian case was presented in Byron 1976 from two different angles: On the one hand opposing views are presented as to the criteria for selection of the Tosk or the Geg dialect; on the other hand, subsequent views center around the issue of what the base component of standard Albanian is; i.e. whether it is one dialect, or the other, or the combination of both.

The paper will also try to bring into the discussion the experiences of other African countries, including that of some of the so-called Francophone countries in relation to 'Francophonie'. Before doing so, let us, first, take a look at the various contributions made to LP/LP.

2. Review of literature

Even though there is a dearth of literature on LP/LP for Pulaar, the issues of LP/LP have been investigated quite thoroughly by a number of linguists around the world.

2.1. General literature on LP/LP

The term language planning refers to the organized pursuit of solutions to language problems, typically at the national level (Jernudd & Gupta cited by Fishman 1972).

'Language planning is an area of linguistics which allows us to consider language as one more object of human manipulation — not only by language specialists but also by lay persons who may change its basic structure through their attitudes and myths about the language, and their subjective reactions to language' (Rubin 1973).

Language planning has also been defined as 'the conscious, predictive approach to changes in language and language use' (Jernudd & Rubin1971). Language planning could be the result of a formal or informal process. The latter consists of a de facto assignment of certain functions, positions, or statuses to certain languages in a multi-lingual community. The first is the result of a deliberate and conscious effort on the part of decision-making bodies in a particular country — an effort geared towards promoting a certain language or certain languages (and, thereby, demoting others?) to a certain status in order to fulfill certain specific functions.

It has been suggested in the sociolinguistic literature on LP/LP that formal language planning encompasses two major steps which can be further analyzed into substeps. The first step is referred to as status planning. As initially suggested this step is concerned with decision-making generally by a government body and often at the national level; it is the political phase of language planning. It involves three different but related subphases:

- the perceived need to elevate one selected language as the national and/or official language of communication;
- an attitudinal survey in order to identify attitudes of individuals vis-à-vis the language(s) selected;
- a collection of demographic data on the languages.

The second step, called corpus-planning, involves the actual planning/articulation of the language. This is the linguistic phase of language planning. Four different but not mutually exclusive substeps can be identified within this step:

- identification and selection of a particular language or dialect (notice that this stage is very similar in essence to the third subphase of the political phase);
- codification, a substep involving the adoption of an alphabet. Thus it is often referred to as graphicization. Codification also consists of selecting lexical items wherever there are some that are competing, writing of grammars, compilation of dictionaries, etc.
- implementation: or the adoption of what has been proposed;
- elaboration of the language functions and development of new registers.

While the subphases of the political stage of language planning can be followed with 'relative' ease, the four substeps of the corpus planning (i.e. identification, codification, implementation, and elaboration) are highly theoretical and harder to apply to language planning in real life. Proper application of language planning would certainly require the right kind of information about the sociolinguistic habits of the target population and about the social basis for language policy in order to 'project productive directions of change' (Rubin 1973).

Another issue that a number of prominent sociolinguists deal with in the literature on LP/LP is the question of the criteria for evaluation (e.g. Ray 1963, Haugen 1966b, Neustupny 1970). Ray (1963), for example, proposes three criteria: efficiency, rationality, and commonalty. The criterion which is directly relevant to the case under study in this paper is that of efficiency as it refers to the relative cost, in time and other resources, of learning and maintaining one form (for that matter, perhaps one dialect) as opposed to another. However, this is not to say that the other criteria are not important at all. They certainly are, but perhaps to a different degree. The second criterion, rationality ('adequacy' in Haugen's terms (1966b)) is certainly an important one in that it has to do with the capacity of a language to function in a wide variety of styles, genres, levels of discourse, etc. Neustupny (1968), on the other hand postulates four national needs which, in his opinion, will determine the ultimate fate of the lingua franca. These are:

- raising the standard of living
- extension of literacy
- development of national consciousness
- cooperation with neighboring nations.

These needs could in fact be regarded as criteria of evaluation to some extent — in that language planners would have to make some predictions and anticipations as to what dialectal variety would be capable of achieving such ideals with the 'minimum' amount of sacrifice.

Another set of criteria often discussed in the literature has to do with the issue of whether the standard should reflect the speech of an elite, i.e. a minority, or that of the majority. In the case under study it is not quite sure whether this is a relevant issue at all, given the multitude of dialects and the widely varying experiences from one country to another.

Although the three linguists mentioned above use almost the same terminology to discuss the notion of criteria of evaluation, they all disagree with respect to what each criterion encompasses. What these disagreements reveal is the fact that the relative role of linguistic vs. non-linguistic factors in language planning is not entirely clear. What, for example, are the 'patterns' and 'features' which Neustupny assures us constitute the field in which evaluation functions? Are phonological problems as tangential to language planning as some linguists (including Ray) contend? Does usage precede evaluation, or vice versa? These as well as other issues are still unaddressed.

Given the outline that I just presented of the general literature on LP/LP and on criteria for evaluation, let us now briefly discuss LP/LP as it relates to the African continent.

2.2. LP/LP in African societies

There is a large body of literature on LP/LP in Africa. Among some of the most well-known I can cite Whitely 1966, 1971, Fishman 1972, Bokamba 1984a, 1984b, 1976, Bokamba & Tlou 1977; Labouret 1952, Sow 1977, Bamgbose 1991, Gnalibouly 1988, Ka 1983, etc. to name a few. To this one can add numerous conferences by the UNESCO and the ACCT. Spencer (1971) gives a historical perspective of what 'led' to LP/LP in Africa — tracing contacts with Europeans that date back as early as 1445. He gives us an analysis of French/Portuguese policies in contrast with British/Belgian policies with respect to attitudes towards the so-called 'vernacular' languages. The former was that of indifference and restrictive assimilation, while the latter displayed a paternalistic flavor. According to Bokamba (1984b:2) French was imposed because of:

- politico-cultural strategies;
- African intellectuals' wish to 'be educated' the way people are educated in France;
- various local linguistic factors.

The French colonial language policy reached its highest stage when the metropolitan ordinance of Villers-Cotteret, which was issued in 1539 by King François I, was extended to the colonies in order to ensure that French must be the only language used in schools by students.

Despite apparently different philosophies of imperialism, these cases share the similarity that LP is the result of a colonial policy whose very foundation rested on a theory of 'separate development' for the different races in contact with Africa. Sow (1977) takes us back some three to four decades to trace the first initiatives taken for the creation of the West African Linguistic Survey in 1956, followed by the creation of 'la Société Linguistique d'Afrique Occidentale' (SLAO). Under the auspices of these organizations a number of initiatives were taken with respect to LP/LP at different meetings. One of the objectives of the organization was to study the Fula culture and language. The 'SLAO' came up with recommendations following a meeting in Accra on April 12, 1965. These recommendations concerned the implementation of an alphabet and an orthography for Pulaar. Recommendations were also made with respect to various aspects of implementation of the language, but Sow does not make it clear which dialect was selected and elevated to the status of standard. That issue is as unclear as the question of what the procedures were for choosing one variety. For the sound transcription Sow goes on to mention that '.... il en est resulté l'adoption, apres vote, d'un systeme qui peut se definir comme suit... ' (Sow 1977:101). Thus decisions made with respect to that particular issue followed a vote rather than empirical/scientific research. One cannot help but question the validity of such an enterprise.

Following a 1979 meeting by the UNESCO in Bamako (Mali), a recommendation was made to the effect of promoting the languages spoken across national boundaries as a means of inter-African communication and their use for administrative and economic purposes. That meeting was notorious for its lack of specifics and, as a consequence, there was no follow-up. The OAU Language Plan of Action for Africa was also one such attempt at planning in Africa. The plan proposes five specific goals, the first of which was to promote 'the use of certain viable African languages at national, regional, and international levels as official languages in place of non-indigenous official languages currently being used, and the adoption of such languages as working languages by national, regional, and continental institutions [....]' (Bamgboṣe 1991:127). This plan, in its internal anatomy (cf. Bamgboṣe 1991:125), has departed from the traditional ill-defined, pitifully vague pronouncements; but so far, it has failed to change attitudes of governments and policy makers in favor of realizing the projected goals.

Other recent initiatives outlined by UNESCO include its 1985 periodical on 'La définition d'une stratégie relative à la promotion des langues Africaines'. Those initiatives were taken after a meeting of experts on African languages in Port-au-Prince, Haiti in 1975. Following that meeting two conferences were held in Yaounde (Cameroon) and Niamey (Niger). During those conferences a number of goals were outlined one of which was to promote languages that have 'une vocation régionale', Pulaar and Mandingue.

Another study that dealt with the issues of LP/LP in Africa is published in collaboration with L'Agence de Cooperation Culturelle et Technique. Proposals that appear in that review are made following conferences of Secretaries of Education in various French West African States. For work in the area of LP/LP directly affecting Pulaar one can cite the Pulaar-Russian-French dictionary by Zoubko (1980); the ACCT's specialized lexicon that deals with teaching the natural sciences, grammar and linguistics, history and geography, politics, administration, and justice; the recent publication of an English-Pulaar dictionary by a Japanese linguist (Eguchi 1986); and finally, a UNESCO-sponsored dictionary of elementary Pulaar-French-English. Future projects of the Pulaarophones from Paris include helping researchers from France, England, Japan, and various countries in Africa gather the data necessary for making an interdialectal dictionary, dialectal dictionaries, and bilingual interdialectal dictionaries. Unfortunately, as pointed out by Gnalibouly (1988), the initiative is not progressing satisfactorily.

It is evident from the preceding discussion that the issue of LP/LP in Africa has changed very little. As a matter of fact, as mentioned by Bamgboṣe (1991), several African governments appear to employ avoidance or vagueness techniques, as can be illustrated by the fact that very few countries have definitive statements of language policy. But avoidance is in itself policy. An example of a

vague policy cited by Bamgboṣe (1991) is Kenya's decision to adopt Swahili as its national language. The immediate motivation was political. The vagueness is reflected in the implementation steps recommended, the details of which were not given. In addition, the country's official language in which records are kept and administration conducted is still English. To this one can add cases (e.g. Nigeria) where a policy is declared without the possibility to implement it due to improper circumstances for implementation, built-in escape clauses that give excuses for not implementing the policy in question, or unspecified procedures. To all of these problems related to language planning in Africa one can also add problems caused by fluctuation in language policy due to such factors as changes in government or party policies, advice from various foreign organizations, etc. Ghana is a case in point for such fluctuations (cf. Bamgbose 1991 for further discussion).

In the next section I will try to give a sketch of some of the most crucial and specific problems underlying the development of a language policy for Pulaar, based on what has already been achieved.

3. Language planning in West Africa: the case of Pulaar

At this juncture we should address at least two questions. First, what problems would a language policy proposed for Fula encounter? Second, how would one deal with these problems? Assuming that the problems can be addressed successfully, what benefits would the population derive from the language policy in question? To get a better picture of the situation at hand I give a profile of Pulaar.

There are many names used to refer to this language: Fula, Fulani, Fulfulde, Pulaar, Haalpular, Toucouleur Fulacunda, Fellata, etc. In this paper I shall refer to it as Pulaar or Fula. Pulaar is a West Atlantic branch of the Niger-Congo family, which is in turn member of the larger family of the Congo-Kordofanian (cf. Greenberg 1955). It is spoken in at least the following countries in West and Central Africa: Benin, Burkina Faso, Cameroon, the Central African Republic, Chad, Gambia, Guinea, Guinea-Bissau, Mali, Mauritania, Niger, Nigeria, Senegal, Sierra Leone, Togo, and to a lesser extent in Sudan and Ethiopia.

Because of differences in geographical distance/proximity it should come as no surprise that certain dialects of Fula are more mutually intelligible than others. Some of them are almost heading towards mutual unintelligibility. Not only do we notice considerable dialectal variations among its different regional dialects, but also within each country where it is spoken there is some kind of dialectal 'atomisation' that is worthy of notice in some cases (e.g. the difference between Pulaar as spoken in Northern Senegal and that which is spoken in the Casamance (southern) region of the same country).

While it is clear that HaalPulaar'en (native speakers of Pulaar), who were for the most part nomad cattle herders, have migrated over centuries all over the African continent, settling wherever they could take advantage of rain and green pasture for their cattle, it is not yet quite certain where they exactly departed from. However, Ka (1983) quoting Hama (1968) who himself quoted Delafosse (1912) seems to support the thesis that Haalpulaar'en originated from somewhere between Mauritania, Senegal, and Mali. The issue is moot for the purposes of this paper and I do not intend to pursue it further.

A number of countries where Pulaar is spoken have been or are currently involved in various attempts to implement concrete guidelines for LP/LP (Niger, Senegal, Mali, Mauritania, Guinea, etc.). The 'PROJET MAPE' is a materialization of such attempts. This project is primarily intended for LP/LP of a number of languages in that particular region of Africa. In the case of Pulaar a number of countries are using the language (still at the experimental stage) both for purposes of education (school curricula) and in the media (radio and TV); cf. the countries just cited in addition to Egypt for radio. However, there does not seem to be any literature that shows where the language fits into the overall framework of LP/LP in other countries, such as Sierra Leone, Sudan, etc.

The language is 'divided' into two major geolinguistic groups: the Eastern Dialects and the Western Dialects. There are some straightforward structural differences between the two.

One straightforward structural difference between these regional dialects is the fact that although all the dialects of Pulaar exhibit stem-initial as well as stem-final consonant alternations, some dialects show more alternations than others, especially in word-initial position (perhaps alternation is avoided for the purposes of 'simplification'). To illustrate, consider the following example. In the Westernmost dialects (e.g. Mauritania, or part of Senegal), the verbal stem *yah-* 'go' exhibits the following alternation (initially) in certain contexts such as plural (1b) or subject pronoun post-position (1c) below:

(1) a. mi yah+ii Kahaydi
1sg. go+asp
'I went to Kahaydi.'

b. be njah+ii Kahaydi
3sg. go+asp
'They went to Kahaydi.'

c. njah mi ko Kahaydi
go 1sg. foc.
'It's Kahaydi I went to.'
Or: 'I went to **Kahaydi**.'

In some of the dialects this alternation is avoided (most of the time) by preserving the same stem shape *yah* in both contexts (after singular and plural object pronouns).

Another difference that is worth mentioning and that is noted even within the same country (e.g. Guinea) is the distinction between degrees of formality. In Fuuta Jalon the plural for respect is used any time one is talking to an elderly person; failure to do so is considered rude. In 'Haute Guinée', on the other hand, such is not the case. While degree of formality may not be justification for dividing dialects into major groups there is yet another more crucial difference between the two major dialect groups, viz. in the form of the infinitive. In the Eastern dialects the infinitive of a verb is formed by adding *gol* to the verbal root (e.g. *yah-gol* 'to go'), whereas in the Western dialect this is achieved by adding the suffix *de* (e.g. *yah-de* 'to go').

There are other dialectal differences that are worth considering but for the purposes of this paper I will not dwell on them. Besides, they are problematic both for native speakers themselves and for lexicographers working on Pulaar (cf. Gnalibouly 1988:10 for discussion).

As we shall see later on, the issues I have briefly outlined here are going to have to be taken into consideration in the choice of a standard dialect. At this juncture I would like to get into issues pertaining, directly or indirectly, to supranational policies for the development of standard Pulaar.

3.1. The development of a supranational language policy for Fulani

The need in the African continent for urgent language planning strategies is certainly worthy of note. Though this paper is not an account of the history of the African continent I cannot, however, avoid dealing with aspects of the legacy of France to West Africa. French colonial language policies have had the detrimental effect of undermining the importance of promoting 'indigenous' African languages to serve national functions (or international ones, for that matter). There was very little (if any) positive action taken to encourage the so-called 'vernacular' literatures or to standardize and extend 'major' languages; i.e., those that had a considerable number of native speakers and dialects or which had the potential of being languages of wider communication. One of the damages of these policies that Africans are still paying a heavy price for is the soaring illiteracy rate in (West) Africa. It was assumed that only those who read and/or wrote and/or spoke French were educated. Thus, no attention was paid to the idea of educating people in their native languages because, after all, writing/reading in Fula or any other African language was not something that was regarded as crucial for literacy. Emphasis may have been put on French to the detriment of native languages for various reasons but the most important ones are mentioned in Bokamba 1984b:2.

When we are involved in the process of LP one of the questions we should ask ourselves is 'why plan at all?' In order to answer that question we should answer another question, which is that of who is it that we are trying to plan for. Thus, it looks as though the answers that we seek lie within the projected long-term goals that we set to achieve. We may need to plan languages in Africa for various but closely intertwined reasons — some of which I would like to mention here. First, where most European countries may have not more than three or four languages to contend with, African countries may well have more than a hundred languages within the state or country, sometimes belonging to several different language families. In such a situation for people to communicate with one another, recourse must be had to some kind of lingua franca. Second, the pre-independence experience of African countries, in addition to political/artificial boundaries of the post-independence era, prevented any language (or dialects thereof) from assuming a pre-eminent national or supra-national position. Third and last, but not least, political and economic atomization, in addition to climatological conditions (in the case of nomad Fulanis) accentuated territorially based dialect distinctions.

After this brief outline, I would like to continue this section with the strictest ideological and political neutrality and discuss the issue as an outsider. With that in mind here are what, I think, should be six of the various needs for which I feel a standard Pulaar is necessary:

First, for purposes of creating communication between, most importantly, native speakers of the language (promotion of intracultural communication) wherever they might be; and then, between native and non-native speakers. As Cooper (1989) suggests, an understanding of language planning demands an understanding of the social changes which promote it. More and more nomad Pulaars are forced (by climatological conditions) to settle, become farmers, or go to modern schools, thereby interacting and intermarrying with their sedentary 'brethren' who for the most part speak a different variety of Pulaar than theirs. This social change, while capable of bringing about the proper setting for a de facto standard at a local level, cannot do so at a regional level.

Second, extension of literacy; mass literacy, that is. The more educated the people are the more likely it would be for them to take an active role in the development process of their nation. Thus literacy leads towards greater productivity.

Third, 'international unity'. Linguistic diversity is as much a source of cultural strength as a potential source of disunity and politico-economic weakness. Thus, dialects of one and the same language, at least if they threaten to become languages, are potentially disruptive forces for a people that is in search of unity because the dialects appeal to local loyalties. If language

planning within a particular country is a major step toward unification of the people of that country, language planning across political boundaries in Africa ought to have the effect (be it immediate or not) of helping toward greater homogeneity and unity of African nations. (One might claim that this is a naive statement but it is only so when we look at language planning and its desired result as a short term relief operation — which it is not.)

Religion (e.g. translation of the Koran in Pulaar since a great number of Haalpulaar'en are Muslims).

'Deethnicization'. The dialectal atomization of Fula would ipso facto imply a certain degree of cultural divergence as well, which itself may mean that certain Pulaar linguistic groups may evolve into totally separate ethnic groups. Language standardization would, in theory, contribute toward helping native speakers of the different dialects of Pulaar to, in the long run, unite within one speech community.

Language maintenance. This is directly correlated to (and could be a result of) the need outlined above under 'deethnicization'.

We need to plan and standardize and enrich our languages because not doing so would have the necessary corollary that they would, in Fishman's (1972) terms 'forever remain in the intellectual and pragmatic shadows of others that had been fortunate enough to undergo slow but sure enrichment and standardization, both consciously and unconsciously, in prior generations'. In the following section I suggest ways of starting such a standardization process.

4. Suggestions towards the development of standard Pulaar

For the choice of a standard dialect of Pulaar, while we cannot yet establish which dialect to choose, we can bear in mind two things. First, for purposes of learnability the choice of a standard dialect should consider dialects that show the least amount of consonant alternations possible since those alternations are a great problem even for native speakers. Second, for the sake of authenticity the choice of vocabulary (codification phase) should be based on the question of 'which words are the most authentic ones to the language?' as opposed to those that mean the same thing and yet are borrowed from Arabic or French. I am not, however, by any means trying to exclude the possibility of borrowing from other languages (of course, interdialectal borrowing would be a first priority). Jernudd (1989:53) mentioned, in the case of French standardization, that one general goal was 'the maintenance and extension of a standard spoken and written French language purified of unacceptable English language borrowings and local idiosyncrasies'. While I do agree with getting rid of as many local idiosyncrasies as is feasible, it is not quite clear what is meant by 'unacceptable' in the above quote. It is certainly

not my suggestion to proceed with a puristic attitude intended to rid Pulaar of borrowings. They are, after all, unavoidable. Furthermore, when it comes to matters of (language) purism, as noted by Jernudd, there is no such thing as absolute purism; but purism has to be viewed vis-à-vis the challenging language. It is not quite sure whether there is a particular language that can be said to constitute a challenging language to Pulaar in the same way English is to French. The nature of foreign borrowings depends on such factors as the language that Pulaar comes in contact with (Wolof, Hausa, French, Arabic, etc.).

The following is an enumeration of some of the suggestion that I would like to make with respect to the standardization and implementation of Pulaar. We need:

- to (first and foremost) gain greater insight into the role of Pulaar in any context in the continent of Africa, using an analytical framework referred to as a sociolinguistic profile (cf. Ferguson 1966 for an outline). This insight will not only document attitudes toward the different dialects of Pulaar, but it would also help, as pointed out by Berns (1992), make important decisions that have to do with curriculum development, material design, and the setting of goals and expectations;

- to conduct an attitude survey before we set to decide on a standard variety, followed by descriptive studies and pilot projects;

- to combine, after the first step is taken, some of the various dialects on a rational and scientific basis, and construct a unified grammar from the compounded dialects without necessarily oversimplifying (and therefore trivializing) that grammar. After this step is taken let us then be concerned with the following:

- to work with people who are not necessarily professional linguists but who should participate in the regularization of the language through their own specialties, e.g. law, economics, agriculture, writing, etc. After all, they, also, are language 'cultivators';

- to create primary and secondary school facilities as well as incentives for going to school. Initially these would be the only places where the standard would be taught;

- to create terminology commissions (e.g. a supranational terminology bank) the purpose of which will be to make lists of needed Pulaar terms in various areas of specialization. For example, as mentioned in Gnalibouly 1988 some concepts, though commonly used nowadays such as: feelings, color taste, nature, animal, insect, gas, liquid, or solid, etc. have no satisfactory equivalents in Pulaar. The Pulaar lexicon needs to be expanded to include these and other terms in order for the language to be taken to higher altitudes in the communications

race. The terminology commissions can use a number of techniques for creating new terms, some of which are outlined in Mezei's paper in the case of Somali (1989): using Pulaar roots or derivations, reviving archaic words from classical poetry and songs, semantic expansion, compounding or reduplication, suffixation, etc. The terms would be used in government-supported schools and institutions, in the first place;

- to organize frequent meetings between Fulanists and, say, Kiswahilists in order to exchange ideas and to learn from each other's experience;

- to materialize initiatives such as that taken by the Pulaarphones of Paris which consists of trying to come up with a General Dictionary of the Pulaar World. Such a piece of work (I do recognize, it would take tremendous effort to realize) is vital for the Pulaarophone world, in the absence, for the time being, of a standard dialect;

- to create committees (in each country involved) in charge of supervising the creation (and co-ordination with various other countries) of new technical terms and promoting linguistic interaction. When things go wrong those committees will determine what went wrong and who is to be held responsible for that;

- to draw some useful conclusions from the Guinean experience. Guinea (cf. Fall 1981 for more discussion) is one of the few countries in (West) Africa to develop and implement a national language policy in its early independence years whereby the country was linguistically divided into geographical areas that reflected linguistic distribution. Pulaar was taught in 'Moyenne Guinée'. Therefore, Guinea may have a lot to offer in terms of pedagogy and teaching experience

Let me mention, however, that all our efforts for planning a standard dialect would be in vain if we do not take a major preliminary step which is outlined in the lines to follow. I suggest that all governments involved take one major step without which nothing can be achieved: promoting language institutes. These language institutes should hire professionally-trained linguists whose job would be to work in ways that would help write consistent, coherent, and well motivated grammars of each separate dialect of Fula. In this way we are going to gain more insights into the structural description of each dialect and therefore, be in a better position for evaluating the dialects with a view to determining which of them to choose. Furthermore, these descriptions would certainly help decide what adequate theoretical framework ought to be used for standardization of Pulaar. There have been some efforts in some countries (Mauritania, Senegal) to promote the existence of language institutes; but, for the time being, they employ many unskilled 'linguists' whose job (though not trivial) is reduced to basic descriptive analyses of the language without any theoretical framework.

Such language institutes would have to work closely together in order to yield the best results. What would then be the implications of such an apparently ambitious enterprise? I address this and other questions in the next section.

4.1. Implications

The significance of language planning to the development of Africa is beyond any shadow of doubt. One of the contributions would be the positive effect that it would have, in the long run, on literacy and education, especially in 'Francophone' West Africa. According to Bokamba (1984b), the illiteracy rate in that part of Africa is one of the highest in the continent. French is still a psychological barrier to the implementation of African languages in Africa. The French colonial language policy is still spreading its effects on the continent in term of the psychosociological behavior of its inhabitants. A lot of people in the continent still feel that being educated in French is more prestigious and is more of a priority than instruction in native languages.

Bokamba (1984b) quoting Fishman (1971) (in an attempt to argue against the latter's typology) outlined three factors determining the type of language policy a developing country may adopt. One of those has some implications for the present study; it is the type c decisions. Type c decisions are those that intervene in a situation where there are conflicting or competing great traditions, particularly in the absence of a superordinate threat. Language planning across political boundaries of the type proposed in this paper has the advantage of helping us avoid certain conflicts (though we all know it would certainly have some backlash effects). I believe that it is possible, though costly, to achieve unity and integration through lessening of linguistic rivalries by using authentic African languages.

One major implication that we should hope to be a corollary of LP/LP for Pulaar is that along with the experience of Swahili and other 'successful' experiences in Africa, the Pulaar experience might contribute toward the creation of major pan-African (or pan-West-African) languages of wider communication whose purpose, among others, is to create major communication networks in the continent and prepare African languages to face the challenging task of expressing the Arts and Sciences — with all the advantages that might encompass.

The objectives, ideals, and projected results and implications outlined here cannot be reached easily. Some of the problems that Pulaar planners may be faced with are outlined in what follows.

4.2. Issues/problems underlying the development of a language policy for Pulaar

There are various problems underlying the development of a language policy for Pulaar or any other language of wider commu-

nication (LWC). One of the most important of these (which is also thoroughly discussed in Berns (1992)) is that of 'outsider' attitude. The fact that Pulaar is a LWC will certainly raise concern on the part of speakers of lesser widespread/known languages who might view any attempt to develop it for standardized use at the regional level as the beginning of the 'death' of their own languages. Berns discusses similar attitudes towards English within the EC countries. Various EC countries have expressed their concern over the spread of the English language (whether it is the British or American variety) in their community in general and in their respective countries in particular. For instance, she mentions the fact that Mr. Jack Lang, France's Minister of Culture, has been quite vocal about English hegemony and '... lack of European identity in the face of the linguistic dominance of English.' One way the Community tries to deal with such attitudes is by exploring various strategies to teach English without giving (non-native) learners the 'impression' that they are being taught a language associated with international communication. Rather, they are being taught a language associated with British Culture just in the same way they could be taught Spanish (associated with the culture of Spain), or French (representing the culture of France), i.e. 'another variety of European culture and society' (Berns 1992:12). What this shows is that people's attitudes are heavily influenced by the status given to the language they are being taught, especially if such language is not their own. Speakers of languages other than Fula may not care too much to be taught a foreign (African) language that has the bold ambition of becoming 'their language of international communication'. Likewise, speakers of different dialects of Fula may resent a particular dialect being imposed on them as their standard. In the remainder of this section I explore these ideas in more details.

4.2.1. (Language) nationalism

(Language) nationalism, to some extent, can be similar to language loyalty. However, while the latter might essentially refer to attitude toward, for example, one language, the former is of a greater dimension and involves the attitude of an entire nation towards a certain issue. (Of course, the obvious question one is tempted to ask is 'what is a nation?' The paper does not intend/pretend to answer this question.) In the case of India, for example, LP involved a choice between totally different languages, thus, totally different identities. In the case under study the choice is that of a variety among competing varieties of the same language. The issue at stake here is twofold.

On the one hand we have an incompatibility between reliance on a foreign language (as expressed in borrowing) and maintenance of (the nation's) self-esteem. The more borrowings we find in the language (from a foreign language), it seems, the more native speakers of that language are going to feel somewhat 'dependent' on the 'donor' language. The issue here, however, is that of whether we can

really do much about borrowings. Sir Samuel Johnson illustrated our helplessness vis-à-vis borrowings better than anyone else. He declared that 'the project for an English Academy failed in the face of what he termed 'the Spirit of English Liberty ...', and that '... sounds are too volatile and subtle for (legal) restraints; to enchain syllables and to lash the wind are equally undertaking of pride' (Johnson, cited in Fishman 1971).

On the other hand, the tension between the requirements of modernization and those of authentification creates a sort of dialectic: The more stress we exert on real authenticity, therefore, the greater the risk of regionalism. The more stress on modernization, the greater the risk of loss of maintenance of self-esteem. Achieving a balance between the two is certainly something worthy of recognition.

4.2.2. Language loyalty

History suggests that human nature is essentially refractory in language matters. Language planning in general involves choosing one variety as opposed to another or other varieties as a norm. To the extent that this is true it gives the people who speak the chosen variety a prestigious position because it raises their status to that of norm-bearers and, therefore, may ultimately, though not necessarily intentionally, give them power. If we agree that language is an expression of personality and a sign of identity and pride, then LP 'should seek a balance between uniformity and diversity of code' (Haugen 1966b:59). Of course that balance is easier talked about than achieved. Thus, it is often the case that those involved in the process of LP adopt conciliatory attitudes towards the other less 'fortunate' varieties.

When we talk about loyalty in the context under consideration, we should distinguish between two kinds of loyalties (or perhaps attitudes):

- loyalty towards one's dialectal variety
- loyalty towards the colonial language.

The first kind has, to some extent, been discussed in the previous paragraph and basically involves attachment to the variety that we grow up speaking and that we are so emotionally and dearly attached to. Thus, choosing another variety would ipso facto undermine the salience of the dialect that we speak. However, typological data of several developed standards (e.g. Albanian) demonstrate that after a certain point in time we are no longer able to claim with certainty and precision that the standard comes exactly from a particular variety; and that is when speakers of the language identify themselves with the standard irrespective of their previous dialectal background. Of course, this is not something that will be seen for quite some time after the implementation of a norm. But at least it is something to look forward to. After all, LP is not an emergency relief

operation. It is my position that dialectal preferences should be suppressed in the interest of important supranational needs. The standard language should be a unifying cultural force; and linguistic unity should imply unity of the nation — whatever that 'magic' word might mean.

The second kind of loyalty, on the other hand, is manifested (for instrumental purposes?) toward a foreign language or languages. Thus, apart from some African supporters of the widespread use of national languages in education and administration during the colonial era in British West Africa, there was a tendency among educated Africans to see in their use the danger that progress for the African peoples and their integration into the modern world would thereby be impaired. This fear still persists. Haugen (1966a) also mentions the case of speakers of Garani who would rather use Spanish for public affairs. Taken from a purely instrumental point of view those attitudes can be somewhat justified. To this array of obstacles one should also add problems that draw their roots from egotistic battles of political leaders, the subject of the next section.

4.2.3. Power conflict/elite closure

'Everything is suffocated if one's own way is not sought and if another nation is blindly taken as a model' (Fishman:1972:8). This quote captures the essence of how decision-makers in the various countries involved in the LP under discussion in this paper might feel vis-à-vis a decision that affects their internal policies in one way or the other, and yet is not totally theirs. Their fear is that of loss of power, prerogatives, and control; control in terms of manipulating decisions. Thus, countries that regard themselves as 'leaders' would not want to see themselves lose the power of decision and have decisions made for them from without as opposed to from within. Furthermore, the differences in the situation in each country and in the philosophy of the government, in part, explain the differences in policy. This is why a policy that works in one country may fail hopelessly in another. To this, one has to add further complications brought about by revolving-door governments which proceed, generally, with reshuffling and ignoring decisions made by previous governments; and this complicates language planning at the supranational level. (However, in the case under study I am assuming language planning under 'ideal circumstances'.)

Another fear (on the part of decision-makers) comes from a perceived 'threat' that such a 'nationalist(ic)' enterprise might be a major step toward 'linking' of the masses to successively higher levels of social and political authority.

In any case the point is that the reasons behind that power conflict could stem from external or internal pressure, or a combination of both. An accurate investigation into the nature of this problem is above and beyond the scope of this paper.

For the present time one of the biggest problems facing linguists who are interested in the project outlined here would be to gain the cooperation of different nations involved; and one of the challenges that they face is to make scholarly publications about LP/LP comprehensible to policy makers so as to assume their constructive participation. This could perhaps be achieved by having follow-up sessions (after conferences) with policy makers and staff members of various ministries of education. Political leaders and decision-makers in Africa sometimes drag their feet when faced with issues of the kind presented in this paper, where Africans are given the option of working towards promoting their languages to fulfill the functions that were primarily assigned to the language of the colonials. Such proposals are often qualified as 'results of leftist movements' in the political jargon. Therefore, LP faces more resistance/reluctance on the part of politicians than on the part of the people whom it will directly affect. However, once we succeed in getting the decision makers involved, the other problem would be a linguistic one; and it would involve a very lengthy process of selection. As I pointed out earlier, in the case of Pulaar, we are dealing with a language that has very many dialects.

4.2.4. Language maintenance

Another problem that is linguistic in nature has to do with maintenance of the variety that is chosen to be standard. Languages are dynamic and have to be so lest they die. Language as an organism (if I may use the biological metaphor) is subject to the same evolutionary development as other organisms. Therefore, the variety of Fula that would be chosen as a standard would certainly undergo some influences. In the paragraphs to follow I shall briefly outline the problem (if it is a problem at all) and try to suggest modest solutions to this particular issue.

Once a variety is chosen and implemented, the next issue is to keep further atomization of the standard from exceeding the boundaries of the 'acceptable and reasonable'. There is no consensus on this issue. Berns (1992) discusses two views on this (with respect to English, at least). One view (Kachru 1985) takes into account multiple and variable standards while the other (Quirk 1985, 1990) proposes to maintain a single internationally intelligible (native) standard. Given the fact that Fula is spoken in a variety of countries, covering a considerable amount of geographic territory, influence over any of its dialects could come from within as well as from without. The former will be the case where other dialects of the language continue to influence the standard because, after all, in their own homes, native speakers are more than likely to use their own respective dialects because they (may) feel (more) comfortable using them. The latter refers to the case where other neighboring languages exert some influence over the new (and therefore vulnerable) variety. In any case, it seems to me that it is more reasonable at this stage to use one

standard of Pulaar for all, instead of allowing multiple and variable standards, because Pulaar is a long way from enjoying the international and 'universal' use that English has enjoyed for centuries.

The third issue concerns attitudes of parents towards their children if they speak to them in the standard medium that they will acquire at school.

No matter what the attitudes of Pulaar speakers are vis à vis standard Pulaar, varieties will emerge. French (and its academy) illustrates this point. As I suggested earlier, a term used in a particular dialect may enter the Terminology Bank and be used legitimately just in case it has already been widely used, and provided it does not affect the 'standard grammar', or pose a serious threat to an existing term.

By and large these are some of the (many) problems with standardization. In order to reduce the impact of these problems, we need to ask: 'Who is it that we are planning for, and for what reasons?' The reasons could stem from short-term as well as long-term goals. Some of these goals might be:

- opening up levels of communications,
- facilitating the publication of texts,
- gaining a sense of national unity,
- mainstreaming of the educational system.

These goals could be more easily reached (and the problems outlined earlier reduced) through implementation of reference materials, textbooks, de facto language academies, perceived social benefits attached to the use of the standard variety, etc. These will not only be tools for furthering and reinforcing mother-tongue education but they also assure maintenance of the standard. Of course, I am aware of the fact that it is much easier to convince governments to spend time and effort to come up with textbooks and other teaching materials than to find ways of attaching benefits to the use of a particular language. That is political in nature and is left to political decision makers.

5. Conclusion

5.1. Directions for future research

Some modest but significant steps have been taken toward LP/LP for the African continent in general and for Pulaar in particular. However, in order to take those steps further we need to convince decision-makers to rely more and more on the collaboration of linguists, without whom LP would be at jeopardy because it would be carried out empirically (at best) rather than scientifically. Pulaarophones have a lot to learn from other Africanists and linguists from other parts of the world. Future research should first of all take the form of thorough attitude surveys of samples that would genuinely represent the major dialects of the language. This is going to be a lengthy but necessary process. After that, Pulaarophones

need to outline clear goals for standardizing Pulaar and design teams (governmental as well as non-governmental) accordingly.

The project would have to involve linguists from all subfields: phonologists, sociolinguists, semanticists, etc. Of course, they alone cannot achieve anything without help from the economists who are qualified to evaluate the cost-effectiveness of such a project; and this takes us back to the first thing to worry about: funding. Without funds this project will never take off, but will remain dormant on paper like many other projects in the third world.

While it is true that some African languages may die over the years (of natural death?) it is not likely that this would happen to Pulaar, given the number of people who speak it as a native language and the number of countries where it is spoken.

African heads of states and politicians in general have always used the word 'unity' as their 'motto'. But just as freedom is not free, unity is not free either. There is a price to pay for both. If unity is a goal that is so dear to us we should be ready to take realistic views, one of which has been mentioned by Bokamba (1984b:27): 'very few linguists and anthropologists encourage the death of languages, but the multiplicity of African languages makes any realist welcome such development for national integration purposes.'

Should our attempts to unify this huge continent not be doomed to failure we have to have faith in the fact that there is a need to group major geographical areas in Africa under 'major' linguistic entities (phase 3 of the political phase of language planning outlined earlier). Once that step is taken, the next equally important step would consist of assigning (or associating) important functions to those languages in the society. This may not be something we can achieve in the immediate future, but if achieved, it would most certainly help Africans realize that their languages are capable of carrying out the functions fulfilled by French (or English, Spanish, Portuguese, etc.). Who knows, it might even be the beginning of the relegation to lower status of French and other colonial languages in the African continent.

Nevertheless, in order to realize most (if not all) of our goals, we need to be aware that it takes efforts from the political scientist, the linguist, and the economist, as well as cooperative participation on the part of the masses to arrive at meaningful results in LP.

Last but not least, efforts to standardize Pulaar should start first with standardizing the name we use to refer to its native speakers: Haalpulaar'en, instead of nonsensical terms such as 'toucouleur', 'Fulani', 'Fellata', etc., most of which not only are created by colonialists but seem to suggest actually separate ethnic groups. It does not help matters when we see in the literature on the origin of HaalPulaar'en many studies that distinguish the 'toucouleurs' from the 'Fulanis' or 'Peulh' (cf. for example Chavane 1985) as if they were

two separate ethnic groups that happened to speak the same language or borrow aspects of each other's culture.

If African decision-makers consciously or unconsciously agree with and maintain the often implied claim that language planning in the form of one language for all is utopian for the continent or its major subregions, they are either accepting, in one way or another, an already existing de facto linguistic imperialism or else they are making the mistake of thinking that LP/LP is a short-term relief operation rather than what it is in reality: a long-term solution for long-standing dilemmas.

NOTE

*The original version of this paper was presented at the 20th ACAL conference at the University of Illinois at Urbana-Champaign. It has been revised several times since then. I benefited from invaluable comments from Eyamba Bokamba and Braj Kachru. To them I would like to express my sincere appreciations. All errors contained in this paper are entirely my own.

REFERENCES

AMMON et. al. 1987. Sociolinguistics: An international handbook of the science of language and society. Berlin: Mouton de Gruyter.

BAMGBOṢE, A. 1991. Language and nation. The language question in sub-saharan Africa. Edinburgh: Edinburgh University Press.

BERNS, Margie 1992. Sociolinguistics and the teaching of English in Europe beyond the 1990s. World Englishes 11:1.3-14.

BOKAMBA, E. G. 1976. Authenticity and the choice of a national language: the case of Zaire. Présence Africaine 99/100.104-143.

-----. 1984a. Language and literacy in West Africa. Annual review of applied linguistics, vol 4: Literacy issues, ed. by R. B. Kaplan, 40-74. Rowley, MA: Newbury House.

-----. 1984b. French colonial language policies in Africa and its legacies. Studies in the Linguistic Sciences 14:2.1-36.

-----, & J. S. Tlou. 1977. The consequences of the language policies of African states vis-a-vis education. Language and linguistic problems in Africa, ed. by P.A. Kotey & H. Der-Houssikian, 35-53. Columbia, SC: Hornbeam Press.

BRETON, R. 1991. The handicaps of language planning in Africa. Language planning: Focusschrift in honor of Joshua A. Fishman, ed. by David F. Marshall, 153-174. Amsterdam: Benjamins.

BYRON, J. L. 1976. Selection among alternates in language standardization: the case of Albanian. The Hague: Mouton.

CHAVANE, Bruno A. 1985. Villages de l'ancien Tekrour. Paris: Karthala.

CONFEMEN. 1986. Promotion et intégration des langues nationales dans les systèmes éducatifs. Bilan et inventaire. Paris: Editions Champion.

COOPER, R. L. 1989. Language planning and social change. Cambridge: Cambridge University Press.
CRYSTAL, D. 1987. The Cambridge Encyclopedia of Language. Cambridge: Cambridge University Press.
DUMONT, P. 1983. Le français et les langues africaines au Sénégal. Paris: Agence de Coopération Culturelle et Téchnique.
EGUCHI, P. 1986. An English-Fulfulde dictionary. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
FALL, Aram 1981. La politique de promotion des langues nationales dans cinq états de l'Afrique occidentale: Gambie, Guinée, Mali, Mauritanie, Sénégal. UNESCO: La définition d'une stratégie rélative à la promotion des langues Africaines. Documents de la réunion d'éxperts qui a eu lieu à Conakry, Guinée, du 21-25 septembre 1981, pp. 192-218. Paris: UNESCO.
FERGUSON, C. 1966. National sociolinguistic profile formulas. Sociolinguistics: Proceedings of the UCLA Sociolinguistics Conference, 1964, ed. by W. Bright, 309-315. The Hague: Mouton.
FISHMAN, J. A. 1971. Sociolinguistics: a brief introduction. Rowley, MA: Newbury House.
-----. 1972. Language and nationalism. Two integrative essays. Rowley, MA: Newbury House.
-----. 1974. Advances in language planning. The Hague: Mouton.
-----, & Gupta. 1968. Readings in the sociology of language. The Hague: Mouton.
GNALIBOULY, B. 1988. Problèmes de léxicologie et de léxicographie: le cas du Fulfulde. Séminaire régional de l'UNESCO sur les problèmes de léxicologie en Afrique. Niamey, Niger, MS.
GREENBERG, J. 1955. Studies in African linguistic classification. New Haven, CT: Yale University Press.
HAUGEN, E. 1966a. Language conflict and language planning. Cambridge, MA: Havard University Press.
-----. 1966b. Linguistics and language planning. Sociolinguistics, ed. by W. Bright, 50-71. The Hague: Mouton.
JERNUDD, B., & J. Rubin. 1971. Can language be planned? Honolulu: University Press of Hawii.
-----, et al. 1989. The politics of language purism. New York: Mouton de Gruyter.
KA, F. 1983. Le Pulaar en Mauritanie. Étude dialectologique Institut des langues nationales. Nouakchott: Institut des Langues Nationales.
KACHRU, B. 1985. Standards, codification, and sociolinguistic realism: the English language in the outer circle. English in the world: Teaching and learning the language and literatures, ed. by R. Quirk & H. Widdowson, 11-30. Cambridge: Cambridge University Press.
LABOURET. 1952. La langue des Peuls ou Foulbé. Mémoires de l'Institut Français d'Afrique Noire 16. Dakar: Institute Français d'Afrique Noire.

MEZEI, R. 1989. Somali language and literacy. Language Problems and Language Planning 13:3.211-223.
NEUSTUPNY, J. V. 1970. Basic types of treatment of language problems. Linguistic Communications (Monash University) 1.77-100.
OAU Bureau of Languages. 1973-80. Reconsideration of African language policies. Kampala: Organization of African Unity.
QUIRK, R. 1985. The English language in a global context. English in the world: Teaching and learning the language and literatures, ed. by R. Quirk & H. Widdowson, 1-6. Cambridge: Cambridge University Press.
-----. 1990. Language varieties and standard language. English Today 6.3-10.
RAY, P. S. 1963. Language standardization. The Hague: Mouton
RUBIN, J. 1973. Language planning: current issues and research. Washington, DC: Georgetown University Press.
-----, & Jernudd Björn (eds.) 1971. Can language be planned? Honolulu, HI: East-West Center Press.
SOW, A. I 1977. Langues et politiques des langues en Afrique noire. Paris: Nubia.
SPENCER, J. (ed.). 1963. Language in Africa. London: Cambridge University Press.
-----. 1971. Colonial language policies and their legacies. Current Trends in Linguistics 7.537-574.
TOLLEFSON, J. W. 1991. Planning language, planning inequality: Language policy in the community. London: Longman.
TURCOTTE, D. 1981. La politique linguistique en Afrique francophone. Quebec: Les Presses de l'Université de LAVAL.
WHITELY, W. H. 1966. Social anthropology, meaning, and linguistics. Man (n.s.) 1.139-157.
-----. 1971. Language policies of independent African states. Linguistics in Sub-Saharan Africa, ed. by T Sebeok. Current Trends in Linguistics 7.548-558.
-----. 1972. To plan is to choose. Bloomington, IN: Indiana University Press.
YANGA, T. 1978. Language planning and onomastics in Zaire. Studies in African Linguistics 9:2 233-244.
ZOUBKO, G. V. 1980. Dictionnaire Peul-Russe-Français. Moscou: Editions Langues Russes.

Studies in the Linguistics Sciences
Volume 23, Number 1, Spring 1993

BIMORAICITY IN MONOSYLLABIC CHICHEWA IDEOPHONES*

Andrew Tilimbe Kulemeka
University of Louisville

Kanerva (1990) has successfully shown that the Chichewa phonological word is disyllabic in underived major-category words such as nouns, verbs, and adjectives. Moreover, in such underived forms a mora is equal to a syllable. It follows that a phonological word is minimally bimoraic.

In Chichewa ideophones a phonological word is also disyllabic. However, underived monosyllabic ideophones are also bimoraic. Thus, in underived monosyllabic ideophones the mora and syllable counts are not identical. This paper presents crucial evidence from cliticization, penultimate lengthening, and facts of pronunciation which reveal the asymmetry between the mora and syllable counts in monosyllabic ideophones. The paper also demonstrates how monosyllabic ideophones become disyllabic. Further, I raise an important question about the interaction between regular and ideophone phonology.

1. Introduction

Ideophones or expressive words have been identified in many languages of the world. For instance, expressive forms exist in Austronesian languages (Brandstetter 1916), including Javanese (Uhlenbeck 1971) and Malay (Collins 1974); in Korean (Lee 1992, Lee 1993); in Dravidian languages, such as Kota, Toda, and Kannada (Emeneau 1969); in Indo-Aryan (Dimock 1957); in Thai (Haas 1946); in Japanese (Alfonso 1966); in Chadic (Newman 1968); and in Niger-Congo (Samarin 1965), particularly in Bantu (Torrend 1891, Samarin 1971, Kulemeka 1993).

The bulk of the literature on ideophones stresses the point that ideophones have a different phonological system from nouns, verbs, adjectives, and so forth (Childs 1988, 1989; Doke 1954; Fivaz 1963; Mphande 1989). While this may be correct to a certain extent, it needs to be substantiated. With respect to Southern Bantu, so far I have not come across any study which systematically investigates the nature and degree of phonological differences and similarities between ideophones on the one hand, and other lexical elements on the other. Most authors on Southern Bantu make statements about the aberrant phonology of ideophones without supporting those claims with empirical investigation. This study fills a gap in our understanding of the ideophone by carrying out an empirical investigation of the structure of monosyllabic ideophones.

Most writers on ideophones in Southern Bantu languages claim that ideophones tend to have long vowels (Fortune 1962; Zondo 1982; Mphande 1989), even though long vowels are often not phonemic in those languages. This is true of Chichewa as well. But careful observation of the distribution of long vowels in Chichewa ideophones clearly reveals the following: (i) All monosyllabic ideophones have underlying long vowels (1a); (ii) disyllabic ideophones may have underlying long or short vowels (1b); (iii) longer, 'polysyllabic' ideophones do NOT have underlying long vowels (2). (An accent indicates high tone; low tone is not marked.)

(1) a. Monosyllabic

/bi:/	'intense dark'
/dú:/	'not saying anything'
/gwá:/	'hard and strong'

b. Disyllabic

/čé:té/	'absolute silence'
/^{n}do:to:/	'something worthwhile'
/kwí:čí:/	'stop suddenly'
/p^{h}očo/	'fall heavily'
/pwimp^{h}wi/	'of a large object sitting'

(2) Polysyllabic

/k^{h}olop^{h}et^{h}e/	'numerous, abundant'
/gobede/	'hard items knocking against each other'
/bálámá^{n}t^{h}ú/	'suddenly appear'

The asymmetrical distribution of long vowels in (1) and (2) raises the question of the possible connection between vowel length and the minimal phonological unit in the ideophone. It seems that monosyllabic ideophones minimally have to be bimoraic (see e.g. Hayes 1981 and Hyman 1985 for issues relating to syllable weight).

The fact that the smallest possible ideophonic form has the structure [(C)μμ] and forms a single syllable is interesting when it is related to the fact that in the rest of the lexicon one mora equals one syllable in underived words (Kanerva 1990). The question that needs to be addressed is whether the minimal bimoraic ideophone is mono- or disyllabic. In order to answer such a question we need to briefly examine Kanerva's 1990 analysis and evidence.

Kanerva demonstrates that underlyingly all Chichewa syllables are monomoraic. Bimoraic syllables are derived through penultimate-syllable vowel lengthening.[1] One piece of evidence for Kanerva's analysis involves the assignment of high tone on the penultimate syllable. I summarize his arguments on high tone assignment immediately because they are relevant to my analysis.

Kanerva 1990 proposes that Chichewa has a tonal rule which assigns a high tone on the penultimate syllable of the verb. One of the triggers for the rule is the negative imperative marker *-sa-*. Kanerva indicates that *-sa-* 'removes any other tone the verb might

have and assigns a penultimate High tone' (1989:33). So, for example, the high tones of the subject and the object marker are assigned to the penultimate syllable as shown in (3). (Here, as elsewhere, 'OM' in the gloss refers to an object marker.)

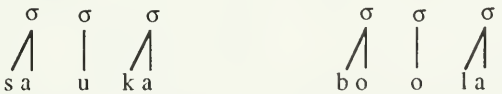
(3)	pʰika	'cook'
	tí-pʰika	'we-cook'
	tí-čí-pʰika	'we-OM-cook'
	ti-sa-či-pʰíke	'we-NegImp-OM-cook'

The pattern of penultimate high tone assignment becomes critical when a word has an underlying sequence of vowels, and it needs to be determined if those vowels form one heavy bimoraic syllable, CVV, or two light syllables, CV.V. The elements /boola/ 'pierce' and /sauka/ 'suffer' are used to illustrate the point. Neither word has an underlying high tone (cf. Kanerva 1989, example 9).

(4)	a.	boola	'pierce'
		tí-boola	'we-pierce'
		tí-čí-boóla	'we-OM-pierce'
		ti-sa-či-boóle	'we-Neg-OM-pierce'
	b.	sauka	'suffer'
		tí-sauka	'we-suffer
		tí-mú-sauka	'we-OM-suffer'
		ti-sa-mu-saúke	'we-Neg-OM-suffer'

As shown by example (4), the rule which assigns high tone to the penultimate syllable reads the sequence of two vowels in /boola/ and /sauka/ as falling into DIFFERENT syllables, schematically represented as follows.

(5) a. H b. H



The penultimate syllable high tone rule is highly productive in Chichewa tonology. Its appearance on the penultimate syllables in the examples in (4) provides one strong piece of evidence for treating the penultimate vowels as separate syllables.

Further, Kanerva uses the penultimate tone assignment rule to also argue that the mora count is exactly identical to the syllable count in underived words in Chichewa. Moreover, he shows that the minimal phonological word for major category words in Chichewa is disyllabic. The last statement entails that the minimal underived phonological word has two moras.

The question, then, arises whether the minimal ideophone also is disyllabic and bimoraic. If so, does this mean that the minimal ideophonic unit is equivalent to a single phonological word?

If the smallest ideophone is equal to a disyllabic word, then it follows that the term 'monosyllabic ideophones' is inappropriate to refer to these forms. In fact it would mean that, with respect to the smallest word possible, ideophones do not violate the rules of regular phonology.

In the following I show that the minimal ideophonic unit is not disyllabic and thus violates the rules of regular phonology. In fact, because monosyllabic ideophones (though bimoraic) do not count as disyllabic, they amount to less than a phonological word in the language, even though they are full morphological words.

Three pieces of evidence support this position: the failure of clitics to attach to monosyllabic ideophones; the position of penultimate vowel lenghthening in ideophones which have become disyllabic; and differentiations of pronunciation between monosyllabic and disyllabic ideophones.

2. Clitic attachment

Nespor & Vogel (1986), Zec & Inkelas (1990), and Kanerva (1990) all agree that clitics attach to full words and no less. Moreover, through a very detailed analysis of different types of bound morphemes, Inkelas (1989) shows that clitics are sensitive to prosodic features, rather than morphological ones. Hence clitics attach to FULL PHONOLOGICAL WORDS. In other words, we can test whether a form counts as a phonological word in a language by attaching a clitic to it.

/nsó/ 'also' and the emphasizer /dí/ are high-toned clitics in Chichewa. Like clitics in other languages 'they may never occur alone; that is they may not be the only element of an utterance' (Nespor & Vogel 1986:149; see also Inkelas & Zec 1990). Chichewa clitics are further distinguished from other forms of affixes, such as verbal extensions, because they alone can attach to all major category words which are minimally disyllabic, such as verbs, nouns, adverbs, possessive pronouns, and adjectives. In addition, clitics such as /-nsó/ or /-dí/ modify the basic meaning of their hosts and sometimes the entire sentence along somewhat predictable lines.

Kanerva (1990) shows that Chichewa has a rule which assigns a high tone to the syllable preceding the clitic. Thus, for example, low tone elements as in (6) acquire high tone through the clitic high tone assignment rule as follows.

In all the cases in (6) we see that a low-toned form acquires a high tone on the syllable preceding the clitic through the clitic high tone assignment rule which assigns a high tone to the left syllable of an element to which the clitic is attached.

Underived monosyllabic ideophones fail to take clitics, as shown in (7a). In contrast, di- and polysyllabic ideophones do accept clitics, cf. (7b).

(6) a. mu-nthu 'person'

mu-nthu-nso 'also a person'

mu-nthu-nso

b. galu 'dog'

galu-nso 'also a dog'

galu-nso

(7) a. Monosyllabic

bu:	'light a fire'	*__bu:__-nso
čí:	'throw'	*__čí:__-nso
mbwe:	'arrive'	*__mbwe:__-nso
myá:	'smooth'	*__myá:__-di
bi:	'intense dark'	*__bi:__-di
bí:	'convincing'	*__bí:__-di

b. Di- & polysyllabic

ndo:to:	'something worthwhile'	**ndo:tó:**-nso	'also something worthwhile'
čololo	'pouring liquid'	**čololó**-nso	'again pour liquid'
psí:tí:	'completely finished'	**pí:tí:**-di	'indeed completely finished'
delú	'pass quickly'	**delú**-di	'indeed pass quickly'
gankha	'to be and look thick headed and ignorant'	**gankhá**-nso	'to also be and look thick headed and ignorant'

As shown in (6) and (7), the high tone which comes with the clitic gets assigned to the syllable preceding the clitic in a di- or polysyllabic form. Such a syllable will naturally be preceded by one or more syllables. In other words, the form to which the clitic attaches is minimally disyllabic, a phonological word. Clitics fail to attach to monosyllabic ideophones. It is only when the structure of monosyllabic ideophones is changed to become disyllabic that clitics attach and assign a high tone on the clitic's preceding vowel (see further below).

The low-toned ideophones in (7a) further demonstrate that the high tone of the clitic fails to appear. For the high tone of the clitic to show up, again a second syllable is required.[2]

If the mora count were identical to the syllable count in underived monosyllabic ideophones, we would expect that clitics would attach to the forms in (7a) because they are bimoraic and thus would

be disyllabic. The high tone of the clitic should surface on the second mora, which would be the second syllable. Since the high tone of the clitic fails to appear on the second mora in low-toned monosyllabic-bimoraic ideophones, we have to conclude that the mora count in monosyllabic ideophones is not identical to the syllable count. Therefore the strict identity between mora count and syllable count holding for other category words, argued for by Kanerva (1990), does not apply to monosyllabic ideophones.

However, it is still interesting that the minimal ideophone is bimoraic. In a way, therefore, the minimal ideophone partially resembles the smallest phonological word, which also has two moras, but in separate syllables.

A further interesting aspect of monosyllabic ideophones is that they do not have an upper limit in terms of the length of their vowels. Thus, depending on expressive needs, the vowel length of the minimal ideophonic unit is completely controlled by the speakers. For example, speakers might say /bi:/ or [bi:::] or [bi:::] or [bi::::::] or [bi::::::::::].[3] In contrast, one cannot expand vowels in verbs, nouns, adverbs, or adjectives indefinitely. Thus, excessive lengthening in words like /gula/ 'buy', /mwana/ 'child', and /táli/ 'tall', such as *[gu::la], *[mwana::], or *[tá:::li:::], is ungrammatical.

Although the length of vowels in the minimal ideophone is unlimited, length by itself does not appear to compensate for phonological wordhood in Chichewa. That is, even when lengthened, monosyllabic ideophones do not accept clitics.

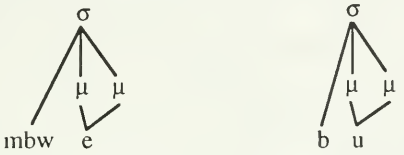
Thus we are forced to conclude that even though the smallest underived ideophone is bimoraic, i.e. contains a long vowel, the two moras comprise one syllable. This is in marked contrast to the behavior of 'ordinary' bimoraic words. Now, the fact that monosyllabic bimoraic ideophones have long vowels is uncontroversial and has been acknowledged by almost all researchers on the subject. What is significant, however, is the asymmetry between 'ordinary' words and ideophones as regards the relation of mora count to syllable structure.

Note further that in order for clitics to attach to it, the monosyllabic ideophone has to become disyllabic. To do this, the second mora to the right delinks to form a separate syllable. The evidence for such a process comes from the interaction of monosyllabic ideophones with high toned clitics like /-nsó/ 'also' and /-dí/ 'indeed', facts of pronunciation, and the occurrence of the phrase-final rule of penultimate lengthening.

Let us first examine the effects of high-toned clitics such as /nsó/ and /dí/. The structural description for cliticization requires a preceding form that is minimally disyllabic [σσ]. Monosyllabic ideophones, however, do not meet this requirement. As shown in (7a), ordinarily these clittics therefore cannot attach to monosyllabic ideophones and thus fail to pass on their high tone.

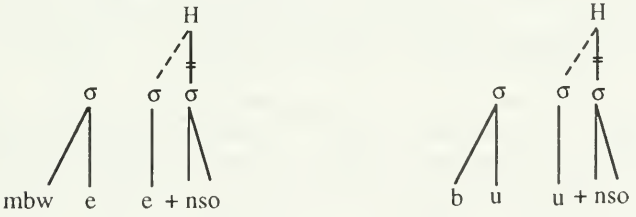
However, if the second mora of a monosyllabic ideophone delinks and forms a separate, second syllable, then the clitics can attach and pass their high tone to the newly created preceding syllable. Note specifically that it is this second syllable alone which bears the high tone of the clitic. Compare the derivations in (8). (The derivations are illustrated with the low-tone monosyllabic ideophones /mbwe:/ 'arrive' and /bu:/ 'light a fire', plus the high-tone clitic /nsó/ 'also'.)

(8) a. Underlying representation



b. Delinking and resyllabification

c. Clitic high tone assignment



d. Surface forms

[**mbwe.é.**nso] [**bu.ú.**nso]

Further support for the claim that the pre-clitic vowel in (8d) is a separate syllable comes from pronunciation. There is an obvious break between it and the preceding vowel, as in [mbwe.é.nso] and [bu.ú.nso]. This inter-vowel break is absent in the corresponding unmodified monosyllabic ideophones /mbwe:/ and /bu:/ (not *[mbwe.e], *[bu.u]). Note that a similar break is evident in 'ordinary' words. For instance, the forms /boola/ and /sauka/ of (4) above are pronounced [bo.o.la] and [sa.u.ka], with an obvious break and transition between the neighboring vowels. This shows that, even when

the two vowels are identical, they are not a single long vowel but form separate syllables.

A final piece of evidence in support of the analysis involves the common phrasal rule of penultimate vowel lengthening. Significantly, in structures like those of (8d), such as [mbwe.é.nso] and [bu.ú.nso], the process only lengthens the penultimate vowel, as in [mbwe.**é:**.nso] and [bu.**ú:**.nso]. If the long vowels of these forms had not undergone resyllabification, we would expect penultimate lengthening to have affected the entire long vowel, yielding *[mbwe:(:).nso] and *[bu:(:).nso]. The fact that we get [mbwe.**é:**.nso] and [bu.**ú:**.nso] therefore provides strong evidence for treating the sequences of two vowels as comprising separate syllables.

Four pieces of evidence, then, provide strong empirical support for arguing that in order to accept clitics, monosyllabic ideophones undergo a process of delinking and resyllabification: (i) rejection of clitics by unmodified monosyllabic ideophones; (ii) placement of the clitic high tone on the SECOND syllable of modified ideophones, (iii) a break and transition between the first and second syllables of such ideophones; and (iv) the fact that penultimate vowel lengthening affects only the second syllable of such structures.

3. Conclusion

This paper has shown that Kanerva's claim (1990) that there is a one-to-one correspondence between mora and syllable count in underived words of Chichewa cannot be fully maintained. While the claim holds true for underived 'ordinary' words, ideophones show a remarkably different phonological pattern: Although bimoraic, they are monosyllabic, not disyllabic as expected. As such they constitute sub-minimal phonological words and therefore cannot host clitics, unless modified by delinking and resyllabification.

These differences in prosodic phonology between ideophones and 'ordinary' words raise important questions which may require a reevaluation of the relationship between ideophone and regular phonology. For instance, to what extent are other conclusions about regular phonology applicable to ideophone rules? Is ideophone phonology distinct or a part of regular phonology? The extent of such a reevaluation can only be gauged after further and more extensive investigation of ideophone behavior.

NOTES

* A version of this paper was presented at the 24 Annual Conference on African Linguistics at Ohio State University, Columbus. I am grateful to Jonni Kanerva, Charles Kisseberth, Lupenga Mpande, Robert Botne, Francis Moto, and Hans Henrich Hock for their very useful comments which have helped me to clarify my points and provide additional supporting evidence. I also thank an anonymous

reviewer for SLS for making pertinent observations. Of course all errors in the paper are entirely my fault.

[1] Kanerva (1990) does not include ideophones in his discussion.

[2] Several native speakers of Chichewa (Dr. Francis Moto, Dr. Lupenga Mphande, and Mr. Simango), who attended the 24th Annual Conference of African Linguistics at Ohio State University, have provided further support for the view that monosyllabic ideophones cannot take clitics. When asked to produce monosyllabic ideophones with clitics, they invariably made the ideophone disyllabic as shown later in this section.

[3] The symbol /::/ represents a vowel whose length is double the usual long length, /:::/ represents thrice the usual long length, etc.

REFERENCES

ALFONSO, A. 1966. Japanese language patterns, volume I. Tokyo: Sophia University Press.

BRANDSTETTER, Renward. 1916. An introduction to Indonesian Linguistics. London: The Royal Asiatic Society.

CHILDS, G. Tucker. 1988. The phonology of Kisi ideophones. Journal of African languages and linguistics 10.165-190.

-----. 1989. African ideophones. Sound symbolism, ed. by Leanne Hinton, John Nichols, & John Ohala. Forthcoming. Cambridge: Cambridge University Press.

COLLINS, James T. 1974. Expressives in Kedah Malay. Pacific linguistics: South East Asian linguistic studies 4.379-406.

DIMOCK, E. C. 1957. Symbolic forms in Bengali. Bulletin of the Deccan College Research Institute 18.22-29.

DOKE, Clement M. 1954. The southern Bantu languages. London: Oxford University Press.

EMENEAU, M. B. 1969. Onomatopoeics in the Indian linguistic area. Language 45.274-299.

FIVAZ, Derek. 1963. Some aspects of the ideophone in Zulu. (Hartford Studies in Linguistics, 4) Hartford: Hartford Seminary Foundation.

FORTUNE, George. 1962. Ideophones in Shona. London: Oxford University Press.

HAAS, Mary. 1946. Techniques of intensifying in Thai. Word 2.127-130.

HAYES, Bruce. 1981. A metrical theory of stress rules. MIT Ph.D. dissertation in Linguistics. (Revised version distributed by the Indiana University Linguistics Club.)

HYMAN, Larry. 1985. A theory of phonological weight. Dordrecht: Foris Publications.

INKELAS, Sharon. 1989. Prosodic constituency in the lexicon. Stanford University Ph.D. dissertation in Linguistics.

KANERVA, Jonni. 1989. Focus and phrasing in Chichewa phonology. Stanford University Ph.D. dissertation in Linguistics.

-----. 1990. Focus and phrasing in Chichewa phonology. New York: Garland.

LEE, Jin-Seong. 1992. Phonology and sound symbolism of Korean ideophones. Indiana University Ph.D. dissertation in Linguistics.

LEE, Yongsung. 1993. Topics in the vowel phonology of Korean. Indiana University Ph.D. dissertation in Linguistics.

KULEMEKA, Andrew T. 1993. The status of the ideophone in Chichewa. Indiana University Ph.D. dissertation in Linguistics.

MCCARTHY, John, & Alan Prince. 1986. Prosodic morphology. University of Massachusetts at Amherst, and Brandeis University, MS.

MPHANDE, Lupenga. 1989. A phonological analysis of the ideophone in Chitumbuka. University of Texas at Austin Ph.D. dissertation in Linguistics.

NESPOR, Marina, & Irene Vogel. 1986. Prosodic phonology. Dordrecht: Foris Publications.

NEWMAN, Paul. 1968. Ideophones from a syntactic point of view. Journal of West African languages 2.107-17.

SAMARIN, William J. 1965. Perspective on African ideophones. African Studies 24.117-121.

-----. 1971. Survey of Bantu ideophones. African language studies 12.130-168.

TORREND, J. 1891. A comparative grammar of the South African Bantu languages. London: Kegan Paul, Trench, Trubner and Co.

UHLENBECK, E. M. 1971. Peripheral verb categories with emotive-expressive or onomatopoeic value in modern Javanese. Travaux linguistiques de Prague 4.145-156.

ZEC, Draga, & Sharon Inkelas. 1990. Prosodically constrained syntax. The phonology-syntax connection, ed. by Sharon Inkelas and Draga Zec, 365-378. Chicago: University of Chicago Press.

ZONDO, J. 1982. Some aspects of the ideophone in Ndebele. Zambezia 10.111-126.

Studies in the Linguistic Sciences
Volume 23, Number 1, Spring 1993

A RE-EXAMINATION OF CARDINAL VOWELS AND AUDITORY EQUIDISTANCE*

Sharon R. Morrison

The present study examines the hypothesis of auditory equidistance of Jones's primary cardinal vowels. Four sets of previously recorded cardinal vowels spoken by Daniel Jones were digitally analyzed. The obtained formant peaks were used to compute Euclidean distance measures between adjacent vowels in a F1/F2 vowel space. ANOVAs of the linear inter-vowel distances (in Hz) and distances in several non-linear scales (Bark, semitone, and auditory-perceptual) were interpreted as indicating that the perceptual inter-vowel distances of the cardinal vowels were significantly different. Separate analyses of the front and back vowel series did not produce consistent results.

1. Introduction

The cardinal vowel system was developed by Daniel Jones in the early 1900's, and the first published description of cardinal vowels appeared in the 1917 edition of An Outline of English Phonetics (see Jones 1956). Cardinal vowels were intended as a descriptive system for vowels based on articulatory positions of the tongue. Jones taught that all the cardinal vowels were articulatorily equidistant, that the difference between any two adjacent vowels involved an equal displacement of the tongue. For example, the difference in tongue position between the first and second cardinal vowels (CV1 and CV2) should be equal to the difference between CV2 and CV3 (see Figure 1). The top and bottom distances between front and back vowel series (CV1 and CV8; CV4 and CV5 respectively) were not specified as equidistant in Jones's system.

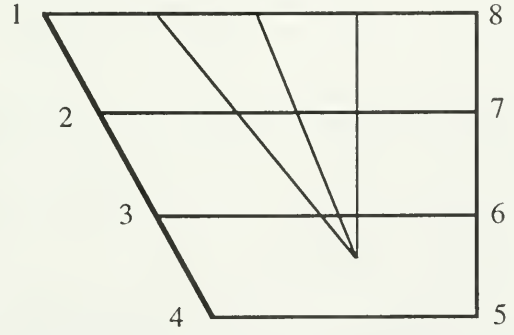


Figure 1. The cardinal vowel figure or trapezium

The cardinal vowels were used almost exclusively in the British school of phonetics. Training in the use of cardinal vowels was a long and laborious procedure. The vowels were to be taught in an oral tradition (by Jones or his students). To use the cardinal vowels as a descriptive technique one learned to produce each vowel, paying particular attention to the position of the tongue. To describe a natural vowel the user would learn to produce that vowel to the satisfaction of the native speaker (informant) and then proprioceptively compare the articulation of that vowel to the closest cardinal vowel. After determining the articulatory relationship between the natural vowel and the cardinal vowel, the natural vowel was placed in the appropriate location on the cardinal vowel figure (trapezium) along with a notation of lip rounding (Abercrombie 1967, 1985). At some point the cardinal vowel figure came to be viewed as an auditory space, not an articulatory space (Ladefoged 1960a, 1967, 1982; O'Connor 1973).

One reason for considering the cardinal vowels as reference points in an auditory space is that x-rays of the cardinal vowels did not show articulatory equidistance of tongue position (S. Jones 1929; no relation to Daniel Jones). Stephen Jones did not publish a scale or measurements of tongue position with his x-rays, and the reproductions were poor. Still, the differences in tongue positions did not appear equidistant, especially for the back vowels (CV 4 - CV8). Ladefoged (1960a, 1967, see also 1982) measured differences in tongue position in S. Jones's (1929) x-rays and found no indication of articulatory equidistance.

Ladefoged (1960b) investigated the usefulness of cardinal vowels as a classification tool. He compared the classification of Gaelic vowels by observers trained in the British school of phonetics (who were extensively trained in the proprioceptive use of cardinal vowels) and those not formally trained in the use of cardinal vowels. Ladefoged concluded that the 15 British-trained observers seemed more consistent, as a group, about the placement of the Gaelic vowels on the trapezium than were the three other subjects not trained extensively in the use of cardinal vowels. Ladefoged's conclusions were interesting, but they were not based on any inferential statistical analyses.

Laver (1965) studied the variability, over time, of the placement of synthetic vowels on the cardinal vowel diagram. All of his subjects were trained extensively in the use of cardinal vowels. Vowel placement was found to vary from trial to trial, though the average locations for each vowel showed a high agreement within the group. Laver also found that, for each subject, there was no significant change in overall variability over time.[1]

The usefulness of cardinal vowels could be furthered by resolution of the issue of equidistance. Obviously, cardinal vowels are not articulatorily equidistant (S. Jones 1929; Ladefoged 1960a; 1967, 1982), but are they equidistant in an auditory frame work?

Ladefoged (1982) indicated that, if cardinal vowels are plotted on a formant chart, there is no way to get the distances between the front vowels to be the same as the distances between the back vowels. His claim is probably based on the spectrographic analysis of cardinal vowels described by Ladefoged, (1960a; 1967). Ladefoged 1960a is a monograph based on his Ph.D. thesis, and a slightly revised version of this monograph appears as the second chapter of Ladefoged 1967. The main difference between the two versions is that the 1960a form has slightly more detail about methods used, larger figures and also includes a table of unscaled cardinal vowel formants.

Ladefoged 1960a, 1967 examined the nature of vowel quality on a number of levels. The relevant section (or chapter) is the third one, 'The acoustic analysis of cardinal vowels'. In this section Ladefoged used cardinal vowels to study the meaning of phonetic quality. Eleven experienced phoneticians, all trained extensively in the use of cardinal vowels by Daniel Jones himself, produced multiple sets of cardinal vowels in various pitches. Each subject was rehearsed in the eight primary cardinal vowels by Jones and then produced at least five sets of cardinal vowels. Each vowel was produced in isolation. Each vowel set was criticized and discussed by Jones and the informant. Jones was the final judge of vowel quality and selected three (sometimes only two) sets of vowels for each informant (a total of 31 sets out of 92 sets) as 'good complete sets not in the extreme pitch ranges' (Ladefoged 1967:78).

The goal of Ladefoged's acoustic analysis was to determine the frequency of each of the first three formants for all 248 vowels. In the course of analysis at least two and often more spectrograms were made of each vowel. All acoustic analyses were done on a modified Kay Sonograph. In addition to the normal frequency scale (1"=2000 Hz) an expanded (1"=1000 Hz) and ultra-expanded (1"=500) scale were added. Acoustic analysis was based on wide band spectrograms and sections (a narrow band spectrum cross section at a given instant in time). Ladefoged (1960a, 1967) reported difficulty finding F1 for the high front and back vowels [i] and [u] (CV1 and CV8) as well as differentiating F1 and F2 when they were close together as found in tokens of [a] and [ɔ] (CV5 and CV6). The second and third formants were also hard to find for some vowels. Ladefoged also reported difficulties with spurious formants in unexpected locations and discussed the poorly defined nature of formants.

Ladefoged presented the results of his analyses of 31 sets of cardinal vowels spoken by 11 trained phoneticians. Many of the vowels could not be adequately described in terms of formant frequencies, especially the back vowel series. For 16 of 31 cases Ladefoged was unable to distinguish the first two formants for [a] (CV5), in 12 out of 31 times he could not determine either F1 or F2 for [ɔ] and [o] (CV6 and CV7), and in 26 out of 31 he could not completely specify formants for [u] (CV8). Formants were presented as hertz and mel (Stevens & Volkmann 1940) values in Ladefoged

1960a and only mel values in Ladefoged 1967. Ladefoged 1960a; 1967 first plotted all sets of vowels in a traditional F1/F2 inverted plot and found that only CV 4 and CV8 ([a] and [u]) did not exhibit overlapping formant regions. CV1 and CV2 and CV6 and CV7 exhibited the most overlap. Ladefoged then tried to use various combinations of the first three formants to 'normalize' formants for the back vowels and thereby differentiate them. None of the attempts appeared successful. Ladefoged discussed the notion that the phonetic quality of vowels apparently does not depend on absolute acoustic value of the formants, but perhaps on the relationship between the formants of all the vowels produced by a given speaker. Following this viewpoint, he plotted each sets of cardinal vowels spoken by each phonetician separately. In the individual plots the relationship between the cardinal vowels was clearer than in the aggregate plot. In the individual plots more vowels were differentiable.

There are several limitations to Ladefoged 1960a, 1967. First, Ladefoged was unable to determine the formants of many back vowels, so his set of formants was not complete. Second, Ladefoged did not attempt any quantitative measures of the distances between cardinal vowels. Another problem with Ladefoged's study may have been how he scaled the cardinal vowel formants. The perception of pitch does not have a linear relationship to frequency measured in hertz (Stevens & Volkmann 1940) and researchers commonly transform formant values in an attempt to plot them in some perceptual space. One of the more commonly used scales is the mel scale of subjective pitch where a 1000-Hz tone, 40 dB above threshold, has a pitch of 1000 mels (Stevens & Volkmann 1940). The mel scale was derived from subjective pitch evaluations of pure tones. Because speech is a complex signal and does not have the same (or even similar) pitch characteristics as pure tones, the transformation of formant frequencies from hertz to mels may be a questionable practice.

Other frequency scales, derived from critical bands and frequency difference limens, do represent the non-linear nature of the human auditory system (Munson & Gardner 1950). The Bark scale (Zwicker 1961), Miller's (1989) Auditory-Perceptual space, and the semitone (semit) scale are all scales which represent frequency in a non-linear fashion. These scales were used in the present study to perceptually scale cardinal vowel formants.

The aim of the present study is to rigorously and quantitatively resolve the issue of auditory equidistance using linear predictive analysis of cardinal vowel formants. Digital analysis is not without its problems, such as varying results depending on the analysis parameters, but the results presented include at least one set of formant values for each cardinal vowel. Calculated measures of the distance between vowels were statistically analyzed for both linear and transformed formant values.

2. Recordings of the cardinal vowels

Each of four, independently acquired, tape-recorded sets of the primary cardinal vowels was digitized and analyzed using linear predictive coding. All four sets of vowels had been spoken by Daniel Jones. Three of the sets were recorded by Peter Ladefoged in 1956 as part of his dissertation (PL sets 3, 6, and 7) (see Ladefoged 1960a, 1967 for recording methodology). The fourth set is from a cassette copy of a recording of the cardinal vowels made by Daniel Jones.[2]

3. Digital analysis

The cassettes were played on a stereo cassette deck (Sony TC-WR950) and bandpass filtered between 100 and 7000 Hz using a 'Brickwall' filter (Wavetek Rockland model 751A). The amplitudes of the PL3 and PL6 sets were amplified 10 dB to equalize the amplitudes for all the vowel sets. The cardinal vowels were then digitized (Data Translation DT2823 A/D-D/A board) at 20,000 samples per second with a resolution of 16 bits.

Interactive Laboratory System (PC-ILS) software (Signal Technology Inc. 1987a) was used for all speech analysis. The temporal centerpoint of each vowel was found, and the frame on each side was analyzed using the ANA command. Points 100 msec on both sides of the midpoint were also analyzed as a consistency check of analysis results. Variations in formants for all vowels were generally less than five percent of the centerpoint formant frequency. The RSO (root solving) command was then used to find the formant peak values (Signal Technology Inc. 1987b).

The results of the digital analysis varied depending on the analysis parameters used in PC-ILS. Most vowels were best analyzed with an 18th order filter, preemphasis of 98% (to compensate for the frequency roll off found in speech signals), a Hamming window, analysis frame size of 200-250 points (samples), and the number of peaks set to eight (8). The above parameters produced reasonable formant values for most vowels.[3] The high back vowel CV8 was analyzed with a higher order filter in an attempt to differentiate between the first two formants.

Because the digital analysis sometimes produced unexpected results, the following post hoc rules were used to make the analysis results more consistent. These rules have no real theoretical basis, but are similar in spirit to the methods used in the art of spectrograph reading.

1. If the calculated bandwidth was equal to or greater than the value of the resonance peak, or the amplitude was very low, then that peak was discarded.
2. Formants with very wide bandwidths were kept only if the other analysis frame for that vowel token had a peak close in value with a narrower bandwidth.

3. A few formant peaks were listed when the value occurred in only one analysis frame if peaks with similar values were found in other samples from the same vowel.
4. When an unexpected formant value was found in an analysis frame, it was dropped if a similar value did not also appear in at least half of the other frames for that token.

4. Formant values

The obtained formant values for the primary cardinal vowels are listed in Table 1. PL indicates that the vowels are from the tape of cardinal vowels recorded by Ladefoged (using his numbering system). LL avg. are the average formants extracted from the duplicate vowel tokens on the tape of Daniel Jones supplied by Lisker.

Formant freq.(Hz): Set		CV1	CV2	CV3	CV4	CV5	CV6	CV7	CV8
F1	LL avg.	233	371	567	889	840	501	342	216
	PL set 3	295	379	476	966	616	476	340	177
	PL set 6	296	427	587	726	666	464	402	?
	PL set 7	304	346	615	983	634	461	327	170
F2	LL avg.	2541	2142	1837	1575	954	893	668	410
	PL set 3	2108	2166	1866	1696	813	785	629	362
	PL set 6	2492	2324	1980	1640	936	709	986	?
	PL set 7	2086	2225	2025	2002	888	704	592	355
F3	LL avg.	3566	2454	2220	1880	2396	2389	2256	2396
	PL set 3	3681	3779	3989	3827	3450	3660	2924	2399
	PL set 6	3808	4109	4126	2107	2310	2703	2982	?
	PL set 7	3357	3513	3383	3390	3304	3322	3371	?

Table 1. Formant frequencies for the primary Cardinal Vowels (in Hertz)

The first formants exhibit the expected inverse relationship with vowel height (Ladefoged 1982). The high vowels, CV1 and CV8, have low F1's and the low vowels, CV4 and CV5, have high first formants. The inverse relationship between vowel height and first formant frequency is displayed in Figure 2. Figure 2 is a traditional, inverse plot of the first two formants of the type first presented by Joos (1948). This linear frequency plot has the advantage of allowing a more direct visual comparison of vowel position in the theoretical vowel trapezium and of obtained vowel formants (Hockett 1955). For example, in the traditional vowel space plot, the high, front vowel CV1 appears in the upper left corner as it does in the trapezium (Figure 1). The other vowels also appear in their correct positions relative to each others' position in the trapezium.

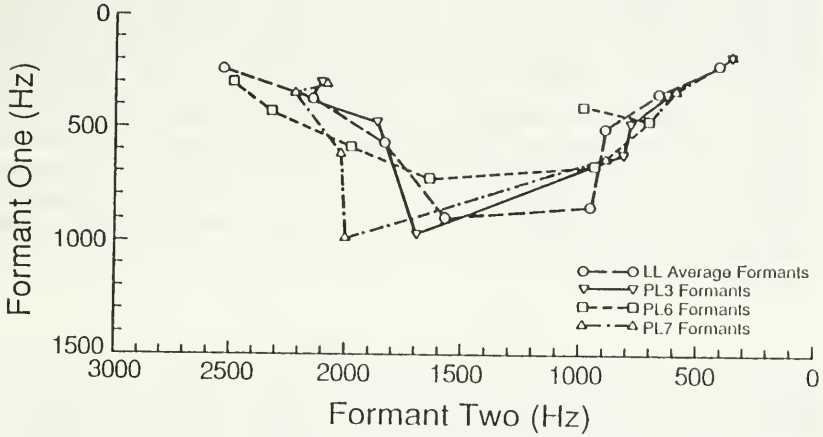


Figure 2. Values of the first two formants of four sets of Jones's cardinal vowels plotted in the traditional inverse F2/F1 style

The high front vowels, CV1 and CV2, show the most F2 variation. Possibly this is related to the greater F1/F2 separation found in the high front vowels. Also, F2 for CV1 is quite close to the F2 for CV2 in the PL sets. The obtained second formants for CV1 are lower than expected for PL sets 3 and 7 (Ladefoged 1960a, 1967). The obtained F2 for CV8 is not quite in its expected location (around 815 Hz, Lisker 1989). These discrepancies may be due to degradation of the recording between the time of recording and the time of the experiment. The second formants show the expected lowering of frequency as Jones moves around the vowel quadrangle (see Figure 2).

Third formant values do not appear very consistent. Extraction of higher vocal tract resonances is more difficult than for lower resonances due to the intensity roll-off found in the speech spectrum. The preemphasis of the signals prior to analysis attempted to counter some of this high frequency loss, but may not have been effective in this case. Some other causes of variation of the higher formants might be the age or quality of the recordings or the fact that the signals were low-pass filtered by the microphones used for recording the original vowel sets.

5. Inter-vowel distances

In the F1/F2 plot of the vowel sets (Figure 2) the spatial distances between the cardinal vowels do not appear equidistant. But, because the issue of equidistance is difficult to resolve solely on the basis of visual inspection, a quantitative method was used to test for equidistance. Euclidean distance measures were computed for

adjacent vowel pairs based on the values of the first two formants for all four vowels sets (see Table 2). (Note that Jones did not specify the intervals between CV4-CV5 and CV1-CV8 as equidistant with the remaining pairs). The final inter-vowel interval, between CV7 and CV8 for PL set 6, could not be computed because no formants were extracted for CV8 of PL set 6. In order to maintain the maximum number of degrees of freedom, the final interval distance for PL set 6 was set at the mean interval size of the other three vowel sets. A repeated measures analysis of variance of the distance measurements showed a significant difference in the distances between the different pairs of cardinal vowels ($F(5,15) = 4.17$, $p<0.02$). The conclusion is that cardinal vowels may not be equidistant in a linear frequency (Hz) space.

		Inter-vowel interval					
Scale	Set	CV1-2	CV2-3	CV3-4	CV5-6	CV6-7	CV7-8
Hertz	LL avg.	422.19	362.55	415.12	344.45	275.50	287.12
	PL set 3	102.08	315.29	518.65	142.77	206.96	312.82
	PL set 6	213.04	379.39	367.32	303.86	283.85	(294.74)[a]
	PL set 7	145.21	335.20	368.72	276.96	174.64	284.12
Bark	LL avg.	1.68	1.93	2.60	2.72	2.20	2.50
	PL set	30.80	1.30	3.82	1.20	1.73	2.30
	PL set	61.28	1.72	1.67	2.36	2.04	(5.20)[a]
	PL set	70.57	2.43	2.69	1.99	1.52	2.57
Semits	LL avg.	8.58	7.81	8.23	9.02	8.30	11.61
	PL set 3	4.36	4.71	12.36	4.51	6.98	14.81
	PL set 6	6.46	6.17	4.92	7.90	6.22	(13.60)[a]
	PL set 7	2.50	10.09	8.12	6.83	6.66	14.38
Auditory	LL avg.	0.81	0.72	0.75	0.69	0.46	0.72
	PL set 3	0.56	0.49	1.07	0.35	0.36	0.75
	PL set 6	0.58	0.60	0.66	0.57	0.55	(1.08)[b]
	PL set 7	0.15	0.88	0.67	0.41	0.43	(0.69)[c]

[a] Average interval distance of the three other sets.
[b] Computed from formant averages of the other sets, F1=188 Hz, F2=376 Hz, F3=2398 Hz.
[c] Computed using formant average of the other sets for F3=2398 Hz.

Table 2. Inter-vowel distances for the primary cardinal vowels using linear and transformed scales

Ladefoged (1960a, 1967) reported that many phoneticians considered the intervals between the first four cardinal vowels to be greater than the intervals between the last four vowels and presented formant charts of 31 sets of cardinal vowels to support this idea. To examine this idea quantitatively, the inter-vowels distances for the front and back vowels were analyzed separately. The ANOVA of the first three intervals (front vowel series) did not produce

significant results, ($F(2,6) = 4.31$, $p<0.07$). An analysis of the last three intervals, (back vowel series) also failed to produce statistically significant results, ($F(2,6) = 1.18$, $p<0.38$). These results support the idea that the front and back cardinal vowels may be independently equidistant, that is, with different distances within each series.

Because frequency and perceived pitch are not linearly related, various attempts were made to perceptually scale the cardinal vowel formants to see if a scaling could be found where all the cardinal vowels were equidistant. The first scale used was the critical band or Bark scale (Zwicker 1961). Zwicker & Terhardt (1980) provided a mathematical approximation to the empirically derived Bark scale:

(1) $\text{Bark} = 13 \arctan(0.76\, f) + 3.5 \arctan(f/7.5)$,
where f is the frequency in kilohertz.

Formant frequencies for all four vowel sets were converted to Barks using Equation (1). As was done for the linear inter-vowel distance computations, the final interval distance for PL set 6 was set at the mean interval size of the other three vowel sets. An analysis of variance revealed a significant difference in the size of the interval between adjacent cardinal vowels ($F(5, 15) = 3.16$, $p<0.04$). The front vowel intervals alone were marginally different from each other ($F(2,6) = 4.77$, $p<0.06$), although the back vowel intervals did not significantly differ from each other ($F(2,6) = 3.12$, $p<0.12$).

Another attempt at perceptually scaling cardinal vowels to find equidistance used the semitone (semits) scaling. Semits are units in the Equal Temperment musical scale and were commonly used by Speech Scientists for traditional inverse formant plots (Hockett 1955). Semits are defined in Equation (2):

(2) $\text{semits} = 12\log_2 (f / 16.35)$

where f is frequency measured in hertz and 16.35 is the reference frequency corresponding to C0 in the Usually Equally Tempered Scale (where an octave is divided into 12 equal intervals) based on A4 = 440 Hz (American National Standards Institute 1960). Formants for all four vowel sets were converted into semits using Equation 2, and Euclidean distances were computed for each adjacent vowel pair. The final interval distance for PL set 6 was set at the mean distance of the other three vowel sets. The results of a repeated measures analysis of variance revealed that the inter-vowel distances using a semitone scaling were significantly different from each other, ($F(5,15) = 6.18$, $p<0.005$). In contrast to the Bark scaling, the front vowel series distances were not significantly different, ($F(2,6) = 0.93$, $p<0.50$), while the back vowel series was different, ($F(2,6) = 18.66$, $p<0.005$).

Miller (1987, 1989) proposed that speech sounds are mapped onto phonetically relevant target zones in an auditory-perceptual space (APS) using log frequency ratios of F0, F1, F2 and F3. Scaling in APS is an attempt to deal with vowel normalization of inter- and

intra-speaker variations in an auditory-perceptual space. Speech is first processed at a sensory level. Miller's (1987, 1989) scalings (see Equations (3)-(6) below) were based on the speech sounds of American English, but the APS framework has been successfully extended to German and Greek vowels (Jongman, Fourakis, & Sereno 1989). Since the APS framework appears to be generally applicable to vowel perception, it was used to examine the equidistance of cardinal vowels. Scaling was done with the following formulas.

(3) $SR = 168\ (GM\ F0 / 168)^{1/3}$

(4) $x = \log (SF3 / SF2)$

(5) $y = \log (SF1 / SR)$

(6) $z = \log (SF2 / SF1)$

SR is the sensory reference computed using the geometric mean of the fundamental frequency (F0 in HZ) of the utterance (here the geometric mean of F0 for all the vowel sets was used, GMF0all = 151). SF1-SF3 are the frequency locations of the first three significant spectral prominences of the short term spectral envelope of the vowel waveform (usually equivalent to the first three formants). Miller's theory posits several levels of processing with almost identical scaling equations (the only difference is that the sensory reference and formants are transformed into perceptual reference and formants). The sensory level equations were arbitrarily chosen for use in this study.

Auditory-perceptual scaling of the four cardinal vowel set formants was done (using Equations (3)-(6)), and Euclidean inter-vowel distances were computed in APS (see Table 2). For computational purposes the first three formants of CV8 in PL set 6 were set at the average of the first three formants of the other three vowel sets, F1=188 Hz, F2=376 Hz, and F3=2398 Hz. The third formant of CV8 in PL set 7 was set to 2398 Hz. A repeated measures analysis of variance indicated that there was a significant difference in the size of the interval between the cardinal vowels ($F(5,15) = 3.00$, $p<0.05$). As found for the semit scaling, the front vowel interval did not significantly vary ($F(2,6) = 1.60$, $p<0.30$), but the back vowel intervals were significantly different ($F(2,6) = 12.37$, $p<0.01$)

6. Discussion

The cardinal vowels analyzed in this study were not equidistant in linear or non-linear (Bark, semit, or APS) vowel space. The separate analyses of the front and back vowel series did not produce consistent results. Both series might be equidistant using linear or Bark scaling. (However the front series could be considered marginally different using either scale. The respective p-values of 0.07 and 0.06 could be considered weak support for the idea that the front vowels series is not equidistant.) The other log-based scales, semitone and APS, indicated that the front vowel series might be equidistant and the back vowel series not equidistant. While it is

possible to conclusively prove that cardinal vowels are not equidistant using standard analysis of variance, negative results are more difficult to interpret. Negative results can only be interpreted as possible support for the equidistance of cardinal vowels.

Neary (1989) and Miller (1989) both suggest that log-based scaling is more appropriate for perceptually normalizing vowels, so the Bark, semit, and auditory-perceptual scaling results probably better represent reality. The distance between cardinal vowels was not equidistant for any of these log-based scales, but conclusions based on the partitioned vowel series are equivocal.

The formant center frequencies obtained in this study may not be correct values due to a number of factors such as poor recordings, problems with digital analysis, or incorrect scaling from the physical to perceptual domain. If new, better recordings of cardinal vowels could be obtained, the analyses described here could be repeated and maybe the results would indicate that cardinal vowels are equidistant. The best way resolve the issue of auditory equidistance would be to do a perceptual or psychophysical study of the perceived distance between cardinal vowels using trained subjects.

Cardinal vowels can still be useful even if they are not perceptually equidistant. Cardinal vowels remain valuable as common reference points for the description and classification of vowels, especially when it is not feasible to do acoustic analyses. In addition, they are appropriate as training tools for students of phonetics (Abercrombie 1985). Cardinal vowels are also suitable when a standard vowel set is needed for a study of some aspect of vowels. Cardinal vowels have been used as the stimuli in a study of rounding (Lisker 1989) and for testing a two-formant model of vowel perception (Bladon & Fant 1978).

NOTES

* This research was submitted in partial fulfillment of the requirements for a Ph.D. in the Department of Speech and Hearing Science, University of Illinois-Champaign. Preparation of this article was supported in part by Department of Health and Human Services, Public Health Services Grant DC 00174-08 to Robert C. Bilger. I would like to thank my advisor, Robert C. Bilger, for his support and understanding, C.-W. Kim for his helpful comments on this project and manuscript, and Cynthia Johnson for comments on the manuscript. In addition, I thank Peter Ladefoged and Leigh Lisker for supplying me with the recordings of cardinal vowels used in this study. Address correspondence to Sharon R. Morrison, Department of Speech and Hearing Science, 901 South Sixth Street, Champaign, IL 61820.

[1] Note, however, that the test used to analyze variability, Bartlett's test, is not a robust test (Neter, Wasserman, & Kutner 1985), so Laver's conclusion may be suspect.

[2] This recording was supplied by Leigh Lisker who reported (Lisker 1989) only that the set was copied onto tape from a commercially available record. (Possibly Linguaphone ENG 254/255, 1955).

[3] Reasonable in this case means close to the formants found in other analyses of cardinal vowels (Ladefoged 1960a, 1967; and Lisker 1989) or for similar sounding English vowels (Peterson & Barney 1952).

REFERENCES

ABERCROMBIE, David. 1967. Elements of general phonetics. Chicago: Aldine Publishing.

-----. 1985. Daniel Jones's teaching. Phonetic linguistics, ed. by V. Fromkin, 15-24. Orlando: Academic Press.

AMERICAN NATIONAL STANDARDS INSTITUTE. 1960. Acoustical Terminology, S 1.1.

BLADON, R. A. W., & Gunnar Fant. 1978. A two formant model and the cardinal vowels. Speech Transmission Laboratory Quarterly Progress and Status Report 1978:1.1-8.

HOCKETT, Charles F. 1955. A manual of phonology. Baltimore, MD: Waverly Press.

JONES, Daniel. 1956. An outline of English phonetics (8th ed.). New York: E. P. Dutton.

JONES, Stephen. 1929. Radiography and pronunciation. British Journal of Radiology 2.149-150.

JONGMAN, Allard, Marios Fourakis, & Joan A. Sereno. 1989. The acoustic vowel space of Modern Greek and German. Language and Speech 32:3.221-248.

JOOS, M. 1948. Acoustic Phonetics. Baltimore, MD: Waverly Press.

LADEFOGED, Peter. 1960a. The nature of vowel quality. Revista do Laboratorio de Fonetica Experimental da Faculdade de Letras da Universidade de Coimbra V.

-----. 1960b. The value of phonetic statements. Language 36.387-396.

-----. 1967. Three areas of experimental phonetics. London: Oxford University Press.

-----. 1982. A course in phonetics (2nd ed.). New York: Harcourt Brace Jovanovich.

LAVER, J. D. M. H. 1965. Variability in vowel perception. Language and Speech 8.95-121.

LISKER, Leigh. 1989. On the interpretation of vowel 'quality': The dimension of rounding. Journal of the International Phonetic Association 19:1.24-30.

MILLER, James D. 1987. The auditory-perceptual theory of phonetic recognition: A synopsis. Central Institute for the Deaf, MS.

-----. 1989. Auditory-perceptual interpretation of the vowel. Journal of the Acoustical Society of America 85:5.2114-2134.

MUSON, W. A.. & M. B. Gardner. 1950. Loudness patterns: a new approach. Journal of the Acoustical Society of America 22:2.177-190.

NEARY, Terrance M. 1989. Static, dynamic, and relational properties in vowel perception. Journal of the Acoustical Society of America 85:5.2088-2113.

NETER, John, William Wasserman, & Michael H. Kutner. 1985. Applied linear statistical models (2nd ed.). Homewood, IL: Richard Irwin.

O'CONNOR, Joseph D. 1973. Phonetics. London: Pelican.

PETERSON, Gordon E., & Harold L. Barney 1952. Control methods used in a study of the identification of vowels. Journal of the Acoustical Society of America 50.678-684.

SIGNAL TECHNOLOGY INC. 1987a. PC-ILS v. 6 [Computer Program]. Goleta, CA: Signal Technology Inc.

-----. 1987b. ILS command reference guide v. 6. Goleta, CA: Signal Technology Inc.

STEVENS, Stanley S., & John Volkmann. 1940. The relation of pitch to frequency: A revised scale. American Journal of Psychology 53:3.329-353.

ZWICKER, Eberhard. 1961. Subdivision of the audible frequency range in critical bands (Frequenzgruppen). Journal of the Acoustical Society of America 33:2.248.

-----, & E. Terhardt. 1980. Analytical expressions for critical-band rate and critical bandwidth as a function of frequency. Journal of the Acoustical Society of America 68:5.1523-1525.

Studies in the Linguistic Sciences
Volume 23, Number 1, Spring 1993

THE VEDIC CLAUSE-INITIAL STRING AND UNIVERSAL GRAMMAR*

Steven Schäufele

Examining a variety of clause-peripheral phenomena in Vedic Sanskrit, this paper argues that they are not amenable to an analysis in terms of projections of a functional head COMP, but rather that such a head, for which evidence exists in other languages, is best regarded as a syntactic coalescence of a variety of syntactic and pragmatic functions which UG allows some languages to keep syntactically distinct.

1. COMP-functions

In various portions of the literature of generative grammar, especially those associated with the Principles-&-Parameters Approach (PPA) and its historical predecessors in the Standard-Theoretical School, the label COMP has been associated with linguistic elements performing a variety of functions, hereafter referred to as 'COMP-Functions'. In PPA, these functions are actually associated with a variety of structural positions associated in different ways with the node-label COMP and its various projections, as will be discussed further in §3. In this paper I argue that in Vedic Sanskrit the syntactic realizations of the COMP-Functions must have been structurally different from those assumed in PPA literature for what some (e.g. Whorf, cited by Bruce Nevin, LINGUIST 3-798[1]) have called 'Standard Average European' languages to the point that the status of the node-label COMP as a universal functional head ought at the very least to be reconsidered.[2] As an alternative, I suggest that UG may allow a range of possibilities of syntactic realizations for these functions.

For the purposes of discussion, I identify the COMP-Functions listed in (1).

(1) a. Syntactic Linking
    b. Pronominal-Fronting
    c. Topicalization
    d. Left-Dislocation

By 'Syntactic Linking', I mean the linking between a verb that subcategorizes for a clausal complement and that complement. Such linking is performed by the bold words in (2).

(2) a. I wonder **if** Hilary will get here before 9:00.
    b. She told Terry **that** I was willing to sell my copy of 'Stairway to Heaven'.

c. I can't see **whether** Phil took the blue bell or the red hammer.

'Pronominal-Fronting' refers to the movement to the beginning of a clause of various pronominals as in (3). In modern English such movement affects only interrogative and relative pronominals, most of which begin orthographically with the digraph 'wh'; hence the phenomenon is typically known in generative literature as 'wh-fronting', even in discussions of languages in which interrogative and relative pronominals take radically different forms. As will become clear in the next section, however, cross-linguistically what I am calling 'pronominal-fronting' can have much wider application than it does in English, therefore justifying the broader label.

(3) a. **Whom** did Sam give the money to?
b. **Where** did Morgan find the keys?
c. Hilary showed it to the farmer **whom** you had seen in the drugstore yesterday.

'Topicalization' refers to the fronting of non-pronominals to clause-initial position as in (4). Generally, there is some special pragmatic function associated with such fronting, though it may be difficult to define.[3]

(4) a. Peas_i I like $ə_i$.
b. [Mary Harper]$_i$ I have yet to see $ə_i$ play Rosalind.
c. [Mary Harper]$_i$ I have yet to see a more enchanting actress than $ə_i$.

'Left-Dislocation' differs from Topicalization chiefly in that it involves a pronominal 'copy' of the fronted constituent, as in (5).[4]

(5) a. [That man]$_i$, I don't like him$_i$.
b. [That man]$_i$, I don't like his$_i$ dog.
c. [Ralph]$_i$, I wish he$_i$ would take his opinions elsewhere.

In generative literature, as noted above, the notion of a functional head labelled COMP has been invoked in describing all of these phenomena, though not all at the same time. Early discussion within generative literature of such a notion (e.g., Rosenbaum 1967, Bresnan 1972) was concerned primarily with Syntactic Linking. Chomsky 1977 focussed on Wh-Fronting, with some discussion of Topicalization and Left-Dislocation. Subsequent work has tended to focus on Wh-Fronting and Topicalization, while acknowledging Syntactic Linking as relevant.

In this paper I show that in Vedic Sanskrit most of these functions were syntactically distinct, and that therefore the use of the notion of a structural node labelled COMP and/or its various projections in describing that language may be misleading and therefore inappropriate.[5]

2. Vedic COMP-realizations

In this section I discuss how the different COMP-Functions identified above are realized in Vedic Sanskrit. The lexical and syntactic strategies discussed here will be referred to as 'COMP-Realizations'.

Sanskrit, like many other South Asian languages (cf. Hock 1989), avoids syntactic subordination, and in fact as noted in Schäufele 1990b:64-66 and references cited there, it has no subordination of finite clauses and, to the extent it has verbs subcategorizing for clausal complements, they are verbs used in reporting quoted discourse and the various metaphorical extensions thereof, as discussed in Hock 1982b; cf. (6). In such constructions, the syntactic link between the quoted material (represented in (6) by the material between quotes in the Sanskrit text) and the 'matrix' (represented in (6) by everything else) is effected by the word *íti*.

(6) a. yáḥ índrāya 'sunávāma' **íti** āha
REL-N / Indra-D / press-1p.impv. / Q / say-3s.pres.
'Who says to Indra, "Let us press"' (RV 5.37.1d)

b. tám āhuḥ 'suprajā́ḥ' **íti**
DEM-Ams. / say-3p.pres. / rich in progeny-N / Q
'They call him "rich in progeny".' (RV 9.114.1c)

c. yát vā pravṛddha satpate 'ná marai'
REL / or / great-V / sovereign-V / not /die-1s.subj.
íti mányase
Q / think-2s.pres.
'Or when, great sovereign, you think "I shall not die"' (RV 8.93.5ab)

d. 'bhū́miḥ' **íti** tvā́m abhi-prá-manvate jánāḥ 'nírṛtiḥ' **íti**
earth-Ns. / Q / 2s.pro.-A / think-3p.pres. / people-Np. / Nirṛti-N / Q
tvā ahám pári-veda sarvátaḥ
2s.cl.-A / 1s.pro.-N / know-1s.pres. / completely
'People think of you as "earth", (but) I know you completely as "Nirṛti".' (AV 6.84.1cd)

e. svayám u hí evá etát véda 'idám
himself-N / PTCL / PTCL / PTCL / DEM-A / know-3s.pres. / DEM-N
átaḥ kárma kartávyam **íti**
now / deed-N / to be done-N / Q
'For he himself knows that this deed is now to be done.' (ŚB 1.2.5.21)

Most if not all Indo-European languages routinely front certain pronominals into clause-initial position, and Vedic was no exception. Vedic pronominal-fronting differs from English wh-fronting in three obvious ways (cf. Schäufele (1988) for further discussion).

Whereas in English only 'wh-elements', i.e., interrogatives and relatives, are fronted, (cf. (7)), in Vedic demonstratives were affected in exactly the same way; cf. (8).

(7) a. [Which book]$_i$ did Lynn buy ə$_i$?
b. The book that$_i$ Lynn bought ə$_i$ has been missing for 3 days.
c. ?*Did Lynn buy which book?
d. Lynn bought that book.
e. ?[That book]$_i$ Lynn bought ə$_i$.
f. *[The book]$_i$ Lynn bought ə$_i$ that has been missing for 3 days.

(8) a. **téṣu**$_i$ evá enām$_j$ etát ə$_i$ pari-ā́sīneṣu
DEM-Lp. / PTCL / 3f.s.cl-A / DEM / ə / sitting around-Lp.
ánagnāṃ ə$_j$ karoti
non-naked-A / ə / make-3s.pres.
'He makes her (to be) non-naked among those sitting around.' (ŚB 1.3.3.8)

b. **etā́m**$_i$ u evá **eṣáḥ**$_j$ **etásmai**$_k$ ə$_j$ víṣṇuḥ
DEM-Af. / PTCL / PTCL / DEM-Nm. / DEM-Dm. / ə / Viṣṇu-N
yajñáḥ ə$_k$ ə$_i$ vikrā́ntim ví-kramati
sacrifice-Nm. / ə / ə / victorious step-Af. / stride-3s.pres.
'This sacrifice, Viṣṇu, strides this victorious step for him.' (ŚB 1.1.2.13)

Whereas in English only one wh-element can be fronted (cf. (9)), in Vedic an indefinite number can be fronted; cf. (8b). In this respect, Vedic resembles other languages which have been discussed by e.g. Rudin (1988) and McDaniel (1989); cf. (10).

(9) a. **Where** did **who** give **what**?
b. ***Where who what** did give?

(10) a. **Ko kas** mislinol so o Demìri cuminja? (Romani)
who / whom / think-3s. / that / Demìri-N / kissed
'Who thinks that Demir kissed whom?'

b. **Kto čto kogda** skazal? (Russian)
who / what / when / said
'Who said what to whom?'

c. **Koj kakvo na kogo** e dal? (Bulgarian)
who / what / to / whom / has / given
'Who gave what to whom?'

d. **Kto komu jaką** by napisał ksiażkę? (Polish)
who / to whom / what kind / would / write / book
'Who would write what kind of book to whom?'

e. **Cine cui ce** ziceai că i a promis? (Rumanian)
who / whom-O / what / said-2s. / that / him-O / has / promised
'Who did you say promised what to whom?'

However, whereas in English any clause including one wh-element is unacceptable unless that wh-element is either fronted or highly emphasized (cf. (11)), the fronting of pronominals is not obligatory in Vedic; cf. (12). It is even possible to have in the same clause some pronominals fronted into clause-initial position while others remain in their DS positions; cf. (13).

(11) a. *This man here is who?
b. Who is this man here?
c. This man here is **who**?

(12) a. áheḥ yātā́ram **kám** apaśya indra
serpent-G / avenger-A / INT-A / see-2s.impf. / Indra-V
'Which avenger of the serpent did you see, O Indra? = Whom did you see, O Indra, as avenger of the serpent?' (RV 1.32.14a)
b. hárī índrasya ní-cikāya **káḥ** svit
horses-Ad. / Indra-G / perceive-3s.perf. / INT-N / PTCL
'Who perceived the two steeds of Indra?' (RV 10.114.9d)
c. **té** devā́ḥ **etát** yájuḥ apaśyan
dDEM-N / gods-N / dDEM-A / Yajus-A / see-3p.impf.
'The gods saw this Yajus.' (TS 5.2.3.1)
d. evam eva **etat** akṣaram **etā́bhiḥ** devatābhiḥ
thus / PTCL / dDEM-N / syllable-N / dDEM-I / deities-I
sam-sṛjyate
unite-3s.pres.pass.
'Thus this syllable is united with these deities.' (JUB 1.1.3.8)

(13) **etám** evá asmin **etát** sántatam
DEM-Am. / PTCL / 3s.cl.-L / DEM-Adv. / extended-Am.
ávyavachinnam dadhāti **etát** anuvácanam
not cut off-Am. / grant-3s.pres. / dDEM-Nn. / recitation-Nn.
'So this recitation grants him (the state of being) extended (in lifespan), not cut off (by death).' (ŚB 1.3.5.13)

Whereas in English what is fronted is an entire NP or PP containing a wh-element (cf. (14)), in Vedic only the pronominal word itself is fronted, leaving behind the NP for which it serves as determiner; cf. (15).

(14) a. Whom$_i$ did you give ə$_i$ the books?
b. [Which books]$_i$ did you give your sister ə$_i$?
c. [How many papers on Celtic verb-fronting]$_i$ did you receive ə$_i$?
d. [From how many of your students]$_i$ did you get papers ə$_i$?

(15) a. **etát**$_i$ ha vaí devā́ḥ [ə$_i$ vratám]$_{NP}$
dDEM-As. / PTCL / PTCL / gods-Np. / ə / vow-As.
caranti
undertake-3p.pres.
'The gods undertake this vow.' (ŚB 1.1.1.5)
b. **yā́bhyaḥ**$_i$ eva **tā́ni**$_j$ [ə$_i$ devatābhyaḥ]$_{NP}$ [ə$_j$ havī́ṁṣi]$_{NP}$
REL-Dp. / PTCL / dDEM-Np. / ə / deities-Dp. / ə / oblations-Np.
gṛhītāni bhavanti
take-pass.part.-Np. / be-3p.pres.
'For which deities those oblations are taken' (AB 7.2.3)
c. **káti**$_i$ **ayám**$_j$ adyá [ə$_j$ udgātā́]$_{NP}$ [asmín yajñé]$_{NP}$
INT-Ap. / DEM-Ns. / today / ə / Udgātṛ-Ns. / DEM-Ls. / sacrifice-Ls.

[ə$_i$ stotríyāḥ]$_{NP}$ stoṣyati
ə / stotriya-Ap. / sing-3s.fut.
'How many stotriyas will the Udgātṛ sing in this sacrifice today?' (ŚB 14.4.6.12)

An additional difference between English and Vedic, well known to Indo-Europeanists, is that clitic pronouns in Vedic also routinely get fronted into the clause-initial string in conformity with Wackernagel's Law (cf. (16)),[6] while in English (cf. (17)) they cliticize obligatorily to the elements that govern them for purposes of θ- and Case-assignment.

(16) a. ā́$_i$ **vām**$_j$ váyaḥ áśvāsaḥ
Pfx. / 2d.cl.-A / winged-Np. / horses-Np.
váhiṣṭhāḥ [abhí práyaḥ nāsatyā
finest draught animals-Np. / to-Adp(A) / offering-A / Nāsatyas-V
[ə$_i$-vahantu = ə$_j$]$_V$]$_{VP}$
ə / bring3p.impv. / ə
'May the winged horses, the very finest draught animals, bring you to the offering, O Nāsatyas.' (RV 6.63.7ab)

b. prá **vaḥ**$_i$ śárdhāya ghṛ́ṣvaye tveṣádyumnāya śuṣmíṇe
Pfx. / 2p.cl.-G / host-D / vigorous-D / fierce-D / wild-D
[deváttam ə$_i$ bráhma]$_{NP}$ gāyata
inspired-A / ə / hymn-A / sing-2p.impv.
'Sing your inspired hymn to the vigorous, fierce, wild host.' (RV 1.37.4)

c. vṛtrám hí **asmai**$_i$ etát jaghnúṣe
Vṛtra-A / PTCL / 3s.cl.-D / DEM-An. / slay-perf.part.-D
āpyā́yanam [ákurvan = ə$_i$]$_V$
envigorating draught-An. / make-3p.impf. / ə
'For him, (the one) having slain Vṛtra, they prepared this envigorating draught.' (ŚB 1.6.4.12)

d. té **enam**$_i$ etát vratám upayántam
DEM-Np. / 3s.cl.-A / DEM-As. / vow-A / undertaking-A
[viduḥ = ə$_i$]$_V$
know-3p.pres. / ə
'They know him (to be) the one undertaking this vow.' (ŚB 1.1.1.7)

(17) a. Tomorrow mornin' Claude's gonna sell='**em** to **y'r**=neighbors.

b. *Tomorrow mornin' '**em** Claude's gonna sell to **y'r**=neighbors.

c. *Tomorrow '**em** mornin' Claude's gonna sell to **y'r**=neighbors.

d. *Tomorrow mornin' **y'r** Claude's gonna sell='**em** to neighbors.

e. *Tomorrow **y'r** mornin' Claude's gonna sell='**em** to neighbors.

f. *Tomorrow mornin' **y'r** '**em** Claude's gonna sell to neighbors.

g. *Tomorrow **y'r 'em** mornin' Claude's gonna sell to neighbors.

Topicalization is attested in Vedic as in probably all other human languages.[7] But whereas in English, topicalization fronts whole NPs but not single words out of NPs (cf. (18)), in Vedic while the fronting of NPs is possible (cf. (19)) the fronting of single words out of phrases is preferred;[8] cf. (20) and see Hock 1982a and Schäufele 1985, 1986, 1991, and 1993a for further discussion.[9]

(18) a. [The neighbors' dog]$_i$ I haven't heard from ə$_i$ yet.
b. *Dog$_i$ I haven't heard from the neighbors' ə$_i$ yet.
c. *Neighbors'$_i$ I haven't heard from the ə$_i$ dog yet.

(19) a. [áṅgirasaḥ suvargám lokám yatáḥ]$_i$
Aṅgirases-Ap. / heavenly-As. / world-As. / go-pres.part.-Ap.
puroḍā́śaḥ kūrmáḥ bhūtvā́ ə$_i$ ánu prá-asarpat
cake-N / tortoise-N / become-ger. / ə / after-Adp(A) / crawl-3s.impf.
'Becoming a tortoise, the cake crawled after the Aṅgirases (who were) going to heaven.' (TS 5.2.8.4)
b. [rājñā somena]$_i$ tat vayam ə$_i$ asmāsu dhārayāmasi
king-I / Soma-I / DEM-A / 1p.pro.-N / ə / 1p.pro-L / guard-1p.pres.
'We guard it for ourselves by means of King Soma.' (KS 29.2)

(20) a. dviṣántam$_i$ ha asya tát [[ə$_i$ bhrā́tṛvyam]$_{NP}$
hateful-A / PTCL / 3s.cl.-G / DEM-N / ə / foe-A
abhi-áti-ricyate]$_{VP}$
remain for-3s.pres.
'That remains for his hateful foe.' (ŚB 1.9.1.18)
b. óṣadhīnām$_i$ vaí sáḥ [[ə$_i$ mū́lāni]$_{NP}$ úpa-amlocat]$_{VP}$
plants-G / PTCL / DEM-N / ə / roots-A / hide among-3s.impf.
'He hid among the roots of plants.' (ŚB 1.2.5.10)
c. ā́$_i$ ha vaí asmin svā́ḥ ca níṣṭyāḥ ca
Pfx. / PTCL / PTCL / 3s.cl.-L / own-Np. / & / strangers-Np / &
ə$_i$-śaṁsate
ə / place trust-3p.pres.mid.
'In him both his own (relatives) and strangers place their trust.' (ŚB 1.6.4.17)
d. mánaḥ$_i$ ha vaí devā́ḥ [[manuṣyàsya ə$_i$]$_{NP}$ ā́-jānanti]$_{VP}$
mind-A / PTCL / PTCL / gods-N / man-G / ə / know-3p.pres.
'The gods know the mind of man' (ŚB 1.1.1.7)
e. brahma$_i$ ca vai idam agre
Brahman-N / PTCL / PTCL / here / beginning-L
[ə$_i$ subrahma ca]$_{NP}$ āstā m
ə / Subrahman-N/PTCL / be-2d.impf.
'In the beginning, there were here Brahman and Subrahman.' (ṢB 1.1.1)
f. ṛ́ṣibhyaḥ$_i$ ca evá enam etát [ə$_i$ devébhyaḥ ca]$_{NP}$
Ṛsis-Dp. / PTCL / PTCL / 3s.cl.-A / DEM / ə / gods-D / PTCL

ní-vedayati
introduce-3s.pres.
'He introduces him to the Ṛsis and to the gods.' (ŚB 1.4.2.3)

And finally, Left-Dislocation is attested in Vedic; cf. (21) and see Oertel 1926 for a detailed discussion.

(21) a. [datvaḥ ca ha sautemanasaḥ mitravit ca
Datva-N / & / PTCL / Sautemanasa-N / Mitravit-N / &
daṁṣṭradyumnaḥ]$_{NP}$ [tau ha pratidarśasya
Daṁṣṭradyumna-N / DEM-Nd. / PTCL / Pratidarśa-G
vaibhāvatasya svaiknasya rājñaḥ brahmacāriṇau āsatuḥ]$_{S}$
Vaibhāvata-G / Śvaikna-G / king-G / pupil-Nd. / be-3d.perf.
'Datva Sautemanasa and Mitravit Daṁṣṭradyumna — these were King Pratidarśa Vaibhāvata Śvaikna's pupils.' (JB 2.274)

b. [átha etát sárīram]$_{NP}$ [tásmin ná rásaḥ asti]$_{S}$
NEXUS / DEM-Nn. / body-Nn. / DEM-Ln. / not / sap-N / be-3s.pres.
'Now this body — there is no sap in it.' (ŚB 4.4.5.1)

c. [yajñāya yajamānāya ātmane]$_{NP}$ [tebhyaḥ eva āśiṣam
sacrifice-Ds. / sacrificer-Ds. / self-Ds. / DEM-Dp. / PTCL / blessing-A
ā-śāste]$_{S}$
invoke-3s.pres.
'For the sacrifice, for the sacrificer, for himself — for these he invokes a blessing.' (TB 2.6.9.3)

3. The syntax of COMP-realizations

3.1. Standard PPA assumptions

In PPA literature (cf. e.g. Chomsky 1986:3) a constituent COMP is typically assumed. Furthermore, this constituent is defined as a Bar-Ø head of a complete projection, CP, with its own complement and specifier. In line with standard $\bar{X}$-theory (cf. Chomsky, ibid.), both complement and specifier are assumed to be themselves maximal projections. COMP° itself is the site of Syntactic-Linkers, while its complement is the clause (IP) 'introduced' thereby. Cf. (22).

(22) a. [$_{IP}$ I [$_{VP}$ wonder [$_{CP}$[$_{COMP}$ if] [$_{IP}$ Hilary will get here before 9:00]]]]?

b. [$_{IP}$ She [$_{VP}$ told Terry [$_{CP}$[$_{COMP}$ that] [$_{IP}$ I was willing to sell my copy of 'Stairway to Heaven']]]].

c. [$_{IP}$ I can't [$_{VP}$ see [$_{CP}$[$_{COMP}$ whether] [$_{IP}$ Phil took the blue bell or the red hammer]]]].

Spec of COMP is assumed to be the landing-site of wh-movement as in (23). The fact that a Spec position is a maximal projection would account for the fact that in English and other Standard Average European languages pronominal-fronting affects whole phrases as in (24).

(23) a. [$_{CP}$[$_{Spec}$ Whom$_i$] [$_{COMP}$ did$_j$] [$_{IP}$ Sam ə$_j$ [$_{VP}$ give the money to ə$_i$]]]?

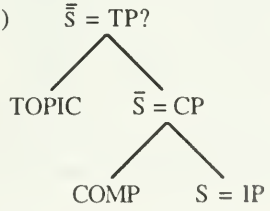
b. [$_{CP}$[$_{Spec}$ Where$_i$] [$_{COMP}$ did$_j$] [$_{IP}$ Morgan ə$_j$ [$_{VP}$ find the keys ə$_i$]]]?

(24) a. [CP[Spec[Which books]$_i$] [COMP did$_j$] [IP you e_j [VP give your sister e_i]]]?
b. [CP[Spec[How many papers on Celtic verb-fronting]$_i$] [COMP did$_j$] [IP you e_j [VP receive e_i]]]?
c. [CP[Spec[From how many of your students]$_i$] [COMP did$_j$] [IP you e_j [VP get papers e_i]]]?

The landing-site of Topicalization and Left-Dislocation may also be Spec of COMP.[10] Or it may be an adjunction-site to IP or CP. Note that in PPA (cf. Chomsky 1986:88) adjunction is typically assumed to be of maximal projections to maximal projections. Adjunction of X°-heads to heads is also possible; cf. Baker (1988). But the consensus is that an X^{max} can only accept another X^{max} in adjunction, and likewise an X° another X°. This is consistent with the Structure-Preservation Constraint (cf. Emonds 1976). Thus, if Topicalization and/or Left-Dislocation are examples of adjunction then they involve maximal projections. The same conclusion, of course, follows from the definition of Topicalization and/or Left-Dislocation as movement to Spec of COMP.

An alternative hypothesis, proposed in Chomsky 1977 and mentioned here for the sake of completeness and because it will be at least tangentially relevant to later discussion, would define an extra TOPIC position as aunt to (i.e., sister to the mother of) COMP. Cf. (25).[11] On the face of it, this hypothesis would appear to define TOPIC as the Bar-Ø head of a separate phrasal projection, with its own specifier and complement, which latter would presumably be CP.

(25)



3.2. The Vedic situation

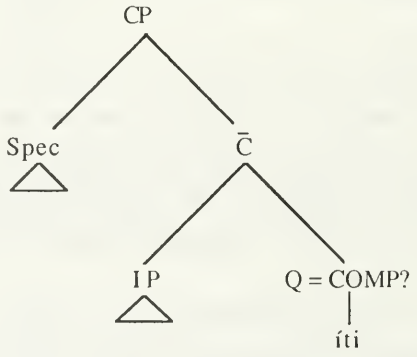
It will be noticed that the theory outlined in the previous section posits a relatively large number of maximal projections and/or landing-sites for maximal projections and a relatively small number of Bar-Ø nodes. COMP itself is a Bar-Ø node and a second Bar-Ø node may exist in TOPIC, but each of these has its own Specifier position, presumably X^{max}, and the possibility of adjunction creates more landing sites that, according to the assumption already mentioned, are restricted to maximal projections.

On the other hand, the Vedic COMP-Realizations discussed in §2 are mostly Bar-Ø in nature. While phrasal topicalization is attested and was therefore presumably possible, the corpus as a whole and

especially the prose portions thereof show a strong if not overwhelming preference for X°-fronting. And while clitic pronouns are syntactically NPs in their own right, the demonstratives, interrogatives, and relatives whose fronting is Sanskrit's equivalent of 'wh-fronting' are merely NP-determiners, not phrases themselves (cf. Schäufele 1988 for further discussion). And, as mentioned earlier (n. 7) and discussed in Schäufele 1991, 1993a, to assert that in spite of appearances all these fronted words are really bare-headed phrases is to drain the theory of falsifiability and therefore of interest. Ideally, what would be desired is a number of Bar-Ø landing sites (which might themselves serve as adjunction sites) without accompanying specifier and complement positions. Cf. further discussion in §5.

The question of the location of these landing-sites vis-à-vis IP raises further complications. It will be noted that *íti*, which is apparently the prime candidate for the realization of the function of Syntactic Linking, is generally in Vedic a clause-final element.[12] It therefore differs from English COMP in this respect. Generative linguists studying Japanese, which like Sanskrit is an obviously head-final language, have claimed that COMP in that language is post-clausal rather than pre-clausal as it is in English. However, as may be obvious from the examples above and will be discussed below, all the other 'COMP-Realizations' in Vedic Sanskrit are clause-initial phenomena. It has been suggested (Davison, Yoon p. c.) that *íti* might indeed represent a clause-final COMP, with a pre-clausal Spec position serving the other 'COMP-Realizations'; cf. the structure outlined in (26). This would be consistent with common assumptions about the internal structure of head-final, left-branching constituents and in fact with the structure assumed in Schäufele 1990b for NPs and other projections of lexical categories in Vedic, except that the analysis presented there seeks to avoid any intermediate level of structure between the Bar-Ø head and the maximal projection thereof. I discuss some of the possible inadequacies of the analysis in (26) later, but I will assume it for the immediate discussion.[13]

(26)



The other COMP-Realizations listed in §2 need to be classified as CLAUSE-INITIAL vs. PRE-CLAUSAL. Clause-initial elements are treated by the syntax as part of the clause but happen to surface at its left periphery; pre-clausal elements are treated by the syntax as outside the clause altogether.

The principal diagnostic in Vedic for the clause-boundary, in terms of which clause-initiality and pre-clausality should of course be defined, is the location of 'sentential particles'. These elements, labelled 'PTCL' in all the examples in this paper, are words with pragmatic rather than semantic content.[14] Syntactically they routinely follow immediately the first phonological word (cf. Schäufele 1986, 1990b for discussion of various complications in the definition thereof) in whatever larger syntactic constituent defines their 'pragmatic scope', most typically the clause.[15] Ergo, the (phonological) word immediately before any set of sentential particles is, by definition, the first word in the clause; anything before it must therefore be outside the clause.

On the basis of this criterion, the Vedic realization of Left-Dislocation must be viewed as pre-clausal. In (21a), repeated below as (27), the two occurrences of the sentential particle *ha* must belong to distinct 'clauses'; the word immediately preceding the second *ha*, the resumptive pronoun *tau*, must therefore be the first word of its clause and all the material to its left must be extra-clausal.[16]

(27) [datvaḥ ca ha sautemanasaḥ mitravit ca
Datva-N / & / PTCL / Sautemanasa -N / Mitravit-N / &
daṁṣṭradyumnaḥ]$_{NP}$ [tau ha pratidarśasya
Daṁṣṭradyumna-N / DEM-Nd. / PTCL / Pratidarśa-G
vaibhāvatasya śvaiknasya rājñaḥ brahmacāriṇau āsatuḥ]$_{S}$
Vaibhāvata-G / Śvaikna-G / king-G / pupil-Nd. / be-3d.perf.
'Datva Sautemanasa and Mitravit Daṁṣṭradyumna — these were King Pratidarśa Vaibhāvata Śvaikna's pupils.' (JB 2.274)

Topicalization and pronominal-fronting in Vedic, on the other hand, are clearly associated with a clause-initial landing site.[17] As can be seen from (20), the topicalized word is routinely followed immediately by the sentential-particle string. As for pronominal fronting, the examples repeated below as (28) show that the particle string is likely to follow immediately a fronted pronominal (28a); if there is more than one fronted pronominal, then the particle string will follow the first one (28b).

(28) a. etát$_{i}$ ha vaí devā́ḥ [ə$_{i}$ vratám]$_{NP}$
dDEM-As. / PTCL / PTCL / gods-Np. / ə / vow-As.
caranti
undertake-3p.pres.
'The gods undertake this vow.' (ŚB 1.1.1.5)

b. etắm$_i$ u evá eṣáḥ$_j$ etásmai$_k$ ə$_j$ víṣṇuḥ
DEM-Af. / PTCL / PTCL / DEM-Nm. / DEM-Dm. / ə / Viṣṇu-N
yajñáḥ ə$_k$ ə$_i$ vikrắntim ví-kramati
sacrifice-Nm. / ə / ə / victorious step-Af. / stride-3s.pres.
'This sacrifice, Viṣṇu, strides this victorious step for him.' (ŚB 1.1.2.13)

4. Discourse Linking

A brief digression is in order to discuss an extra function which has not been among the foci of interest in the study of Standard Average European languages with regard to COMP, but which Vedic syntax invites into this examination. This is 'Discourse Linking', by which I mean the linking together (often by means of a special class of adverbs) of sentences which are syntactically and semantically complete in themselves but which are pragmatically linked in a discourse context such that said linking is a necessary aspect of their relevance and cooperativeness. Cf. the bold words in (29).

(29) Greg had been wanting for weeks to buy a new pair of shoes. **But** Irene had told him that the local shoe-store would be having a sale next month. **So** Greg chose to wait a while longer. **However**, unbeknownst to both Irene and Greg, circumstances were compelling the store owner to change his plans and to delay the anticipated sale. **Meanwhile**, on the other side of town a new shoe store was opening, with a major sale of new stock. **Fortunately,** Greg's friend Hilary, who occasionally had business to do in that area, found out about this and told Greg. **So** Greg was able to get new shoes in spite of the delay at the local shoe store.

Syntactically and semantically, every sentence in (29) could stand by itself without the bold words.

Discourse Linking in Vedic is done by means of such words as *átha*, *táthā*, and *tátaḥ*. Cf. (30). These words (which in this role I shall label 'nexus', following Minard 1936 and Klein 1987) can be related to pronominal stems. Some nexus words are more obviously derived from pronominal stems. As Delbrück (1888:§140) notes, the nominative masculine and neuter singular demonstratives *sáḥ* and *tát* can serve as nexus (though he doesn't use that label); cf. (31). The ablative singular *tásmāt* can be similarly used; cf. (32). By means of such words, extensive discussions can be held together; cf. (33).

(30) a. sáḥ ha uvāca videgháḥ māthavàḥ ... átha
DEM-N / PTCL / speak-3s.perf. / Videgha / Māthava-N / ... / NEXUS
ha uvāca gótamaḥ rāhūgaṇáḥ ...
PTCL / speak-3s.perf. / Gotama / Rāhūgana-N
'Videgha Māthava said ... Then Gotama Rāhūgana said ...' (ŚB 1.4.1.17-18)

b. táthā u evá eṣáḥ eténa vájreṇa ā́jyena
NEXUS / PTCL / DEM-N / DEM-I / thunderbolt-I / butter-I
ṛtū́n samvatsarám prá-jayati
seasons-A / year-A / win-3s.pres.
'Likewise with that thunderbolt (which is) the butter he wins the seasons, the year.' (ŚB 1.5.3.4)

c. tátaḥ dvā́bhyām brāhmaṇā́ḥ yajñé cáranti
NEXUS / two-I / Brahmin-Np. / sacrifice-L / use-3p.pres.
'Therefore, Brahmins use two in the sacrifice.' (ŚB 1.2.4.2)

(31) a. sáḥ yátra ha evám ṛtvíjaḥ samvidānā́ḥ yajñéna
NEXUS / REL / PTCL / thus / priests-Np. / agreeing-Np. / sacrifice-I
cáranti
perform-3p.pres.
'And when the priests in complete agreement thus perform the sacrifice ... ' (ŚB 1.5.2.15)

b. tát svéna evá enam etát páyasā devā́ḥ
NEXUS / own-I / PTCL / 3s.cl.-A / thus / juice-I / gods-N
sví-akurvata
acquire-3p.impf.mid.
'So the gods acquired it by means of its own juice.'
(ŚB 1.5.3.5)

(32) tásmāt gardabháḥ ápi anāleśé ati
NEXUS / donkey-N / even / poor grazing-L / beyond-Adp(A)
anyā́n paśū́n medyati
other-A / animals-A / get fat-3s.pres.
'Therefore the donkey, even in conditions of poor grazing, gets fat beyond other animals.' (TS 5.1.5.5)

(33) sáḥ [yáḥ kapā́lāni upa-dádhāti], [sáḥ
NEXUS / REL-N / potsherds-A / put down-3s.pres. / DEM-N /
upaveṣám a-datte] ['dhṛ́ṣṭiḥ asi'] íti. sáḥ
Upaveṣa-A / pick up-3s.pres. / bold-N / be-2s.pres. / Q / NEXUS
[yát enena agním dhṛ́ṣṭi iva apa-cárati], [téna
REL / 3s.cl.-I / fire-A / boldly / PTCL / attack-3s.pres. / DEM-I
átha dhṛ́ṣṭiḥ]. [yát enena yajñé upa-ā-lábhate], [úpa
NEXUS / bold-N / REL / 3s.cl.-I / sacrifice-L / touch-3s.pres. / Pfx.
evá vaí eténa etát veveṣṭi], tásmāt [upaveṣáḥ
PTCL / DEM-I / DEM-A / attend-3s.pres. / NEXUS / Upaveṣa-N
nā́ma].
by name
'Now he who puts down the potsherds picks up the Upaveṣa (shovel) (saying) 'You are bold'. Now because with it he attacks the fire boldly (dhṛṣṭi), as it were, that's why (it is called) 'Dhṛṣṭi'. And because with it he touches (the coals) in the sacrifice, he attends (úpa-viṣ) it (the fire) with it (the shovel), that's why it (the shovel) is called 'Upaveṣa'.
(ŚB 1.2.1.3)

To the extent that constituents like the bold words in (29) have been discussed at all in the Standard-Theoretical literature, it tends to be assumed that they are merely clause-introductory adverbials. As noted above, note 4, the quotative marker *íti* which has been identified as Vedic's only realization of the Syntactic-Linking function also serves a similar function on occasion, as in (34).

(34) íti bravīti vaktárī rárāṇaḥ vásoḥ vasutvā́
Q / say-3s.pres. / speaker-N / generous-N / good-G / goodness-I
kā́rávaḥ anehā́ḥ
singers-N / guiltless-Np.
'(Thus) says the generous speaker, "Through the goodness of the good, the singers are guiltless."' (RV 10.61.12bc)

However, such collocations do not excuse us from including *íti* in a discussion of Vedic COMP-Realizations. For one thing, as noted above they are very rare compared with those exemplified in (6) in which *íti* is used as a link between quoted material and the matrix. Furthermore, there is another usage of *íti* which is in my opinion more essentially syntactic. As noted by Hock (1982b:55-57), at least by the time of the Atharva-Veda, i.e., mid-late Vedic period, *íti* was used as a means of assimilating onomatopoetic elements and other non-Sanskrit lexemes into the syntax of a Vedic clause without having to bother with fitting them out with Sanskritic morphological trappings as was the common strategy in the RV (cf. Hock 1982b:56). Cf. (35).

(35) a. pṛthivyā́m te nipécanam bahíḥ te astu
ground-L / 2s.cl.-G / outpouring-N / outside / 2s.cl.-G / be-3s.impv.
'bā́l' íti
splash / Q
'May your outpouring be on the ground, outside of you, "splash".' (AV 1.3.1-9 refrain)

b. ajéna kṛṇvántaḥ śītám vṛṣéṇa ukṣantu
goat-I / making-Np. / cool-A / rain-I / sprinkle-3p.impv.
'bā́l' íti
splash / Q
'Making (you) cool with the goat, let them sprinkle (you) with rain, "splash".' (AV 18.2.22)

c. 'bhúk' íti abhí-gataḥ 'śā́l' íti apá-krāntaḥ 'phā́l'
bounce / Q / come-Ns.part. / whist / Q / go-Ns.part. / bang
íti abhí-ṣṭhitaḥ
Q / tread-Ns.part.
'"Bounce" it has come; "whist" it has gone; "bang" it has trodden.' (AV 20.135.1)

In these clauses *íti* is performing a purely grammatical function, in a sense 'naturalizing' elements that are foreign to the grammar. Therefore, there can be no doubt that, at least in cases like these, its presence and position must receive a grammatical account as a functional category.

Returning to the question of Discourse Linkers, it is not enough to characterize them as clause-initial adverbials. In English, different Discourse Linkers are treated differently by the grammar. Witness the freedom of ordering exhibited by the discourse-linking pronominal *therefore* that is not enjoyed by the conjunction *but* in (36). Presumably, this difference relates to the syntactic difference between conjunctions, which typically and cross-linguistically are severally restricted in position, and other 'adverbials'. Even vocative adjuncts cannot precede clause-conjunctions in English; cf. (37). Clearly, returning to the distinction introduced earlier, clause-conjunctions in English must be pre-clausal.

(36) a. i. Therefore, there's no way of being certain.
ii. But there's no way of being certain.
b. i. There's therefore no way of being certain.
ii. *There's but no way of being certain.
c. i. There's no way therefore of being certain.
ii. *There's no way but of being certain.
d. i. There's no way of being certain therefore.
ii. *There's no way of being certain but.

(37) a. But there's no way of ever being certain, one way or another, Commander.
b. But there's no way of ever being certain, Commander, one way or another.
c. But there's no way, Commander, of ever being certain, one way or another.
d. But, Commander, there's no way of ever being certain, one way or another.
e. *Commander, but there's no way of ever being certain, one way or another.

In Vedic, nexus words always occur at the left periphery of their associated clauses. But are they clause-initial or pre-clausal? In (38) we see examples of nexus words followed by other words which are in turn followed by sentential particles. So evidently in cases like these it is not the nexus word, the discourse-linking element, that is the first word in the clause but the word immediately following it.

(38) a. átha kím u yáḥ devéṣu ánaśnatsu
NEXUS / INT / PTCL / PTCL / REL-N / gods-L / eating-L
púrvaḥ aśnīyā́t
before-Adp(L) / eat-3s.opt.
'But what if he should eat before the gods eat?' (ŚB 1.1.1.8)
b. átho manasā́ vaí prajā́patiḥ yajñám
NEXUS / mind-I / PTCL / Prajāpati-N / sacrifice-A
atanuta
perform-3s.impf.
'Then Prajāpati performed the sacrifice with his mind.' (TS 1.6.8.4)

c. sáḥ yáthā iva ha tát agnéḥ bhávati
NEXUS / REL-Adv. / PTCL / PTCL / DEM-N / fire-G / be-3s.pres.
'In that case, that (feud) is (= would be) (on the part) of the fire.' (ŚB 1.1.1.21)

However, it should be noted that there is some variation with regard to discourse linkers. It is possible for 'nexus' words to be followed immediately by sentential particles. Cf. (39).[18] This variation is doubtless due to the fact that these elements, like their equivalents in other Indo-European languages (such as 'therefore' in (36)), are etymologically pronouns. In some cases they may refer directly to (possibly empty) constituents in the clauses they introduce, in which case they may be treated syntactically as real members of those clauses, while in others they may be serving solely as discourse linkers, which if they refer to anything at all are referring to the previous discourse context in general. Similar variation obtains in English; cf. (40). In (40a), parallel to constructions like (41), the word 'then' is a pronominal referring to a specific time that has (presumably) been identified in the preceding discourse and which is functioning in this sentence as an adverbial but is overtly represented solely by this pronominal; being a pronominal, in this sentence 'then' is immediately followed by the verb, which has been fronted ahead of the subject in the manner typical of questions. (40b), on the other hand, is not asking 'will you listen to me at a particular time' but 'you are going to listen to me, aren't you? That's what I deduce from your previous remarks'. In which case the word 'then' if it is referential at all refers to 'the previous remarks', i.e. the entire discourse context. In this case, the verb is not fronted ahead of the subject, and 'then' is acting as a mere discourse linker.

(39) a. utá sma asya panayanti jánāḥ jūtím
NEXUS / PTCL / 3s.cl.-G / praise-3p.pres. / people-N / zeal-A
'And the people praise his zeal.' (RV 4.38.9ab)
b. átha ha sómaḥ uvāca
NEXUS / PTCL /Soma -N / say-3s.perf.
'Then Soma said' (ŚB 1.6.3.21)

(40) a. Then will you listen to me?
b. Then you will listen to me?

(41) a. When will you listen to me?
b. Will you listen to me this time?

5. On the structural organization of Vedic COMP-realizations

To summarize, Vedic COMP-Realizations divide into three syntactic groups in terms of constituent order. The quotative marker *íti*, the only realization of syntactic linking, is a clause-final element. The 'landing-site' of left-dislocation is clearly pre-clausal; the 'nexus' words that serve as discourse linkers are variously pre-clausal or clause-initial, presumably due to their etymological status as

pronominals. The landing-sites of topicalization and pronominal fronting, on the other hand, are clearly clause-initial.

It is in relation to this last fact that Hale's (1987) attempt to shoehorn Vedic Sanskrit into the general assumptions of PPA are most clearly at fault. Hale first fails to recognize that in Vedic 'wh-fronting' is not restricted to relative and interrogative pronominals as it is in Standard Average European languages, but that demonstratives based on the stem *ta-* are subject to the same effects. He further assumes without question that the landing-site of 'wh-fronting' in Vedic as in Standard Average European languages is COMP.[19] He then defines the landing-site of topicalization as a distinct structural position TOPIC which he assumes is to the left of COMP, consistently with the above-mentioned position advocated in Chomsky 1977. This assumption is also consistent with the fact that, in the Vedic corpus, whenever a clause shows both pronominal fronting and topicalization, the topicalized word USUALLY precedes any and all fronted pronominals. This is, however, not always true. Especially if the topicalized word is a verb of speech, as discussed in Hock 1982a, it is not unheard of for it to follow a fronted pronominal, even though it precedes all non-fronted elements and is clearly included in the clause-initial string. Cf. (42).

(42) etat$_i$ ha vai uvāca$_j$ śaṅkaḥ kauṣyaḥ ə$_i$
DEM-A / PTCL / PTCL / say-3s.perf. / Śaṅka-N / Kauṣya-N / ə
putram ə$_j$
son-A / ə
'On this (subject), Śaṅka Kauṣya said to his son ...' (KS 22.6)

Hale argues that his analysis is supported by the facts of clitic-pronominal ordering. He claims that clitic pronouns, when they are fronted into the clause-initial string, cliticize specifically to COMP, and that his analysis, wherein wh-fronting goes to COMP while topicalization lands in a TOPIC position to the left of COMP, is supported by the fact that in the RV clitic pronouns never come between the topicalized word and a fronted 'wh-word', i.e. strings like (43b) are unattested.

(43) a. brahmā́ káḥ vaḥ saparyati
priest-N / INT-N / 2p.cl.-A / honor-3s.pres.
'Which priest honors you?' (RV 8.7.20c)
b. *brahmā́ vaḥ káḥ saparyati
priest-N / 2p.cl.-A / INT-N / honor-3s.pres.

As far as I know, this claim is true of the 'Family Books' of the RV (*maṇḍalāḥ* 2-7), but it is not true of either the younger portions of the RV (cf. (44)) nor of Vedic Prose (cf. (45); presumably, Hale overlooked examples such as these because they involve demonstratives rather than relative or interrogative pronominals.[20]) In (44a) a (topicalized?) subject is followed by a relative pronominal which in turn is followed, not preceded as Hale's theory would predict, by a sentential particle. In (44b) a pronominal clitic comes between a

clause-initial (topicalized?) subject and a relative pronominal in third position. In (45a-b), *etát* is not associated with any nominal and must therefore be read as a sentential adverbial; it is therefore indubitably to be included in the clause-initial string, in which it is clearly preceded by the clitics *asyai* and *enam* respectively. In (45c), it is associated with the noun *āpyā́yanam* 'draught' from which it has been fronted into the clause-initial string, in which it is preceded by the clitic *asmai*. Such examples are enough to falsify Hale's analysis.

(44) a. víṣṇuḥ yát ha ā́vat vṛ́ṣaṇam madacyútam
Viṣṇu-N / REL / PTCL / assist-3s.impf. / bull-A / intoxicated-A
'As Viṣṇu assisted the intoxicated bull' (RV 1.85.7c)

b. pūṣā́ naḥ yáthā védasām ásat vṛdhé
Pūṣan-N / 1p.cl.-D / REL / goods-G / be-3p.subj. / increase-D
rakṣitā́ pāyúḥ ádabdhaḥ svastáye
protector-N / guardian-N / indeceivable-N / well-being-D
'So that Pūṣan, the indeceivable protector and guardian for (our) well-being, may be for us increase of goods' (RV 1.89.5cd)

(45) a. krūrám iva vaí asyai etát karoti
cruel-A / PTCL / PTCL / 3s.f.cl.-D / DEM / do-3s.pres.
'He does a cruel (thing), so to speak, to her.' (TS 5.1.5.1)

b. tát svéna evá enam etát páyasā devā́ḥ svá́
NEXUS / own-I / PTCL / 3s.cl.-A / DEM / juice-I / gods-N / own
akurvata
make-3p.impf.mid.
'So the gods made own = gained possession of it by means of its own juice.' (ŚB 1.5.3.5)

c. vṛtrám hí asmai etát$_i$ jaghnúṣe ə$_i$
Vṛtra-A / PTCL / 3s.cl.-D / DEM-An. / slay-perf.part.-D / ə
āpyā́yanam ákurvan
envigorating draught-An. / make-3p.impf.
'For him, (the one) having slain Vṛtra, they prepared this envigorating draught.' (ŚB 1.6.4.12)

As Hock (1989) has made clear, the fact that clauses like (43b) and (45) are not attested in the RV is due not to a structural distinction between the landing-sites of topicalization and 'wh-fronting' and a constraint on the landing-site of clitic-fronting but on a slight distinction between Mantra Vedic and Vedic Prose in the details of the ordering of constituents within the clause-initial string. Whereas in Vedic Prose, constituents in the clause-initial string stack up according to the five-place template described in Hock (1982a) and outlined in (46a), with all accented fronted words except one going into the final position of the clause-initial string, in Mantra Vedic a variant template operates (46b), in which accented fronted pronominals that do not go into first position go into the third position which in Vedic Prose is reserved for accented particles. For details and further discussion cf. Schäufele 1990b:171-179, 1993b, and Hock 1992:60-71.

(46) a. Ẃ/Pró PTCL PTĆL Pro Pró (Vedic Prose)
b. Ẃ/Pró PTCL Pró/PTĆL Pro (Mantra Vedic)

Hale says nothing at all about Left-Dislocation, and although he notes that a 'clause-connector' (=nexus) will precede a fronted pronominal, Hale fails to acknowledge that it will also precede a topicalized word as in (47). In Schäufele 1986 I argued from this circumstance that, contrary to Hale's and Chomsky's assumption, TOPIC in Vedic must follow, not precede COMP, identifying COMP as the location for nexus. As I have been saying throughout this paper, this is an oversimplification.

(47) a. átho manasā vaí prajā́patiḥ yajñám
NEXUS / mind-I / PTCL / Prajāpati-N / sacrifice-A
atanuta
perform-3s.impf.
'Then Prajāpati performed the sacrifice with his mind.'
(TS 1.6.8.4)
b. tát svéna evá enam etát páyasā devā́ḥ
NEXUS / own-I / PTCL / 3s.cl.-A / thus / juice-I / gods-N
svī́-akurvata
acquire-3p.impf.mid.
'So the gods acquired it by means of its own juice.'
(ŚB 1.5.3.5)

As noted earlier, standard PPA, especially with the assumption that adjunction sites are only for maximal projections plus the 'binary-branching' assumption according to which any non-terminal node has at most two daughters, seems to provide an excessive number of landing-sites for maximal-projections and/or an insufficient number of Bar-Ø landing-sites. What is needed for an adequate description of Vedic is:

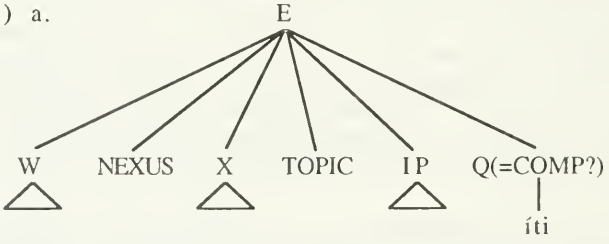
i. A Bar-Ø landing-site to the left of the 'clause proper' (to be defined somehow; cf. below) for topicalized words and fronted pronominals;
ii. A landing-site to the left of the 'clause-proper' for topicalized PHRASES as in (19);
iii. A DISTINCT 'landing-site' to the left of the 'clause proper' for Left-Dislocation (cf. notes 8, 15);
iv. A (presumably Bar-Ø) site for the generation of nexus words, which may be either clause-initial or pre-clausal;
v. A site to the right of the clause proper for the generation of the quotative marker *íti*. This site is likewise presumably Bar-Ø.

It must be further noted that none of these elements can plausibly be said to be subcategorized for by anything or, in other words, to be part of any other constituent's thematic structure. It may be plausible to suppose that nexus words and the quotative marker are base-generated (although a case can be made for the insertion of *íti*,

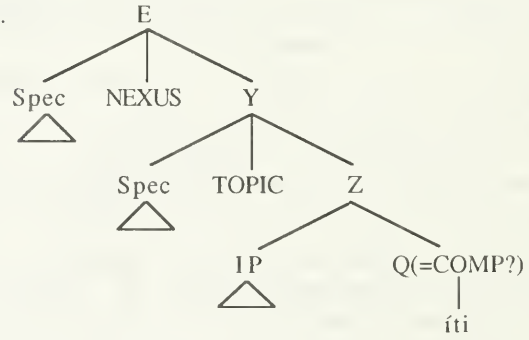
rather in the manner of a case-marker, by the grammar in cases like (35)), but in the absence of such words there may be no reason a priori to expect the base-generation of nodes corresponding to them, and in any case the landing-sites of topicalization and Left-Dislocation are clearly not mandated by θ-Theory.[21] Yet Bar-Ø nodes are manifestly necessary in order to describe these phenomena. So I will for the moment assume that such nodes are base-generated by the grammar of Vedic, probably as functional heads (cf. Gelderen 1993 and references cited there), at least as an option.[22]

There are at least two options for the organization of these various nodes vis-à-vis each other and IP, represented in (48). (48a) represents a relatively flat structure, in which the various clause-peripheral sites are all sisters to each other and to IP. In this diagram, W is the landing-site for Left-Dislocation while X is the landing-site for phrasal topicalization and TOPIC for Bar-Ø topicalization. (The relative ordering of W and NEXUS is unclear and I don't know of anything in the data that would clarify it.) E, as in some PPA literature, represents a supra-clausal 'expression'. In (48b), the nodes labelled 'NEXUS' and 'TOPIC' have associated 'Specifier' nodes, representing the landing-sites for Left-Dislocation and phrasal topicalization respectively. The status of the nodes Y and Z is problematical and will be discussed below.

(48) a.



b.



Either hypothesis needs to account for pronominal fronting as well as topicalization. In Schäufele 1990b I suggested that one single non-pronominal word could be fronted into the TOPIC° position, but that any number of pronominals could be adjoined to this position. It may be objected that, if I am going to posit Bar-Ø-adjunction as a

feature of Vedic grammar, why can't I assume such adjunction to, e.g., IP to account for topicalization instead of positing extra functional heads for which there is little or no evidence in Standard Average European languages? This suggestion runs into two problems, however. One is that, as noted earlier, adjunction of X° constituents to X^{max} nodes seems to violate the Structure-Preservation Constraint.[23] The other is that adjunction is supposed to be capable of unlimited iteration; if one constituent can be adjoined, why not two or more? But in Vedic unlimited fronting is characteristic peculiarly of pronominals, which is why I have suggested that specifically pronominal-fronting is adjunction to the node that serves as landing-site for lexical topicalization. This hypothesis, that (Bar-Ø) topicalization involves movement to a Bar-Ø node while pronominal-fronting involves adjunction to the same node, accounts for:

i. The fact that, while an unlimited number of pronominals can be fronted, at most one non-pronominal can be;
ii. The fact that both the topicalized word and the fronted pronominals are Bar-Ø constituents;
iii. The pragmatic parallel between topicalization and pronominal-fronting.

Of the two hypotheses diagrammed in (48), (48b) has the advantage of providing a structural distinction between 'clause-initial' topicalization and pronominal-fronting on the one hand and 'pre-clausal' Left-Dislocation and nexus on the other. The facts of Vedic syntax require us to distinguish between the clause-initial string, the 'clause proper' i.e., everything within the clause that is not included in the clause-initial string, and the aggregate of the two, which we might call the 'maximal clause'. In the diagram in (48b), IP is to be identified as the 'clause proper'; all of the 'clause-initial string' discussed in Hock 1982a is ultimately contained under TOPIC or its associated Specifier position, and the 'maximal clause' is to be identified with Y. The placement of sentential particles, in terms of which the clause-initial/pre-clausal distinction was defined, is to be defined in terms not of IP, the 'clause proper', but of Y, the 'maximal clause'. Of course, if there has been no topicalization or pronominal-fronting as in (49) then the TOPIC node will be empty and Y will reduce to Z or IP. Meanwhile, Left-Dislocation would place constituents to the left of Y, which would account for such constituents' being treated as outside Y (the 'maximal clause') for purposes of particle-placement. Nexus words likewise would be generated outside of Y, as required, though (being pronominal) they could also appear under TOPIC, like fronted pronominals, hence serving as hosts for particle-placement as in (39). (48a) does not account for this distinction between topicalization and Left-Dislocation, between clause-initial and pre-clausal sites, which is why I have been inclined to favour (48b).

(49) [Y[S[NP janakáḥ ha vaídehaḥ] [VP bahudakṣiṇéna
Janaka-N / PTCL / Vaideha-N / lavish-I
yajñéna ĩje]]]
sacrifice-I / worship-3s.perf.
'Janaka Vaideha worshipped with = performed a lavish sacrifice' (ŚB 14.6.6.1)

The weakness of the analysis in (48b) is that it suggests that *íti*, TOPIC, and NEXUS are 'heads' of Z, Y, and E respectively. This suggestion has in my opinion little if any justification. In English, COMP words like *for* in (50) have the power to assign Case and otherwise license the generation of structures and constituents; in other words, they behave like lexical heads. To the extent that *íti* 'licenses' the generation of constituents within larger constituents, it too behaves like a head, in which case the node labelled 'Z' in (48b) may actually correspond to CP or some such standard phrasal constituent in PPA. But note that the 'complement' of *íti* isn't necessarily a clause; it may be, as in (6d) or (35), a single word. This fact casts in doubt the precise status of Z.

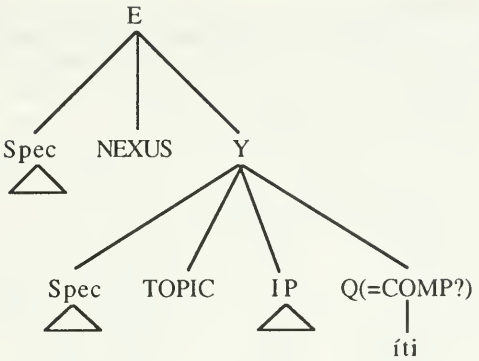
(50) We were waiting for him to bring the tapes.

As for Y and E, there is no recognizable sense in which the nodes labelled TOPIC and NEXUS function as their respective heads. In other words, saying TOPIC is the 'head' of Y or that Y is the maximal projection of TOPIC is to say nothing at all; and likewise for the relationship between NEXUS and E. Furthermore, it is not clear from the data what the relationship is between TOPIC to the left of IP and *íti* to its right. (48b) implies an asymmetrical aunt-niece relationship (cf. also (25) and accompanying discussion), but they might as well be sisters, as in (48a). Standard PPA would strongly prefer the analysis in (48b) because it conforms better to the assumptions underlying the 'Universal Base Hypothesis', outlined in Travis 1984 and Speas 1990, according to which all lexical or Bar-Ø categories share the same projection structure. But standard PPA would necessarily read into this analysis implications about headship, etc. that are unwarranted. For the reason given above, I prefer the analysis in (48b) over that in (48a), but in doing so I warn stringently against the implications it would seem to bear within the context of $\bar{X}$-Theory; an adaptation of the analysis in (48b) such as that in (51) might actually be preferable, as far as the data are concerned. This is why I have used the non-committal labels 'E', 'Y', and 'Z', and have striven to avoid the suggestion that NEXUS, TOPIC, or even *íti* is a 'head' in any grammatically meaningful sense.

6. Conclusion

In this paper I have argued that the functions which the Standard-Theoretical school has associated with the structural node-label COMP and its projections were in Vedic Sanskrit associated with a variety of structural positions which cannot plausibly be viewed as

(51)



belonging to the projection of any one node. This is especially true given the necessity of two Bar-Ø nodes at opposite ends of the clause, the location of the quotative-marker *íti* and the landing-site of topicalization and pronominal-fronting, each of which serves some of the COMP-Functions, and the implausibility of regarding some of the Bar-Ø nodes discussed in this paper as 'heads' in the X-theoretical sense.

For the reasons reported in this paper, I am inclined to reject the notion that UG defines a single Bar-Ø node and its set of projections as the syntactic realization of the 'COMP-Functions' listed in (1), as is typical of the Standard Average European languages. Instead, I suggest that there is a (presumably universal) set of 'COMP-Functions', and that UG allows for a range of possible syntactic structures realizing them. It seems to me likely that all of these options involve clause-peripheral positions, though even this hypothesis may prove too narrow. This is an empirical question.

Of course, since the syntactic and pragmatic functions discussed in this paper are called 'COMP-Functions' because of their theoretical association with a specific node-label 'COMP', if we reject that node-label with its implications as a linguistic universal we should probably find another name for the collective functions it tends to serve. However, if broad-based and thorough comparative and typological studies indicate that there is a strong cross-linguistic tendency for these functions to be 'gathered together' into the purview of a single Bar-Ø node and its projections, then we would be justified not only in calling that node 'COMP' but in using that label for the set of syntactic and pragmatic functions associated with it as well, and languages like Sanskrit would serve syntactic theory primarily as exotic reminders that this association is not, as standard PPA assumes, inherent in UG.

NOTES

* An earlier version of this paper was presented at the Thirteenth South Asian Languages Roundtable, 26 May 1991, Urbana,

IL. I would like to acknowledge the helpful comments of Rakesh Bhatt, Hans Henrich Hock, Gillian Ramchand, Nalini Rau-Murthy, and James Yoon, and to absolve them of any responsibility for this paper. In the examples, all inter-word sandhi has been undone, and accent is indicated if it is given in the primary source.

The examples in this paper make use of the following abbreviations:

GRAMMATICAL ABBREVIATIONS

A	accusative case	cl.	clitic pronoun
Ab.	ablative case	d.	dual
Adp.	adposition	f.	feminine gender
Adp(X)	adposition assigning case X	fut.	future tense
		ger.	gerundive
Adv.	adverb	impf.	imperfect tense
D	dative case	impv.	imperative mood
DEM	demonstrative	m.	masculine gender
G	genitive case	mid.	middle voice
I	instrumental case	n.	neuter gender
INT	interrogative	opt.	optative mood
L	locative case	p.	plural
N	nominative case	part.	participle
O	oblique case	pass.	passive voice
Pfx.	verbal prefix	perf.	perfect tense
PTCL	particle	pres.	present tense
Q	quotative marker	pro.	personal pronoun
REL	relative pronominal	s.	singular
V	vocative case	subj.	subjunctive mood

TEXTUAL ABBREVIATIONS

AB: Aitareya Brāhmaṇa
AV: Atharva-Veda
JB: Jaiminīya Brāhmaṇa
JUB: Jaiminīya Upaniṣad-Brāhmaṇa
KS: Kāṭhaka Saṁhitā
RV: Ṛg-Veda
ṢB: Ṣaḍviṁśati Brāhmaṇa
ŚB: Satapatha Brāhmaṇa
TB: Taittirīya Brāhmaṇa
TS: Taittirīya Saṁhitā

[1] Here, as elsewhere, references to LINGUIST list contributions are cited in this format.

[2] One reviewer challenged the claim that COMP is generally regarded in the PPA community as a linguistic universal. While works independent of mine (e.g. Bhatt & Yoon 1991) have challenged this assumption, in practice the assumption is strong enough to be introduced immediately into any discussion of clause-peripheral phenomena in any language; cf., for the issues relevant to this paper,

especially Hale (1987, 1993). I am here recommending that, in any language outside Standard Average European, the existence of COMP must first be justified before it is used in any further syntactic discussion.

[3] That the fronting of non-pronominals, in this paper for convenience labelled 'topicalization', can have a variety of communicative functions in addition to topicalization properly so called (for which (4b-c) may serve as good examples) can be seen briefly from the following examples. It may involve (contrastive) focus as in (i) = (4a); there are, of course, alternative strategies for expressing such focus; cf. (ii-iii). Such strategies include not only contrastive accent (in speech) or left-dislocation, represented in (ii-iii), but, at least in verb-final languages, focus-movement; for discussion cf. Schäufele 1990a and references cited there. 'Topicalization' may also serve a stage-setting function as in (iv-v). It can also serve a 'backgrounding' function, deemphasizing a constituent or at least denying it an emphasis it might get in a more 'normal' position. Cf. (vi) vs. (vii).

(i) Peas I like.
(ii) I like **peas**.
(iii) As for peas, I like them.
(iv) In the morning, Terry left the package at the corner.
(v) In the jungle, the mighty jungle, the lion sleeps tonight.
(vi) With the claw (of the hammer), Brent pried out several nails. (most felicitous if the nails are the topic of subsequent narrative/discussion)
(vii) Brent pried out several nails with the claw of the hammer. (most felicitous if the subsequent topic is either the hammer or the claw)

Cf. Schäufele 1991, 1993a for a discussion of the variety of pragmatic consequences 'topicalization' seems to have had in Vedic.

Of course, pronominal-fronting also entails certain pragmatic consequences, often parallel to those of 'topicalization'. As will be further discussed in this paper, this is one reason why I judge it desirable to define both kinds of fronting as having the same, or at least related, landing-site(s).

[4] 'Left-Dislocation' is known in more traditional grammars of Indo-European languages by the Greek name 'prolepsis'.

[5] Cf. also Bhatt & Yoon 1991 for a more general discussion of the necessity of distributing COMP-Functions amongst distinct structural positions in a variety of languages, and Gelderen 1993 for the inappropriateness of certain functional categories, allowed and defined by UG, in the description of certain languages.

[6] In (16), I have represented the 'traces' of clitic pronouns that represent verbal arguments as being 'cliticized' to the verbs under the lexical node V, since if these pronouns had not been fronted they would be in these positions, forming phonological words with the

verbs. Cf. Schäufele 1993b and references cited there for further discussion.

[7] Vedic topicalization differs from pronominal-fronting most obviously in that while an unlimited number of pronouns can be fronted in Vedic (cf. above, (8b)), at most one non-pronominal can be topicalized. The ramifications of this distinction will be discussed below.

[8] As noted in Schäufele 1991, 1993a, the general Vedic preference for X°-fronting over the fronting of multi-word phrases is reflected to a great extent in text frequency; the overwhelming majority of fronted constituents in the entire corpus are single words. But it is also reflected in genre differences: In the metrical texts of the Mantra Vedic corpora, although the vast majority of fronted constituents are single words, the fronting of whole phrases is not uncommon. However, in the pedantic style of the Vedic Prose texts the (unambiguous) fronting of whole phrases is of vanishingly small frequency.

[9] As noted in Schäufele 1991, 1993a, Webelhuth & den Besten (1987) have suggested that some apparently Bar-Ø topicalized constituents in, e.g., Germanic are actually X^{max} constituents. In those papers I noted that, given that in Vedic any (non-clitic) word can be topicalized (including verbal prefixes and members of conjuncts, cf. (20c,f)), the hypothesis proposed by Webelhuth & den Besten as a 'safety valve' for the theoretical assumption that only maximal projections can be topicalized would require that in Vedic every word be treated as a maximal projection. Given the evidence for hierarchical structure adduced in Schäufele 1990b, this hypothesis seems not only ad hoc but absurd. Cf. Schäufele 1991, 1993a for further discussion.

[10] Even if, as one reviewer pointed out, Left-Dislocation is assumed to be a base-generated structure (a very plausible assumption, even 'pre-/extratheoretically'), the material to the left of the clause must have a structural location. For convenience, I will in this paper refer to this location as a 'landing-site.'

[11] The assumption that TOPIC, if it exists, is to the left of COMP has not gone unchallenged in PPA. Baltin (1982:17-22) argues on the basis of sentences like the one below that TOPIC must be to the RIGHT of COMP. Huang (1982:152-53) has also argued that in Chinese TOPIC must be to the right of COMP. Cf. Lasnik & Saito (1992:cap. 4) for further discussion.

It's obvious [[that]COMP [Mary$_i$ he can't stand ə$_i$]S]$\bar{S}$

[12] Hock (1982b:45-47) notes that occasions in which *íti* precedes the quoted material, while they can be found in the RV, are 'definitely ... in the minority compared to those in which *íti* follows' the quoted material. He tallies five occurrences of pre-quote *íti* in the RV, as opposed to 32 occurrences of post-quote *íti*; however, the only

example he gives in his text is one in which the order is *íti*—matrix verb—matrix subject—quote; cf. below. Thus, it is not obvious that *íti* in this example bears any syntactic relation to the quoted material; as Hock himself notes, *íti* in this example could at least equally plausibly be glossed by the pronominal 'thus'.

> íti bravīti vaktárī rárāṇaḥ vásoḥ vasutvā́
> Q / say-3s.pres. / speaker-N / generous-N / good-G / goodness-I
> kārávaḥ / anehā́ḥ
> singers-N / guiltless-Np.
> '(Thus) says the generous speaker, "Through the goodness of the good, the singers are guiltless."' (RV 10.61.12bc)

[13] One reviewer suggested the following alternatives to the structure in (26):

(i) Y [XP Topic [$\bar{X}$ PTCL IP]]
in which Y is the landing-site for Left-Dislocation, and extra 'topic' positions can be generated by adjunction to IP.

(ii) [TopP Spec [$\bar{T}$ Topic [CP Spec [$\bar{C}$ IP [C íti]]]]]

Suggestion (i) misses the point that any 'topic' nodes in Vedic ought preferentially to be Bar-Ø, not Spec nodes or adjunction to other X^{max} nodes. Suggestion (ii) avoids this problem by defining a Bar-Ø Topic node as a functional head. However, assuming that this node is the landing-site for topicalization and/or pronominal fronting, as suggested later in this paper, while its Spec position is the landing-site for Left-Dislocation, it fails to account for the fact that the Topic node must be clause-internal, as discussed immediately below, while its Spec node is extra-clausal.

[14] Cf. Klein 1978, Hock 1982a, and Schäufele 1985, 1986, 1988, and 1990b, and, for more general discussion of such elements, Kendall & Yoon 1986, Schourup 1983, Zwicky 1985, and Underhill 1988.

[15] It should be noted that Vedic also has a small number of particles such as *evá* and *iva* that are phrase-bound rather than clause-bound; cf. especially the discussion in Schäufele 1985, 1990b on this subject. These particles are also glossed 'PTCL' in the examples in this paper, but the distinction between them and the more important clausal particles should in any given case be either clear or irrelevant.

[16] The word immediately preceding the first *ha, ca* '&', itself belongs to the particle class and therefore the two together form a particle string *ca ha* identifying the preceding word *datvaḥ* as the first word in the 'clause'.

[17] As is clear from this discussion, topicalization and Left-Dislocation are distinct in Vedic: whereas Left-Dislocation not only involves resumptive pronouns (cf. Oertel 1926) but is clearly pre-clausal, witness the evidence of distinct sentential-particle strings,

topicalization is clearly clause-initial by the contrary evidence. Pragmatically, the two phenomena are probably in complementary distribution; I have not come across any sentences in the Vedic corpus involving both. Be it noted that, while topicalization is often used to highlight or emphasize a constituent, Vedic grammar had other means of accomplishing this end, such as use of the emphatic phrasal particle *evá* or focussing, as discussed in Schäufele 1990a.

[18] One reviewer asked if it is possible for particles ever to be clause-initial in Vedic. Generally, it is not; while not all Vedic discourse particles are clitics, they all behave like clitics in requiring preceding hosts; cf. Schäufele 1993b for further discussion. Some particles, e.g., *nú* and *evá*, have in the RV 'long-vowel' doublets (i.e., *nā́* and *evā́*) which act as full adverbial clause-connectors; cf. Schäufele 1986 for further discussion. But these are not relevant to this paper.

[19] Actually, of course, as noted earlier, modern PPA defines the landing-site of wh-movement as 'SPEC of COMP', rather than COMP°. This distinction is, however, irrelevant for Hale's discussion.

[20] Overwhelmingly in Vedic Prose relative and interrogative pronominals go into clause-initial position, so one rarely has occasion to find such a pronominal preceded by anything at all in a clause in Vedic Prose. It is not impossible, however; witness the following example:

índraḥ ha yátra vṛtrā́ya vájram pra-jahā́ra
Indra-N / PTCL / REL / Vṛtra-D / thunderbolt-A / throw-3s.perf.
'When Indra had thrown (his) thunderbolt at Vṛtra' (ŚB 1.2.4.1)

[21] As noted in Schäufele 1991, 1993a, topicalization in Vedic, whatever its peculiarities, is an orthodox example of Move α in that it is movement to an Ā-position.

[22] I have no reason to suppose that such nodes have been generated in sentences in which they are not occupied

[23] Note that the Bar-Ø adjunction posited in Schäufele 1990a to account for focus-movement in Vedic, as well as the 'incorporation' processes discussed in Baker 1988, all make use of Bar-Ø hosts.

REFERENCES

BAKER, Mark C. 1988. Incorporation: a theory of grammatical function changing. Chicago: University of Chicago Press.

BALTIN, Mark R. 1982. A landing site theory of movement rules. Linguistic Inquiry 13.1-38.

BHATT, Rakesh, & James Yoon. 1991. On the composition of COMP and parameters of V2. Proceedings of the Tenth West Coast Conference on Formal Linguistics, ed. by Dawn Dates.

BRESNAN, Joan. 1972. Theory of complementation in English syntax. MIT Ph. D. dissertation in Linguistics.

CHOMSKY, Noam. 1977. On *wh*-movement. Formal syntax, ed. by Peter W. Culicover, Thomas Wasow, & Adrian Akmajian, 71-132. New York: Academic Press.

-----. 1986. Barriers. Cambridge, MA: MIT Press.

DELBRÜCK, Berthold. 1888. Altindische Syntax. (Syntaktische Forschungen, 5.) Halle: Buchhandlung des Waisenhauses.

EMONDS, Joseph E. 1976. A transformational approach to English syntax: root, structure-preserving, and local transformations. New York: Academic Press.

GELDEREN, Elly van. 1993. The rise of functional categories. Amsterdam: Benjamins.

HALE, Mark Robert. 1987. Studies in the comparative syntax of the oldest Indo-Iranian languages. Harvard Ph. D. dissertation in Linguistics.

-----. 1993. Deriving Wackernagel's Law: prosodic and syntactic factors determining clitic placement in the language of the Rigveda. Paper read at the 2nd-Position Clitic Workshop, The Ohio State University, Columbus, Ohio, 11 July 1993. To appear in the Proceedings.

HOCK, Hans Henrich. 1982a. Clitic verbs in PIE or discourse-based verb fronting? Sanskrit *sá hovāca gā́rgyaḥ* and congeners in Avestan and Homeric Greek. Studies in the Linguistic Sciences 12:2.1-38.

-----. 1982b. The Sanskrit quotative: a historical and comparative study. Studies in the Linguistic Sciences 12:2.39-85.

-----. 1989. Conjoined we stand: theoretical implications of Sanskrit relative structures. Studies in the Linguistic Sciences 19:1.93-126.

----- (ed.). 1991. Studies in Sanskrit syntax. Delhi: Motilal Banarsidass.

-----. 1992. What's a nice word like you doing in a place like this? Syntax vs. Phonological Form. Studies in the Linguistic Sciences 22:1.39-87.

HUANG, Cheng-Teh James. 1982. Logical relations in Chinese and the theory of grammar. MIT Ph. D. dissertation in Linguistics.

KENDALL, Sue Ann, & James H. Yoon. 1986. Sentence particles as evidence for morphosyntactic interactions with pragmatics. Studies in the Linguistic Sciences 16:1.53-77.

KLEIN, Jared S. 1978. The particle *u* in the Rigveda: a synchronic and diachronic study. Göttingen: Vandenhoeck & Ruprecht.

-----. 1987. Syntactic and discourse correlates of verb-initial sentences in the Rigveda. Paper read at the Second Sanskrit Syntax Symposium, Ninth South Asian Languages Analysis Roundtable, Syracuse, NY, 7 June 1987. Hock 1991.123-143.

LASNIK, Howard, & Mamoru Saito. 1992. Move α. (Current Studies in Linguistics, 22.) Cambridge, MA: MIT Press.

MCDANIEL, Dana. 1989. Partial and multiple wh-movement. Natural Language and Linguistic Theory 7.565-604.

MINARD, Armand. 1936. La subordination dans la prose védique. (Études sur le Śatapatha-Brāhmaṇa, 1. Annales de l'Université de Lyon, 2:3.) Paris: Belles Lettres.

OERTEL, Hanns. 1926. The syntax of cases in the narrative and descriptive prose of the Brāhmaṇas. I. The disjunct use of cases. (Indogermanische Bibliothek. 1. Abt. Sammlung indogermanischer Lehr- und Handbücher. 1. Reihe. Grammatiken, 18.) Heidelberg: Carl Winter's Universitätsbuchhandlung.

ROSENBAUM, Peter S. 1967. The grammar of English predicate complement constructions. Cambridge, MA: MIT Press.

RUDIN, Catherine. 1988. On multiple question and multiple wh fronting. Natural Language and Linguistic Theory 6.445-501.

SCHÄUFELE, Steven. 1985. Fronting/topicalization and sentential particles in Vedic prose. Department of Linguistics, University of Illinois at Urbana-Champaign, MS.

-----. 1986. Single-word topicalization in Vedic Prose: a challenge to Government & Binding? Paper read at the First Sanskrit Syntax Symposium, Eighth South Asian Languages Analysis Roundtable, University of Illinois, Urbana, IL, 30 May 1986. Hock 1991.153-175.

-----. 1988. Where's my NP? Non-transformational analyses of Vedic pronominal fronting. Studies in the Linguistic Sciences 18:2.129-162.

-----. 1990a. A 'focus' position for subjects within the Vedic VP. Paper read at the Twelfth Annual South Asian Languages Analysis Roundtable, University of California, Berkeley, CA, 3 June 1990.

-----. 1990b. Free word-order syntax: the challenge from Vedic Sanskrit to contemporary formal syntactic theory. University of Illinois at Urbana-Champaign Ph.D. dissertation in Linguistics. Available from University Microfilms.

-----. 1991. Lexical topicalization in Vedic Sanskrit and the Head Movement Constraint. Paper read at the 44th Kentucky Foreign Languages Conference, Lexington, KY, 26 April 1991.

-----. 1993a. X°-fronting in Vedic Sanskrit and Binding Theory. Paper read at the 15th Annual South Asian Languages Analysis Roundtable, Iowa City, IA, 22 May 1993.

-----. 1993b. Now that we're all here, where do we sit? Phonological ordering in the Vedic clause-initial string. Paper read at the 2nd-Position Clitic Workshop, The Ohio State University, Columbus, Ohio, 11 July 1993. To appear in the Proceedings.

SCHOURUP, Laurence C. 1983. Common discourse particles in English conversation. Ohio State University Working Papers in Linguistics, 28. Columbus: Ohio State University.

SPEAS, Margaret J. 1990. Phrase structure in natural language. Studies in natural language and linguistic theory 21. Dordrecht: Kluwer.

TRAVIS, Lise DeMena. 1984. Parameters and effects of word order variation. MIT Ph. D. dissertation in Linguistics.

UNDERHILL, Robert. 1988. *Like* is, like, focus. American Speech 63. 234-246.

WEBELHUTH, Gert, & Hans den Besten. 1987. Adjunction and remnant topicalization in the Germanic SOV languages. Paper presented at the GLOW Conference, 30 March-2 April 1987, Venice.

ZWICKY, Arnold M. 1985. Clitics and particles. Language 61.283-305.

Studies in the Linguistic Sciences
Volume 23, Number 1, Spring 1993

SQUIB

Discourse markers and sentential syntax*

M. Lynne Murphy

The notion Discourse Marker or Discourse Particle (henceforth, DM) has, of late, claimed the attention of many linguists, who have tried to define the meanings or uses of these lexical items. Relatively little has been said, however, about the syntactic status of such pieces of language as *oh*, *like*, *y'know*, and *well*, though several syntacticians (e.g. Ross (1973), Emonds (1976), McCawley (1982), Espinal (1991)) have focused their concern on the Discourse Marker's wordier cousin, the parenthetical remark. Though Deborah James noted many of the syntactic peculiarities of DMs in her 1972 and 1973 CLS papers, her data and provocative arguments for syntactic treatment of DMs have not received much attention in the current flurry of literature on DMs and parentheticals. The aim of this writing is to examine and refute James's argument that DMs are syntactically sensitive to the sentences they inhabit. Instead, I maintain that semantic well-formedness, rather than syntactic well-formedness, determines where a DM can occur. Thus, DMs are best treated as independent of the syntactic structure of the sentence.

Discourse Markers are those lexical items which are used by the speaker to comment upon the discourse plan and goals. This definition, from Andrea Dunn's 1990 dissertation on Swahili DMs, covers a large assortment of lexical items in English, which do not otherwise fall into traditional parts of speech, such as *oh*, *ah*, *uh*, certain uses of *well*, *say*, *y'know*, *like*, and non-conjunctive uses of *so*, *and*, *but*, among others. Just a few examples of DMs are given in (1)-(10).

(1) **Oh**, I wish I were an Oscar Mayer wiener.
(2) **Ah**, I see you've found the wine cellar.
(3) I, **like**, couldn't believe that Elvis was dead.
(4) **Well**, I always meant to buy flood insurance.
(5) If we had, **say**, a jelly roll, we'd be having a lot more fun.
(6) **Now**, there comes a time in every linguist's life when she must read Chomsky.
(7) I never meant to electrocute my brother's goldfish, **y'know**.
(8) **But** you said I could have a cookie!
(9) **And** how am I supposed to pay for this?
(10) **So** you think I'm made of money?

DMs have a number of characteristics in common which are symptomatic of their discourse-comment purpose. First, they do not generally contribute to the truth-conditional semantics of their host sentences. This can be seen in comparing the DM *y'know* with the

subject-verb sequence *you know*. In (11), where *know* is the matrix verb and *you* its subject, the sentence is true if and only if my mother's maiden name is among the things that you know. On the other hand, sentence (12) with DM use of *y'know* (12) is truth conditionally equivalent to (13), which has no *you know* at all. And in fact, (12) is felicitously used even, or especially, if (11) is false.

(11) You know (that) my mother's maiden name is Bangs.

(12) **Y'know**, my mother's maiden name is Bangs.

(13) My mother's maiden name is Bangs.

Another characteristic of DMs is their freedom in syntactic ordering. Any syntactic position may be preceded or followed by some Discourse Marker. The DM *like* can be used at any point in the sentence, to the frustration of parents of teenagers everywhere. Though it is unlikely that sentence (14a) would be uttered with all of the possible *like*s indicated, with the right intonational patterns, it can support any one, or even several, of the *like*s, as in (14b) or (14c), for example.

(14) a. (Like) it's (like) been (like) forever (like) since (like) I've (like) been (like) to (like) the (like) mall (like).
b. **Like** it's been forever since I've **like** been to the mall.
c. It's been **like** forever since **like** I've been to **like** the mall.

But not every DM is so flexible in its placement. For instance, while *now* and *ah* are both acceptable at the beginning of sentences, as shown in (15), *now*, but not *ah*, is acceptable at the end of a sentence, as in (16).

(15) a. **Now**, the Jolly Green Giant doesn't like lima beans.
b. **Ah**, the Jolly Green Giant doesn't like lima beans!

(16) a. The Jolly Green Giant doesn't like lima beans, **now**.
b. #The Jolly Green Giant doesn't like lima beans, **ah**.

These sorts of facts about the distribution of DMs in host sentences inspired James to try to account for DM distribution within the schemata of sentential syntax. In her 1972 CLS paper, James makes two major claims. First, that DMs have meaning, in that they can be distinguished from one another by what they communicate. And second, that there are syntactic constraints on the distribution of DMs. James starts to catalogue some of these syntactic constraints, such as: DMs cannot break up idiomatic phrases, as in (17), or follow preposed manner adverbials, as in (18).

(17) *John kicked, **oh**, the bucket.

(18) *Wisely, **why**, Jan left early.

In her 1973 CLS paper, James further develops her 1972 claims and asserts that DMs refer to parts of the sentence, and the parts of the sentence to which they refer are necessarily syntactic con-

stituents. James's use of the term "refer" here is not the usual semantic notion of reference. Instead, it is what I have termed here as "comment". So, to rephrase James's claim, DMs comment upon parts of the sentence, and the parts upon which they comment must be syntactic constituents. This can be characterized as a kind of scopal relation — a DM has scope over the constituent it comments upon, although DM scope should not be equated with other types of scopal relations such as quantifier scope. James's claim for the syntactic sensitivity of DM scope is motivated by sentences such as (19) and (20), in which the scopes of the DMs are just those complete constituents that occur immediately to the right of the DM, as indicated by the square brackets.

(19) The woman who said she liked, [DM-scope **oh**, [VP to read]] sang a song.
(20) That Kim is thinking of moving to [DM-scope **ah**, [NP San Francisco]] is considered likely.

Since the scope of *oh* in (19) cannot be *to read sang a song*, which is not a constituent, James concludes that the objects over which DMs have scope must be well-formed syntactic units. However, we can put James's claim to rest with one simple counter-example, stated in (21), where the DM simultaneously comments on more than one syntactic unit in the sentence.

(21) Jackie should drive, [DM-scope **oh**, [PP from Nova Scotia] [PP to Arizona]], just for the experience at the wheel.

In (21), *from Nova Scotia to Arizona*, is not a syntactic constituent, rather it is a series of two prepositional phrases, each of which is a daughter of the larger VP. However, this string is most likely what the speaker intends the *oh* to comment upon, if we read the *oh* as indicating that the content of the string *from Nova Scotia to Arizona* was made up at the spur of the moment as exemplary start and finish points for Jackie's trip. Thus, James's claim that syntactic constituency is relevant to the scopal properties of DMs is falsified, although we still need to explain why the scopes of DMs are so frequently linked to syntactic constituency, as in (19) and (20). If syntactic constituency is irrelevant, then why is it that non-constituents like *to read sang a song* are not granted DM scope?

The answer to this question lies in the fact that well-formed syntactic units reflect well-formed semantic units. That is, syntactic units are meaningful, and only units of meaning are commented upon by Discourse Markers. This jibes with Dunn's definition of DMs as commenting on the speaker's goals or plan in the discourse. It is only by communicating with the audience via meaningful units of language that the speaker can successfully complete her discourse plan, or parts thereof. Thus, DMs comment upon meaningful units of the discourse in order to clarify how the statement of those units contributes to the speaker's goals. Although it is not a syntactic constituent, *from Nova Scotia to Arizona* is a meaningful unit in the

discourse, referring to a path by designating its two extremes. In commenting upon this string, *oh* indicates that the specifics of the path were extemporaneously produced, and therefore it may be the case that those specifics were not terribly relevant to the successfulness of the discourse plan. In other words, the *oh* indicates that the path from Nova Scotia to Arizona is a good one for exemplifying the speaker's point, but that some other path might have worked equally as well.

Having established that DMs are not syntactically sensitive in their scopal properties, we are left with no motivation for attributing syntactic properties to DMs. Recall that DMs defy distributional generalizations, making them an unlikely syntactic category. While nouns, for instance, might be syntactically interchangeable in noun phrases, DMs are not necessarily interchangeable in any particular sentence. Furthermore, DMs play no syntactic roles in the sentence, since they are neither subcategorized nor specified for. Finally, there is little evidence for the position of DMs in the hierarchical structure of the sentence. For instance, in (21), what would be the mother of *oh*? The first PP, in which case *oh from Nova Scotia* would be a constituent? The VP (*drive, oh, from Nova Scotia to Arizona*)? The sentence? Each possibility presents problems for syntactic theory, since none of these proposed constituents passes the traditional tests for constituency: they don't form prosodic units, nor are they moved, copied, or deleted as constituents.

The argument that DMs are not syntactic constituents of their host sentences can be extended to parentheticals, which can serve the same purpose as lexical DMs, making meta-discourse comment, as in (22).

(22) Jackie should drive — [DM-scope **what would be a good route** — [PP from Nova Scotia] [PP to Arizona]], just for the experience at the wheel.

Recognition that neither DMs nor parentheticals are properly within the sentential syntax puts to rest the arguments among Ross (1973), Emonds (1976), and McCawley (1982) as to what the constituency status of parentheticals is. Accounts in three-dimensional syntax (Espinal 1991) or autolexical syntax (as formulated by Sadock 1991) may be more promising, as such accounts allow for multiple syntactic dimensions in which DM/parenthetical and sentence structures could exist independently. It is unclear, however, how the separate syntactic structures for the host sentence and the DM/parenthetical would interact semantically such that the theory can represent the DM's scope over semantic units in the sentence.

To conclude, I have shown here, contra James 1973, that there is no motivation for treating DMs and their host sentences as composite syntactic structures. Instead, it seems that DMs interact with their host sentences with reference to semantic rather than syntactic units. While multidimensional theories of language structure hold promise

for the syntactic representation of these sentences, the semantic interaction of DMs and their hosts deserves further investigation.

NOTES

* I am grateful to many members of the University of Illinois Department of Linguistics for comments and discussion concerning this work, especially Jennifer Cole, Georgia Green, and Jerry Morgan.

REFERENCES

DUNN, Andrea Susan. 1990. The pragmatics of selected discourse markers in Swahili. University of Illinois at Urbana-Champaign Ph.D. dissertation in Linguistics.

EMONDS, Joseph E. 1976. A transformational approach to English syntax: Root, structure-preserving, and local transformations. New York: Academic Press.

ESPINAL, M. Teresa. 1991. The representation of disjunct constituents. Language 67:4.726-762.

JAMES, Deborah. 1972. Some aspects of the syntax and semantics of interjections. CLS 8.162-172.

-----. 1973. Another look at, say, some grammatical constraints, on, oh, interjections and hesitations. CLS 9.242-251.

MCCAWLEY, James D. 1982. Parentheticals and discontinuous constituent structure. MS. University of Chicago (printed in Linguistic Inquiry 13).

ROSS, John R. 1973. Slifting. The formal analysis of natural languages, ed. by M. Gross, M. Halle, & M. Schützenberger. The Hague: Mouton.

SADOCK, Jerrold M. 1991. Autolexical syntax. Chicago: University of Chicago Press.

Studies in the Linguistic Sciences
Volume 23, Number 1, Spring 1993

REVIEW

Anvita Abbi. **Reduplication in South Asian languages: An areal, typological, and historical study**. New Delhi: Allied Publishers, 1992; pp. xxii, 193.

Hans Henrich Hock

South Asia is one of the paradigm cases of linguistic convergence, involving the three major linguistic families Indo-Aryan, Dravidian, and Munda, as well as some of the Iranian and Tibeto-Burman languages, Khasi (Austro-Asiatic like Munda, but belonging to a different subfamily), the language isolate Burushaski, and languages such as Nahali whose genetic affiliation is controversial. The features most prominently mentioned as defining the convergence area are (i) a contrast between dental and retroflex consonants; (ii) SOV order; (iii) the widespread use of absolutives and other non-finite devices instead of, or beside, clausal subordination, and (iv) the use of quotative constructions marked by a post-citation quotative particle. It is generally acknowledged that Indo-Aryan, Dravidian, and Munda form the core of the convergence area.[1] See especially the synchronic work of Masica (1988) and Ramanujan & Masica (1969). The historical developments giving rise to convergence are still a matter of controversy; cf. e.g. Emeneau 1954, 1956, 1962, 1980, Kuiper 1967a, b, 1991, and Kaufman & Thomason 1988 vs. Hock 1975, 1982, 1984 and the yet different interpretation in Tikkanen 1987.

The volume under review attempts to demonstrate that another important feature characterizes the South Asian convergence area, viz. REDUPLICATION.

Abbi (A.) uses the term 'reduplication' in a fairly broad sense to refer to all of the highlighted structures in (1) - (4), although sometimes she distinguishes structures of the type (1) and (2) from those of the type (3) and (4) as 'morphological' vs. 'lexical'.[2] (The examples in (1) and (2) are from Sanskrit, (3) from Hindi, and (4) from Khasi.) Her major emphasis is on the type (3); for it is this type, especially structures like (3cd) with iterated non-finite verbs, which she claims characterizes the core of the South Asian convergence area.

(1) Reduplicated structures from Skt. *tar-* 'cross over'
  a. Reduplicated perfect: **ta-tār**-a
  b. Reduplicated desiderative: **ti-tīr**-ṣa-ti
  c. Reduplicated intensive: **tar-tarī**-ti

(2) a. **adhikādhika-** 'very much, very many'
  adhika-adhika- (from adhika- 'much, many')
  b. **ekaika-** 'each one (in turn)'
  eka-eka- (from eka- 'one')

(3) a. rāvaṇ kī **baṛī baṛī** āṁkheṁ thīṁ
Ravana's/big/big/eyes/be-past
'Ravana had big, big (= very big) eyes.'

b. unko **ek ek** phal de do
they-dat/one/one/fruit/give/give
'Give them one fruit each.'

c. vah pān **becte becte** bolā
[verbal adverb]
he/betel/selling/selling/speak-past
'He spoke while selling betel.'

d. vah **baiṭhe baiṭhe** thak gayā
[verbal adverb]
he/sit-perfective/sit-perfective/tired/go
'He got tired from sitting (for a long time).'

(4) bən **Iyait ši Iyait** ŋgala thait
['interfixed reduplication']
infinitive-marker/walk/'infix'-particle/walk/I-past/tired
'I am tired of walking continuously.'

A. bases her claims on an impressive amount of data, much of them coming from her own fieldwork. Within modern South Asia she examines the thirty-three languages and dialects listed in Table I. For comparison and contrast she adds data from earlier stages of South Asian languages and from related languages outside the subcontinent (cf. Table II), with some additions from 'Hokkien and Chinese structures', some of 'the languages of Indonesia, Thailand, Malaysia, the Philippines, and Singapore' (8), as well as ancient Iranian (9). The total number of languages is 'about 48 to 50' (8).

Indo-Aryan:	1.	Assamese	8.	Kashmiri
	2.	Bengali	9.	Marathi
	3.	Dakhini Urdu	10.	Maithili
	4.	Dogri	11.	Oriya
	5.	Gujarati	12.	Panjabi
	6.	Hindi-Urdu	13.	Sindhi
	7.	Konkani	14.	Sadari
Dravidian:	1.	Kannada	4.	Malayalam
	2.	Kodagu	5.	Tamil
	3.	Kurukh	6.	Telugu
Tibeto-Burman:	1.	Gangte	4.	Paite
	2.	Kabui	5.	Thado
	3.	Meitei		
Austro-Asiatic:				
a. Munda:	1.	Birjia	5.	Mundari
	2.	Juang	6.	Santhali
	3.	Kharia	7.	Sora
	4.	Kuruku		
b. non-Munda:	8.	Khasi		

Table I: Modern South Asian languages

Indo-European:				
a. South Asian:	1.	Sanskrit	3.	Pali
	2.	Prakrit		
b. Outside:	4.	Greek	6.	Pashto
	5.	Hittite	7.	Persian
Dravidian:	1.	Old Tamil		
Austro-Asiatic:	1.	Malay	3.	Thai
	2.	Iban	4.	Kadazan

Table II: Ancient and 'outside' languages

One of A.'s major claims is that type (3) (and to some extent, (4)) reduplication of 'verbal adverbs' (VA) has aspectual functions. These are: (i) simultaneity of the actions of the VA and the main verb (MV) of the clause as in (3c); (ii) 'non-precipitation' where the MV action stops the incipient VA action, as in (5a) from Hindi; (iii) 'continuation-duration' as in (3d) and (4); (iv) 'iteration' as in (5b) from Hindi; and (v) 'sequentiality', where the action of the VA is 'long and often continuous', while the MV action is 'short and abrupt' (53); cf. (5c) from Konkani. A. claims that the 'iterative' and 'continuative-durative' functions are found in all of the South Asian languages and therefore must be considered a feature defining the South Asian convergence area.

(5) a. bāriš **hote hote** rah gaī
rain/happening/happening/stop/went
'It was going/about to rain, but it stopped.'

b. vah gānā **sun sun** kar thak gayā
(s)he/song/listen/listen/absolutive-marker/tired/went
'He got tired of listening to songs.'

c. **uleita uleita** taja dọleat dukā aili
speaking/speaking/his-from/eyes/tears/came
'He had been speaking (for a long time) and suddenly he began weeping.'

Chapter IV deals with reduplicated modifiers (adjectives, quantifiers, and adverbs). A. determines four major functions and discusses the degree to which they are present in different modifier subcategories. The four functions are (i) 'emphasis' as in (3a) above, (ii) 'attenuation' as in (6a) from Hindi, (iii) 'exclusiveness' as in (6b) from Hindi, and (iv) the well-known feature of 'distributiveness', cf. (3b) above. She finds that 'emphasis' and 'distributiveness' are most widely attested in South Asia; 'exclusiveness' is rare in Dravidian, more common in Indo-Aryan and Tibeto-Burman, and most common in Munda; while 'attenuation' is most common in Indo-Aryan, rarer in Munda and Dravidian, and least common in Tibeto-Burman.

(6) a. yah phal **khaṭṭā khaṭṭā** hai
this/fruit/sour/sour/is
'This fruit is sourish.'

b. **aurtẽ̀ aurtẽ̀** melā dekhne gaĩ̀
ladies/ladies/fair/see/went
'Only ladies went to see the fair.'

Chapter V examines in greater detail the geographical distribution of the functions distinguished in Chapters III and IV, as they occur with different subcategories of reduplicated structures.

Her findings suggest a core 'Reduplication Area' which includes the Munda languages and the Indo-Aryan languages Hindi-Urdu, Panjabi, Dogri, [Sadari], and Gujarati. The other Indo-Aryan languages, Marathi, Maithili, Oriya, Assamese, Konkani, Sindhi, Bengali, and Kashmiri, corresponding to Grierson's (1921) 'outer circle' of Indo-Aryan, agree with Tibeto-Burman and most of Dravidian in exhibiting reduplicated structures to a lesser degree. Finally, within Dravidian, Tamil and especially Malayalam are most 'resistant' to using reduplication. A. concludes that 'It appears that the structures originated somewhere in central India, perhaps in Munda languages, and then spread outside the Munda speech region' (96).'

Chapter VI is concerned with the function of reduplication in Austro-Asiatic, comparing members of the Munda branch of the family with non-Munda Khasi, with side glances at non-South Asian members of the family, as well as other languages of Southeast Asia, Austronesian, and Chinese.

A. finds that 'morphological' reduplication (of the type (1) and (2) above) plays a much more prominent role in Austro-Asiatic than in (modern) Indo-Aryan, Dravidian, and Tibeto-Burman. Within the Austro-Asiatic languages, however, the Munda branch stands out by making extensive use of 'lexically' reduplicated verbal-adverb structures of the type (3cd) which, A. claims, are absent even in Khasi (138).

This finding, if correct, would have interesting consequences for the historical interpretation of the geographical distribution of South Asian reduplicated structures. However, A.'s own data include the Khasi structure in (4) above (A.'s example (19), p. 130) which suggests that the distinction between Munda and Khasi is not as great as A. claims.

Chapter VII, entitled 'Reduplication in classical languages: Indo-European and Dravidian', attempts to use historical evidence to come to a better understanding of the origins of South Asian reduplication.

A. acknowledges Dressler's (1968) evidence for structures with 'lexical' reduplication in early Indo-European languages, as well as Macdonell's (1916:281-282) more extensive listing of Vedic forms of this type. Now, Macdonell gives only two Vedic structures with iterated verb, both involving imperatives, cf. (7),[3] but none with non-finite verbs (comparable to (3cd) above). This fact, combined with

'[t]he total absence of any discussion on reduplicated verbal adverbs ... by Panini or any other grammarian', leads A 'to believe that, perhaps, up to the period of 600 - 500 B.C. duplication of verbal adverbs ... had not emerged in Vedic or Classical Sanskrit' (150).'[4]

(7) a. píba-piba 'drink, drink' (RV 2.11.11)
b. yájasva-yajasva 'sacrifice, sacrifice' (ŚB 12.3.4.1 (2x))

Chapter VIII presents A.'s major conclusions. Although these are certainly significant and thought-provoking, it is not quite clear how some of them follow from her preceding discussion.

Apparently because of the widespread use of reduplication (in the larger sense) in Austro-Asiatic, A. feels that the reduplicated structures which she has examined are 'the gift of Munda on the Indian soil' (162). On the other hand, she claims that 'the development and widespread use of verbal adverbs ... in reduplicated forms representing various aspectual elements seem to be the internal innovation of Indo-Aryan languages' (162), even though they are not yet found in Vedic and the Classical Sanskrit of Pāṇini (161).

A. considers Dravidian origin of these constructions unlikely for several reasons: (i) Brahui lacks such structures; (ii) the earliest indigenous Dravidian (Tamil) grammar, Tolkāppiam, 'has no reference to the structures under consideration'; (iii) the range of functions is more limited in Dravidian than in Indo-Aryan and Munda; (iv) Dravidian offers fewer subtypes of reduplicated constructions (163).

Further claims include the suggestion that the widespread use of morphological reduplication in Munda and Proto-Indo-European may be attributable to early convergence 'and contact between the speakers of the two families of which we have no knowledge' (160).[5]

Finally, A. reiterates her earlier finding that the 'inner-circle' Indo-Aryan languages, Hindi-Urdu, Panjabi, Dogri, and Gujarati, together with Munda, form the core 'reduplicated area'. The 'outer-circle' Indo-Aryan languages, Konkani, Kashmiri, Maithili, Marathi, Oriya, Bengali, and Assamese, together with much of Dravidian constitute an area with attenuated reduplication. The languages with the least amount of reduplication are the most peripheral. These are Dravidian Malayalam in the extreme southwest, and Austro-Asiatic (but non-Munda) Khasi as well as Tibeto-Burman Thado in the extreme northeast (166).

As noted by Colin P. Masica in the Foreword, A.'s monograph is a 'significant contribution' to the study of the South Asian convergence area, addressing an aspect 'hitherto little explored', based on extensive fieldwork, and 'rich in original data ... It ... will repay careful study' (v). At the same time, as A. herself realizes (e.g. p. 9), a pioneering study of this type cannot yield satisfactory answers for all questions.

In the following I suggest some areas for improvement in a revised edition, when such an edition will be published.[6]

First, A.'s use of the term 'reduplication' to refer to all of the structures in (1) - (4) is unfortunate, since the structures differ considerably, both in form and semantics. Using a single term can blur these distinctions and/or confuse the reader; and A.'s distinction between 'morphological' and 'lexical' reduplication only partly alleviates the problem. At the same time, in spite of their differences, the structures may exhibit various degrees of overlap, especially in function. A fuller discussion of these similarities and differences and a more differentiated terminology would certainly be helpful in establishing a clearer focus for the discussion.

Let us begin by examining the formal similarities and differences between these structures:[7]

The type (1) involves MORPHOLOGICAL DERIVATION similar to affixation. This is the type for which the term 'reduplication' traditionally is most commonly used. Like other types of derivational morphology it may be more or less productive; but it tends to exhibit much of the idiosyncrasy characteristic of derivational morphology. For instance, while in Sanskrit perfects, the vowel of the reduplication syllable is sensitive to the underlying root shape (as in *u-vāc-a* from /**u**ac-/ 'speak', *yu-yoj-a* from /**yu**j-/ 'yoke'), in other forms it is sensitive to the surface root shape (as in the corresponding desideratives *vi-vak-ṣa-ti* 'desires to speak' vs. *yu-yuk-ṣa-ti* 'desires to yoke').

The type (2) also is morphological, but it belongs to the category of COMPOUNDS, formally similar to structures like Skt. *divāniśam* '(by) day (and) night'. Like other compounds these structures constitute single lexical units, with inflectional endings added at the end of the compound, not to each of the component parts. Moreover, excepting lexicalized forms, compounds generally exhibit fewer idiosyncrasies than structures of the type (1). The term 'iterative compound' may be useful to distinguish type (2) from type (1) structures.

Structures of the type (3) look like ITERATIONS of fully inflected words. (For instance, both instances of *baiṭhe* in example (3d) contain the adverbial affix *-e*.) As such, they behave more like syntactic collocations than like morphological compounds. The traditional term for this type in Sanskritist and Indo-Europeanist literature is 'āmreḍita'; a less exotic term would be 'lexical iteration'.

The type (4), finally, looks very similar to SYNTACTICALLY conjoined (or quasi-conjoined) structures of the English type (8) and the Vedic Sanskrit type (9). These two types differ only in terms of the method of signaling coordination. In the very widespread type (8), an overt coordinating conjunction is employed, while in (9) coordination is signaled by the use of a case form (in the present case, the instrumental in 'comitative' function).[8] It is convenient to refer to this type as 'iterative conjuncts'.

(8) a. better **and** better
b. through **and** through
c. Turning **and** turning in the widening gyre | The falcon cannot hear the falconer (Yeats, *The second coming*, 1)

(9) yudh-á́ yudham úpa ... eṣi (RV 1.53.7)
battle-instrumental/battle-accusative/to/you-go
'You go to battle after battle/every battle.'

From the formal perspective, then, the four different structures can be ranked in the order (1) - (2) - (3) - (4), with (1) most prototypically morphological and (4) closest to being fully syntactic.

However, the distinction between these types is not absolute. For instance, in languages without inflectional affixes, the types (2) and (3) are difficult to distinguish. In fact, even the Sanskrit type (3) exhibits behavior close or even identical to that of compounds.[9] Thus, in accented texts, only the first of the iterated items bears accent; cf. (10a).[10] Moreover, clause-second particles follow the entire structure rather than the first element; cf. (10b) vs. (10c).[11]

(10) a. sá vaí **saṁmṛ́jya saṁmṛjya**
he/particle/wipe-absolutive/wipe-absolutive
pratápya pratapya prá yacchati (ŚB 1.3.1.8)
heat-absolutive/heat-absolutive/gives
'He gives (the ladle to the adhvaryu) after each wiping and heating.'
b. **vyatihā́raṁ vyatihāraṁ** h(i)
transpose-absolutive/transp.-absolutive/'for'(P2-particle)/
uttaravedíṁ vyāghāráyanti (ŚB 9.2.1.7)
uttaravedi/over-sprinkle
'For they over-sprinkle the uttaravedi, transposing again and again.'
c. ***vyatihā́raṁ** hi **vyatihāraṁ** ...

On the other hand, where there is no evidence (as in (10)) that structures of the type (3) are compounds, it may be difficult to reject an alternative analysis of these as being 'asyndetic' (i.e. conjunctionless) counterparts of the syntactically (quasi-)conjoined type (4), (8), or (9).

The formal overlaps between the types (2) and (3), as well as (3) and (4), may be significant, suggesting a possible diachronic path of grammaticalization, from the most 'syntactic' type (4), via the quasi-compound type (3), to the full-compound type (2).[12]

Functionally, too, the types (2) - (4) exhibit a great degree of affinity: All three of them most commonly signal either intensitivity (as in (2a), (3a), and (8b)) or distributiveness and iterativity (as in (2b), (3b), (8a), and (9)). By generally exhibiting this very restricted, and highly iconic, range of functions, these types differ markedly from the type (1), whose range of functions is much broader and

generally much less iconic. (The examples in (1) illustrate only a small subset of these functions.)

For these reasons it is legitimate to distinguish types (2) - (4) as a group from type (1), by referring to them as exhibiting 'iteration', while type (1) exhibits the much more clearly morphological process of 'reduplication'.

Functionally, even this distinction is not absolute: As is well known, some subtypes of morphological reduplication exhibit the same iconic functions of intensity and distributeness/iterativity as the types (2) - (4); cf. e.g. the Sanskrit 'intensive' type (1c). Structures of this type, therefore, may be relevant when discussing constructions of the type (2) - (4). Other forms of reduplication, however, would generally be of questionable value.

The differences between these four different types of structure have significant implications for establishing genetic relationship or contact.

In principle, for instance, similarities in the purely morphological type (1) would be most probative, especially if they are highly idiosyncratic (rather than iconic), since such similarities are least likely to result from chance. In this regard, then, A.'s claimed parallelism of Austro-Asiatic and early Indo-European as regards morphological reduplication might be significant. However, this parallelism holds only at a very general level; the details of reduplication differ to such a degree that a special prehistoric relationship is quite unlikely.

On the other side of the spectrum, the iterative conjunct type (4) and (8/9) has the least probative value in comparative work. It is highly iconic (rather than idiosyncratic), draws on the crosslinguistically very productive syntactic process of coordination, and as a consequence is found in language after language after language. Evidence of this type, therefore, is best ignored in trying to establish a special (contact or genetic) relationship, unless the nature of the evidence happens to be highly idiosyncratic.

A.'s best evidence lies in the area of types (2) and (3), i.e. in the area intermediate between the most morphological and most syntactic types. Structures of this type are more likely to exhibit idiosyncrasies than type (4) structures. For instance, while all varieties of English have the 'iterated conjunct' type (8), 'lexical iteration' is relatively restricted, especially in American English, where even the British type *It's very very difficult* tends to be avoided. Even a priori, then, similarities in structures of the type (2) and (3) are more likely to be indicative of a special, contact or genetic, relationship.

In addition, however, A. has been able to isolate a number of specific South Asian idiosyncrasies. One of these is found in the 'attenuated' and 'exclusive' functions of the types (6a) and (6b) which

differ from the normal, iconic, intensive and distributive/iterative functions of these structures.

A second idiosyncrasy consists in the uneven, but patterned, geographical distribution of the structures, which strongly suggests spread from a core area.

A third, and perhaps most significant, idiosyncrasy lies in the existence of the iterated verbal-adverb type (3cd) and its special semantic functions, a phenomenon which does not seem to be found in related languages outside of India.

So far, so good. But several questions arise. One of these is the typological one of how idiosyncratic these structures really are. Another set of interrelated issues concerns the reliability of A.'s data: Are they sufficient to establish that the reduplicated structures originated in Munda, that the 'reduplicated verbal-adverb' type is an Indo-Aryan innovation, and that this type developed in post-Pāṇinian times? And can we be certain that it was absent in early Dravidian?

Now, while 'attenuated' and 'exclusive' functions are not a usual crosslinguistic feature of iterated structures, there are parallels in other languages, as well as historical (near-)antecedents in Sanskrit. Both functions can be documented in English. Attenuation is observed in the expression *so-so* (which has parallels in many other European languages), even though this type is quite marginal. Something close to exclusivity is found in the much more productive, quite colloquial, English pattern (11). But perhaps this pattern is of fairly recent origin.

(11) It's not a mickey-mouse book; it's a **book book**.

More significant is the fact that similar semantic specializations are found relatively early in Sanskrit. Delbrück (1888:55) notes the use of iterated *upári* 'over' and *adhás* 'below' in exclusive function in structures like (12a); and in Delbrück 1900:151 he adds that this fact was already known to Pāṇini (cf. his rule 8.1.7). Upon closer examination, Pāṇini's rules yield another, similar type of iterated adposition with exclusive function (8.1.5), as well as an adjectival iterative compound type with attenuated function, cf. (12b). Attenuation is no doubt also found in the type (12c) mentioned in Delbrück 1888:55, 572, but without discussion of its function. (The attenuated function is adumbrated by Harisvāmin's *alpārthe* 'in the meaning of little' in his commentary on the passage.) An exclusive function of iteration may even be found as early as the Rig-Veda; cf. (12d).

(12) a. átha dákṣiṇāṁ bhrúvam **upáry-upari** lalāṭam úpa spṛśati (ŚB 3.2.1.29)
then/right/brow/above-above/forehead/to/touch
'Then he touches the forehead JUST above the right brow.'

b. prakāre guṇavacanasya (8.1.12)
'[Iteration] of an adjective in the meaning "of sorts".'
Example: paṭu-paṭu- 'rather, somewhat hot/spicy'

c. tásmin **yā́vān vā yāvān vā** rásaḥ sám asravat (ŚB 4.4.3.4)
in-it/how-much/or/how-much/or/juice/together/flowed
'Into it, just so much juice flowed.'

d. **tvā́m-tvaṁ** aharyathā(ḥ) ... (RV 10.96.5)
you-you/were-desired
'Only you were desired ... ' (Geldner: 'immer nur du ... ')

The fact that the attenuated and exclusive functions of iteration thus go back to the earliest stages of Sanskrit might be interpreted as indicating that these functions originated in Indo-Aryan.

This conclusion would be compatible with A.'s observation that the geographic core area of attenuation consists of the Indo-Aryan languages. But it seems to be in conflict with A.'s claim that the core area of exclusivity is Munda. Unfortunately, no Munda texts (or texts from any other South Asian language group other than Indo-Aryan) can match the antiquity of Vedic Sanskrit. As a consequence it is not possible to judge whether at a comparable prehistoric time, Munda did or did not have iterated structures with exclusive (or attenuated) functions. We are thus unable to determine whether A.'s evidence for Munda being the core area of modern South Asian exclusivity should be interpreted as indicating that Munda was the originator of these feature, or whether for some reason the feature may have become more productive in modern Munda.

The situation is similar as regards A.'s claim that iteration in general is of Munda origin. True, A. finds that iterated structures are most common in Munda and the neighboring Indo-Aryan languages. And it is certainly tempting to assume that the greater use of iterated structures in the latter, 'inner-circle' Indo-Aryan languages, as compared to the 'outer-circle' ones, is due to contact with Munda. However, here again we find that iteration is a feature found even in the earliest stratum of Indo-Aryan, Vedic Sanskrit, a period for which we simply have no counterparts in Munda or any of the other South Asian languages. Moreover, it is also found, to varying degrees, in other Indo-European languages, as well as in non-Indo-European languages outside South Asia.

Before giving selected examples and references, however, it is important to note that for many of these languages it is much more difficult to find relevant evidence than for Sanskrit. Several factors seem to be responsible for this fact: In many languages iteration is a feature of the colloquial language (cf. e.g. Hofmann 1951:58-61). As a consequence, it may not be found freely in the formal texts that have been handed down (cf. Bartholomae 1907:166). Moreover, while the tradition of indigenous Sanskrit grammar offered a formal treatment of iteration under the notion āmreḍita (Pāṇini 8.1.1-13) and thus invited discussion of this phenomenon in western accounts, the tendency of traditional western grammars of such languages as Latin and Greek has been to treat iteration informally, under the heading of STYLISTICS (cf. e.g. Hofmann 1965:§45, Menges 1953:551, Schwyzer

1949:613, 699-100). For these reasons, one has to be extremely careful in interpreting the information provided by traditional western accounts of languages other than Sanskrit (and occasionally even of Sanskrit). In some cases it is possible to supplement that information from individual articles or from primary textual research. But where such subsidiary information is not available, the absence of any discussion of iteration, or the presence of just a few cursory notes, should not be taken to indicate that a given language had little or no iteration.

Studies of the Rig-Veda, the oldest stage of Sanskrit, have established that lexical iteration is most common for nouns, pronouns, and preposition-adverbs; it is less common for adjectives and numerals; and exceedingly rare for verbs, for which there is just one certain example with accent only on the first item (cf. (7a) above, as well as note 10). See for instance Collitz 1882, Delbrück 1888:51-55, 1900:141-153, Wackernagel 1905/Wackernagel & Debrunner 1957: 142-148, Wackernagel 1929-30:395-396 (for numerals).

Moreover, the Rig-Veda offers several examples of iteration with accent on BOTH elements (see note 10), a fact which might indicate that the normal type of lexical iteration with accent only on the first element is an innovation. Perhaps significantly, all of these examples are exclamations, and the iteration merely signals emphasis, not the 'aspectual' functions of distributivity or iterativity.

Interestingly, the examples of finite-verb iteration that I have found in the post-Rig-Veda Vedic language preponderantly are imperatives (i.e. exclamatory in nature), and with one possible exception likewise signal emphasis rather than aspectual functions (cf. (7b) above with note 3). Notice in this regard that iteration of imperatives, vocatives, and other 'exclamatories' for emphasis is an extremely common phenomenon, even in languages with much less robust evidence for lexical iteration than Sanskrit; cf. e.g. Hofmann 1951:58-59, 1965:§45.

These facts might support the view that verbal iteration with aspectual functions is in fact an innovation in Sanskrit/Indo-Aryan.

Note however that traditional accounts of Rig-Vedic Sanskrit do not accord special treatment to one type of verbal iteration which does seem to be used in aspectual functions, viz. NON-finite verb iteration; cf. the examples in (13). To the extent that such structures are noted, they are classified as adjectives. And in fact, their use IS rather adjectival.

(13) a. **satáḥ-sataḥ** pratimā́nam̐ parobhū́r (RV 3.31.8a)
of being-being/equality/surpassing
'Equal of every existing one/every being, surpassing ...'

b. **suté-sute** vāvṛdhe ... (RV 3.36.1c)
at pressed-pressed/grows
'He grows at every pressed (soma) ...'

c. **jātó-jāto** jāyate vājy àsya (RV 7.90.2d)
born-born/is-born/victor/his
'Every born (son) of his is born a victor.'

d. **pányaṁ-panyaṁ** ít sotāra ấ dhāvata (RV 8.2.25ab)
to-be-praised=to-be-praised/particle/pressers/up/stir
'Stir up, O pressers, the one to be praised again and again.'

At the same time, however, the iterated lexical items are VERBAL adjectives, and they exhibit the aspectual functions characteristic of the modern South Asian type (3c/d).

The first unambiguous examples of non-finite verb iteration with aspectual functions come from one of the earliest Vedic-Prose texts; cf. (14). Note also the somewhat later examples in (10) above.

(14) taíh **saṁstambhấṁ-saṁstambham** ásurān ajayat (MS 1.4.14)
they-instrumental/suppressing-suppr./Asuras/defeated
'Suppressing the Asuras again and again by means of these (victory sacrifices) he defeated (them).'
(Similarly ibid. and MS 3.8.1, 3.10.5)

Perhaps significantly, these unambiguously aspectual verbal iteration structures involve absolutives, i.e. verbal adverbs, rather than verbal adjectives. If the difference between the Rig-Veda and later Vedic Prose reflects a difference in chronology, rather than genre,[13] this fact may indicate that the iterated verbal-adverb construction is a post-Rig-Vedic, Vedic-Prose innovation. While this would support A.'s claim that the construction is an Indo-Aryan innovation, it dates the innovation at a considerably earlier period.[14]

Now, as is well known, the origin and antiquity of the Sanskrit absolutives is a matter of controversy. Many scholars consider them an Indo-Aryan innovation, resulting from Dravidian influence; cf. e.g. Emeneau 1954, 1956, 1962, 1980, Kuiper 1967a, b, 1991, and Kaufman & Thomason 1988. Following earlier suggestions I have argued that they may be inherited and that, at any rate, the arguments for Dravidian influence have not been established beyond a reasonable doubt (Hock 1975, 1982, 1984). While accepting the latter claim, Tikkanen (1987) postulates a different, northwestern substratum as the source for the Sanskrit absolutive.

Depending on one's position regarding this controversy, the fact that aspectual verbal iteration is first attested in Sanskrit for absolutives may suggest different conclusions. Accepting my arguments would make it relatively easy to follow A. in considering structures of the type (3cd), (10), and (14) an Indo-Aryan innovation, perhaps resulting from extension of the verbal-adjective pattern in (13) to the adverbial absolutives. Accepting the Dravidian-substratum hypothesis might be taken to support Dravidian origin for this type. This view would run into the difficulty that A. finds structures of this sort to be less common and the range of functions more limited in Dravidian than in the Indo-Aryan and Munda core area. But perhaps it

is only the wider range of functions which is an innovation of the core area. Finally, accepting Tikkanen's hypothesis would open up yet further opportunities for speculation.

Unfortunately, here again our ability to decide between these different possible scenarios is limited by the fact that none of the other languages is attested at anything approaching the time depth of Vedic Sanskrit.

On the other hand, comparative Indo-European evidence suggests that iterated verbal-adverb structures are indeed an innovation of early Indo-Aryan. Even Avestan, the most closely related early Iranian language has much more restricted evidence of iteration, with examples limited to nouns, pronouns, and numerals.[15] All of these, to be sure, have the distributive/iterative functions which A. considers a feature of the entire South Asian area. But none of them involve verbs (whether finite or non-finite). At the same time, as Bartholomae (1907:166-167) notes, the dearth of examples may be attributable to the nature of the Avestan texts. More examples of iteration are found in later Iranian. See also Heston 1980.[16]

The majority of the non-Iranian Indo-European languages offer even less evidence for iteration with distributive/iterative function, at least in their earliest stages. In early Greek, for instances, the cited evidence is limited to Mycenaean *we-te-i-we-te-i* 'year after year', Cypriot *a-ma-ti-a-ma-ti* 'day after day' (Dressler 1968), the Homeric iterated verbal prefix *propró-* 'forward and forward' (Collitz 1882: 298, Dressler 1968, Dunkel 1981a), and a few other less obviously distributive forms (for which see e.g. Dressler 1968). To these forms Dunkel (1981a,b) adds a few more early examples of iterated verbal prefixes. Excepting intensive iteration, especially of imperatives, vocatives, and other 'exclamatories', early Latin distributive/iterative examples are essentially limited to iterated interrogative pronouns in indefinite-pronoun function, such as *quisquis* 'whoever'. Dunkel (1981ab) adds a few possible examples with iterated preposition/adverbs. Iterated numerals in distributive function are attested quite late, both in Greek and in Latin.

Interestingly, iteration seems to be much more widely attested in Hittite. Dressler (1968) lists several examples of iterated locative nominals in temporal function, such as KASKAL-*ši* KASKAL-*ši* 'every time'. To these Dunkel adds several examples of iterated preverbs or adverbs in distributive/iterative function (1981a, with references). In addition, Hittite has iterated interrogatives used as indefinites, and at least one example of an iterated numeral; cf. (15a). Dressler (1968) claims that it is not certain that the iterated structures function like compounds as they generally do in Sanskrit. But examples like (15b) suggest that they do, since sentence-initial particles (indicated by underlining) follow the entire structure, rather than the first element.

(15) a. nu **1-aš 1-aš** ... šešuwanzi liebe karaštari
now/one/one/sleep/not/fail
(KUB III 4,5,6,17,18,19 III 5-6)
'Now let them not fail, one by one, to sleep ... '

b. **MI-ti MI-ti**-ma ... (KUB III 4,5,6,17,18,19 III 12-13)
night/night/but
'But night after night ...'

While most Indo-Europeanists agree that the comparative evidence of the early Indo-European languages warrants reconstructing some kind of iterated structure for the proto-language, they disagree as to what structures should be reconstructed. Delbrück (1900:152-153) considered it likely that Proto-Indo-European had iterated pronouns, numerals, and preposition-adverbs. Dressler (1968) claims that iterated locative nominals can be reconstructed, but no other case forms of iterated nominals. Most recently Dunkel (1981a) argues that only iterated prepositions can be reconstructed, since only in this category do we find exact equations (of the type Skt. *prá pra* : Gk. *propró*). Dunkel's position may be overly restrictive: In quasi-syntactic structures of this type it is unrealistic to expect exact equations, just as we should not expect exact equations in sentences. What we can reconstruct in such cases, and even for many productive morphological processes, are structural PATTERNS (cf. Hock 1985 and 1986/1991:576-577). Moreover, Dunkel's preverbs, whose behavior in early Indo-European suggests that they started out as adverbs, are quite similar to Dressler's locative nominals, whose function likewise is adverbial. Hittite evidence of the type (15a), combined with the Indo-Iranian evidence, further suggests that iterated numerals may have been possible in early Indo-European. And the agreement of Indo-Iranian, Latin, and Hittite in using iterated interrogative pronouns as indefinites might perhaps be taken to suggest reconstruction of this type, as well. But note that, if a pattern of iteration was inherited from the proto-language, it is perfectly possible that it was independently extended to any one of the structures just mentioned. That is, although we can be quite certain that Proto-Indo-European had a pattern of lexical iteration, it is much more difficult to know which lexical categories were permitted to undergo iteration.

Significantly, however, there is no comparative evidence for iterated VERBAL structures, either finite or non-finite, with distributive/iterative function.[17] The comparative Indo-European perspective, then, seems to support A.'s claim that iterated verbal adverbs are an Indo-Aryan innovation

Non-Indo-European languages outside South Asia, however, do provide evidence for iterated verbal structures and thus present a challenge to the view that the constructions are a uniquely South Asian phenomenon.

To judge by the sample folk stories from non-literary Uralic languages in Collinder 1965, finite-verb iteration in iterative-contin-

uative function is not uncommon in Uralic. Compare e.g. Mordvin *sokan sokan da sizińʹ koda* 'I plow and plow/I keep plowing, and how tired I got'.[18] For one language, Votyak, Collinder further cites the struture in (16), with iterated 'modal gerundium', a construction remarkably similar to the South Asian type (3cd). Unfortunately, Collinder does not indicate how widespread this structure is in Uralic. If it is quite isolated, it might either be a spontaneous internal development or reflect Turkic influence.

(16) **berdysa berdysa** gurtaz koškem
crying/crying/home/went-narrative
'She went home crying and crying/crying bitterly.'

Most important is the evidence of Turkish. In addition to iterated adverbs and adjectives in intensive/emphatic function (Lewis 1953:56, 59), it offers a large variety of iterated structures in distributive/iterative function, including numerals (Lewis 1953:69), adjectives modifying plural nouns (Lewis 1967:236), nouns (ibid.), finite verbs (ibid.235), or non-finite gerunds in iterated verbal-adverb structures of the type (17) for which see Lewis 1967:175-176.

(17) insan belki **döğüle döğüle** uslanır
person/perhaps/being-beaten/being-beaten/becomes-well-behaved
'Perhaps one becomes well-behaved with being constantly beaten.'

Structures of the latter type are strikingly similar to the South Asian iterated verbal-adverb structures in (3cd). Moreover, like their Sanskrit counterparts, they bear accent only on their first element (Lewis 1953:56, 1967:176), while at the same time fully inflecting both of the iterated elements

These similarities by themselves would be remarkable enough to provide an interesting challenge to the view that lexical iteration, especially of the verbal-adverb type, is a uniquely South Asian feature.

However, these are not the only structural similarities between Turkish (and Turkic in general) and the languages of South Asia. Excepting retroflexion, all the features commonly considered characteristic of the South Asian convergence area recur in Turkish: basic SOV order; the widespread use of absolutives and other non-finite devices instead of, or beside, clausal subordination; and the use of quotative constructions marked by a post-citation quotative particle. Moreover, Turkish shares with Indo-Aryan and Dravidian the use of 'relative-correlative' constructions which, as in Dravidian, are more marginal than relative verbal participle structures. Compare Hock 1988, 1989, 1992, based in part on Steever's pioneering study of finiteness in Dravidian (1988).[19]

In addition to posing a significant challenge to the view that all of these features uniquely define the South Asian convergence area,

these similarities raise questions which, to my knowledge, have not been satisfactorily addressed so far, not just in A.'s monograph, but in any of the studies devoted to South Asia convergence: How do we explain these remakable similarities? Can they be due to chance? Or must they reflect some kind of bilingual contact? If we are really dealing with chance similarities, could some or all of the South Asian similarities likewise be due to chance? If there was bilingual contact, where in time and space should it be located? And how do we account for the fact that the Turkic languages are more similar in their syntax to the Dravidian languages than to Indo-Aryan which, in the form of Sanskrit, had a much better chance of coming into contact with Turkic in Buddhist Central Asia? Or should we assume that all of the features reflect inheritance from an often posited, but so far not well established, common ancestor of Dravidian and Turkic (or Altaic, the larger family of which the Turkic languages are commonly considered to be a member)?

These are difficult questions indeed. And it would be too much to expect A. to find satisfactory answers in a revised edition. Nevertheless, even if no answer is possible at this time, the evidence of Turkic needs to be considered more fully in assessing to what extent iteration, especially the verbal-adverb type, is a feature uniquely defining the South Asian convergence area.

A revised version of A.'s monograph might further benefit from taking another, closer look at what can be known about the earlier history of Dravidian. First, the range of functions in which iterated verbal-adverb structures are used in Dravidian does not seem to differ significantly from their range in Indo-Aryan and Munda. From A.'s discussion it appears that only one function, that of 'non-precipitation' in (5a) above, is absent in several Dravidian languages. Secondly, even if the oldest Dravidian grammar, Tolkāppiam, does not explicitly mention iterated verbal-adverb constructions, this does not necessarily mean that such structures did not exist. It is very well possible that their existence is assumed or subsumed by the general rules that permit lexical iteration. Only an examination of Old Tamil texts, and of the oldest texts in the other Dravidian languages with a long literary tradition, can yield some degree of certainty in this regard (unless there are genre-considerations that would preclude the use of such structures).

At the same time, Tamil, in many respects the most conservative Dravidian language, offers certain forms which might suggest that very different structures, with morphological (quasi-) reduplication, may have originally been used instead of some of the modern iterated structures — or perhaps even all? Beythan (1943: 174) states that intensive adjectives are formed by a process of quasi-reduplication, with infixation of *-n̲n̲aN-* or some other element between the reduplication syllable and the stem; cf. (18a). And certain distributive numerals likewise seem to be formed by some kind of reduplicative prefixation; cf. (18b) from Beythan (1943:143).

(18) a. karutta 'black' : ka-nnaṅ-karutta 'very black'
b. onru 'one' : o-vv-onru 'one each'
beside iterated onru-v-onru
iraṇṭu 'two' : i-vv-iraṇṭu 'two each'
mūnru 'three' : mu-m-mūnru 'three each'

Should structures of this type turn out to be genuine archaisms, rather than relatively recent reductions of originally iterated constructions, this might support A.'s view that the principle of iteration spread from an Indo-Aryan/Munda core area to Dravidian.

Finally, I would like to express the hope that a revised edition will receive greater editorial care by the publishers. There are all too many passages whose interpretation is unclear or which seem to be self-contradictory. (See for instance notes 4 and 5 above.) These passages do injustice to the great amount of research that A. has done and to her extremely interesting conclusions.

As the length of this review may suggest, A.'s monograph makes significant and challenging contributions to the study of South Asian convergence. It should be considered 'must' reading for scholars concerned with this topic, no matter whether they agree with all of her findings or not.

NOTES

[1] But note that Burushaski, in the extreme north of the area, also exhibits the major diagnostic features defining the South Asian convergence area. At the same time, the quotative construction is absent in many of the modern Indo-Aryan languages, as well as in the northwestern Dravidian language Brahui. (It was present in the earliest attested Indo-Aryan language, Sanskrit.) Indo-Aryan Assamese, in the extreme northeast, lacks the retroflex-dental contrast. (Again, the contrast was present in Sanskrit.) And so on. That is, as in many convergence areas, the isoglosses correlated with the presence of particular features do not form neat bundles that boldly define the area but rather, like dialectal isoglosses, more irregular patterns, especially in the transition areas between the core and the peripheral languages outside the area. Cf. e.g. Masica 1976 for South Asia and the discussion in Hock 1988a.

[2] In her discussion in Chapter 2 she further includes 'echo formations' of the type Hindi *pānī-vānī* 'water and the like' (from *pānī* 'water'), Tamil *paṛam-giṛam* 'fruit and the like' (from *paṛam* 'fruit') and compounds of the type Hindi *dhan-daulat* 'wealth' (*dhan* 'wealth' + *daulat* 'wealth, money') or Burmese *thwa la* 'travel' (*thwa* 'come' + *la* 'go'). She admits that the latter type is 'not duplicated on the phonological level'. Moreover, both types of structures do not figure prominently in her further discussion.

The type *dhan daulat*, by her own admission, is not limited to South Asia but is also found in Burmese, Thai, and Malay. In fact, it recurs in Tocharian (Toch. A *ñom-klyu* 'name-fame' = 'fame'), Old Irish (*gaisced* 'weapons' = *gáe* 'spear' + *scíath* 'shield'; Meid 1968), Lithuanian (cf. Senn 1966:351), and many other Indo-European languages (references in Meid 1968), as well as in Uralic (Collinder 1965:49, 63) and Turkish (Lewis 1988:§43, 45). This type therefore is not a unique property of South Asia, except perhaps in terms of frequency of occurrence.

'Echo words' likewise have outside parallels. Compare e.g. Turk. *sıkı fıkı* 'intimate' (from *sıkı* 'close'; Lewis 1988:236); Engl. *hurly-burly, pell-mell*, etc.; Lat. (Plautus) *at bat* (pejorative, from *at* 'but'), *eia beia* (from *eia* 'up and at it'); and especially the Yiddish-based English type *syntax-shmyntax*. Most of these echo formations are not very productive and do not exhibit recurrent morphological characteristics, in stark contrast to the South Asian languages. Moreover, the only well-known PRODUCTIVE type, that of Yiddish, only has negative connotations. But as Steever (1988) observed, the echo formations of Tamil likewise have negative connotations. Moreover, if we expand our horizon beyond the (western) European languages, we do find at least one highly productive parallel outside South Asia, viz. the colloquial Turkish type *Burada **kutu** **mutu** yok* 'there is no box OR THE LIKE here' (Lewis 1953:112); and this type does not seem to have negative connotations. The question whether productive echo formation is an exclusive feature of the South Asian convergence area deserves further study.

[3] Even for finite verbs this list is incomplete. The Śatapatha-Brāhmaṇa additionally offers *yája-yaja* 'sacrifice, sacrifice' (1.5.3.8 (2x), 2.5.2.30,41), *véda-veda* 'I know, I know' (14.6.7.5); the Aitareya-Brāhmaṇa, *aireyathām-aireyathām* 'did you divide, did you divide?' (6.15.12); the Jaiminīya-Brāhmaṇa, *jayati-jayati* 'he wins again and again' or 'he wins, he wins (indeed)' (2.293) and *dadāni-dadāni* 'let me give, let me give' (2.213). A complete investigation of finite-verb iteration in Vedic Sanskrit still is very much a desideratum.

[4] A.'s discussion on this point, unfortunately, is not very clear. She acknowledges one example of an iterated absolutive in the Jaiminīya-Brāhmaṇa which I had brought to her attention, then states that this 'is an absolutive construction which even Panini does not exclude from his treatise', and then continues, 'It is the reduplicated absolutives that he does not discuss at all' (151). Perhaps the latter sentence should read 'It is the reduplicated VERBAL ADVERBS that he does not discuss at all', assuming that A. makes a distinction between reduplicated absolutives and other 'reduplicated verbal adverb' structures.

Such a distinction might make sense for modern Indo-Aryan languages like Hindi; cf. e.g. A.'s example (22) on p. 45 vs. (37) on p. 51. However, to judge by the Tamil, Telugu, and Kannada data, the corresponding Dravidian structures only show past (or perfective) ABSOLUTIVES; and similarly, the Dravidian structures corresponding to modern Indo-Aryan adverbial present-participle constructions like (3c) above have present (or imperfective) absolutives. Moreover, while Sanskrit does have locative (or genitive) absolute constructions involving quasi-adverbial case forms of participles, the absolutive is the main verbal adverb of Sanskrit. And the functions of Sanskrit iterated-absolutive structures are very much the same as A.'s modern verbal-adverb constructions, no matter what specific adverbial form of the verb they might employ. In the subsequent discussion, therefore, I treat Vedic Sanskrit iterated absolutive structures (and other structures with iterated non-finite verbs) as relevant for evaluating A.'s conclusions. (Note incidentally that Pāṇini does make an explicit reference to iterated absolutive structures, but in the section on absolutives (3.4.22), not in the āmreḍita section.)

[5] As in a number of other places, A.'s discussion of this matter is somewhat confusing: She first claims a special affinity between Old Iranian and Munda in respect to morphological reduplication, adding that 'At his juncture ... one cannot be very sure whether such a parallel can be established for Vedic Sanskrit' (§8.3.1). But in the next section (§8.3.2) she points to 'structural parallels between Proto-Munda reduplicated verbs ... and Proto IE reduplicated verbs', and cites (Vedic) Sanskrit examples to illustrate the Indo-European pattern.

[6] For obvious reasons these comments are confined to areas with which I am familiar. It is to be hoped that A.'s book will receive similar comments from experts in other relevant areas.

[7] The following discussion does not make any claims to originality but simply summarizes characteristics that are generally well known and have been widely reported and discussed in the literature.

[8] For this type of structure see Hoffmann 1960.

[9] Pāṇini does not appear to have considered this type a compound: He does not account for it in the compound section (2.1.1-2.38), but in a completely different part of his grammar (8.1.1-8, 57, 2.95, 103, 3.12, 49, plus 3.4.22). On the other hand, the padapāṭha of the Rig-Veda analyzes structures of this type as compounds.

[10] One early exception is Rig-Vedic *stuhí stuhí ‿ id* 'praise, praise indeed' (RV 8.1.30). Here the emphasizer *id* might be responsible for the accentuation of the second element. But this explanation does not apply to the other certain examples with accent on both

elements of the iterated structure, *vā́ṣaṭ vā́ṣaṭ ... námo nám(o)* 'hail, hail ... honor, honor' (RV 10.115.9) and *áraṇyāny áraṇyāny* 'O wild woman, O wild woman' (RV 10.146.1). Wackernagel & Debrunner (1957:144), listing only the last example, wonder whether the single-accent pattern may not be a relatively recent innovation for these structures. Interestingly, all of these exceptions are exclamatory, either imperatives (8.1.30), or sacrificial exclamations (10.115.9), or vocatives (10.146.1).

[11] Elsewhere, the particles can follow the first word, rather than just the first constituent; cf. e.g. Hock 1982, Schäufele 1990.

[12] Wackernagel & Debrunner (1957:146-147) map out several ways in which Sanskrit structures of the type (3), with lexical iteration, can become iterative compounds of the type (2). Similarly, Pāṇini (8.1.9-15) lists certain iterative compounds as special types of āmreḍita, i.e. of lexical iteration. As Hoffmann (1960) points out, the quasi-conjoined type (9) above likewise has furnished iterative compounds in Vedic, a pattern which continues into Middle Indo-Aryan.

[13] Cf. Hock In Press on this difficult issue.

[14] See also note 4 above.

[15] No single publication known to me gives all the examples. The following is a composite listing from Delbrück 1900:144, 148, Bartholomae 1907:166, Wackernagel 1929-30:395, Mayrhofer 1992: 34-35 with references. (Duchesne-Guillemin's listing (1936:43) is disappointingly incomplete.) For complete citations see Bartholomae 1904:svv. The examples in (i) are nominal iterations; those in (ii) iterations of interrogative pronouns used as indefinites (with or without the indefinitizing particle *cit*); (iii) an iterated demonstrative pronoun used as indefinite (for parallels in Sanskrit and other Indo-European languages see Brugmann 1904:130-131); (iv) what appears to be an iterative COMPOUND of a numeral, rather than lexical iteration; and (v) an iterated noun which has been reinterpreted and univerbated as an adverb, apparently in Proto-Indo-Iranian.

(i) nmāne nmāne vīsi vīsi 'in home after home, clan after clan'
narām narəm 'man after man'
manō manō 'attack after attack'
(ii) kaŋhe kaŋhe apaγzāire 'in each gully'
kəm kəm cit aipi nmāne 'with everyone in his house'
(Bartholomae 1904:425-426 lists three similar examples)
(iii) aēm aēm 'anyone'
(iv) baēvarə baēvaranąm 'myriads after myriads'
(v) nanā̆ 'separately', Skt. nā́nā (id.) < *nā́ nā 'man after man'

[16] Modern Ossetic has a fair number of iterative compounds; cf. Abaev 1964:116. Some of these, such as *raxæc-raxæc* 'pulling' are

listed as 'verbal'; and to judge by the gloss 'while hobbling', one example, *k'uilix-k'uilix*, looks like an iterated verbal-adverb structure. However, no context is given, and no indication as to how frequently such structures are used. Independent evidence for Turkic influence on Ossetic (Thordarson 1989:457) makes it at least possible that the Ossetic structures result from contact with Turkic.

[17] Modern Lithuanian (Senn 1966:450, 480, 488) and Russian (Hirt 1937:50) use finite-verb iteration in iterative or continuative function. Perhaps this is an innovation due to convergence with Uralic.

[18] Similar structures are found in Vogul (Collinder 1965:336-338 passim) and Kamassian (ibid.508). Some of the other Uralic languages use similar structures with (borrowed) coordinating conjunctions.

[19] To these similarities must be added the evidence of echo formations etc. discussed in note 2 above.

TEXTUAL ABBREVIATIONS

MS = Maitrāyaṇī Saṁhitā
ŚB = Śatapatha-Brāhmaṇa
RV = Rig-Veda

REFERENCES

ABAEV, V. I. 1964. A grammatical sketch of Ossetic, transl. by S. P. Hill. The Hague: Mouton.

BARTHOLOMAE, Christian. 1904. Altiranisches Wörterbuch. Straßburg: Trübner.

-----. 1907. Review of Wackernagel 1905. Anzeiger für Indogermanische Sprach- und Altertumskunde 20.162-172.

BEYTHAN, Hermann. 1943. Praktische Grammatik der Tamilsprache. Leipzig: Harrassowitz.

BRUGMANN, Karl. 1904. Die Demonstrativpronomina der indogermanischen Sprachen. (Abhandlungen der Königl. Sächsischen Gesellschaft der Wissenschaften, philologisch-historische Klasse, 22:6.) Leipzig: Teubner.

-----. 1907. Die distributiven und kollektiven Numeralia der indogermanischen Sprachen. (Abhandlungen der Königl. Sächsischen Gesellschaft der Wissenschaften, philologisch-historische Klasse, 25:5.) Leipzig: Teubner.

COLLINDER, Björn. 1965. An introduction to the Uralic languages. Berkeley & Los Angeles: University of California Press.

COLLITZ, Hermann. 1882. Ueber eine besondere Art vedischer Composita. Verhandlungen des Fünften Internationalen Orientalisten-Congresses, 2:2.287-298.

DELBRÜCK, Bertold. 1888. Altindische Syntax. (Syntaktische Forschungen, 5.) Halle: Waisenhaus. (Repr. 1968, Darmstadt: Wissenschaftliche Buchgesellschaft.)

-----. 1900. Vergleichende Syntax der indogermanischen Sprachen, 3. (= Vol. 5 of K. Brugmann & B. Delbrück's Grundriß der vergleichenden Grammatik der indogermanischen Sprachen.) Straßburg: Trübner.

DRESSLER, Wolfgang. 1968. Ved. *divé-dive* und die idg. Iterativkomposita. Studien zur Sprachwissenschaft und Kulturkunde: Gedenkschrift für Wilhelm Brandenstein ..., ed. by M. Mayrhofer et al., 39-47. (Innsbrucker Beiträge zur Kulturwissenschaft, 14.) Innsbruck.

DUCHESNE-GUILLEMIN, Jacques. 1936. Les composés de l'avesta. Liège & Paris: Faculté de Philosophie et Lettres/Librairie E. Droz.

DUNKEL, George. 1981a. Āmreḍita and iteration of preverbs in Vedic and Hittite. Zeitschrift für vergleichende Sprachforschung 95. 214-226.

-----. 1981b. Further traces of preverbal āmreḍita in Greek and Latin. Zeitschrift für vergleichende Sprachforschung 95.226-231.

EMENEAU, Murray B. 1954. Linguistic prehistory of India. Proceedings of the American Philosophical Society 98.282-292.

-----. 1956. India as a linguistic area. Language 32.3-16.

-----. 1962. Bilingualism and structural borrowing. Proceedings of the American Philosophical Society 106.430-442.

-----. 1980. Language and linguistic area: Essays selected by A. S. Dil. Stanford, CA: University Press.

GRIERSON, J. H. 1921. Linguistic survey of India, vol. 1:1. Calcutta. Repr. 1968, Delhi: Motilal Banarsidass.

HESTON, Wilma L. 1980. Some areal features: Indian or Indo-Iranian? International Journal of Dravidian Linguistics 10.180-187.

HIRT, Hermann. 1937. Indogermanische Grammatik, 7. Heidelberg: Winter.

HOCK, Hans Henrich. 1975. Substratum influence on (Rig-Vedic) Sanskrit? Studies in the Linguistic Sciences 5:2.76-125.

-----. 1982. Clitic verbs in PIE or discourse-based verb fronting? Sanskrit *sá hovāca gā́rgyaḥ* and congeners in Avestan and Homeric Greek. Studies in the Linguistic Sciences 12:2.1-38.

-----. 1984. (Pre-)Rig-Vedic convergence of Indo-Aryan with Dravidian? Another look at the evidence. Studies in the Linguistic Sciences 14:1.89-107.

-----. 1985. Yes, Virginia, syntactic reconstruction is possible. Studies in the Linguistic Sciences 15:1.49-60.

-----. 1986/1991. Principles of historical linguistics. 2nd rev. edition, 1991. Berlin: Mouton de Gruyter.

-----. 1988a. Historical implications of a dialectological approach to convergence. Historical dialectology, ed. by J. Fisiak, 283-328. Berlin: Mouton de Gruyter.

-----. 1988b. Review article: Finiteness in Dravidian. (Steever 1988.) Studies in the Linguistic Sciences 18:2.211-31.

-----. 1989. Conjoined we stand: Theoretical implications of Sanskrit relative clauses. Studies in the Linguistic Sciences 19:1.93-126.
-----. 1992. Reconstruction and syntactic typology: A plea for a different approach. Explanation in historical linguistics, ed. by G. W. Davis & G. K. Iverson, 105-121. Amsterdam & Philadelphia: Benjamins.
-----. In Press. Chronology or genre? Problems in Vedic syntax. Proceedings of the International Vedic Workshop, Harvard University, June 1989, ed. by M. Witzel.
HOFFMANN, Karl. 1960. Der vedische Typus *menāmenam*. Zeitschrift für vergleichende Sprachforschung 76.242-248. (Repr. 1975, in Karl Hoffmann, Aufsätze zur Indoiranistik, ed. by J. Narten, Wiesbaden: Reichert.)
HOFMANN, J. B. 1951. Lateinische Umgangssprache. 3rd ed. Heidelberg: Winter.
-----. 1965. Lateinische Syntax and Stylistik, rev. by Z. Szantyr. München: Beck.
KRAUSE, Wolfgang, & Werner Thomas. 1960. Tocharisches Elementarbuch, 1. Heidelberg: Winter.
KUIPER, F. B. J. 1967a. The genesis of a linguistic area. Indo-Iranian Journal 10.81-102. (Repr. in International Journal of Dravidian Linguistics 3.135-153 (1974).)
-----. 1967b. The Sanskrit nom. sing. *viṭ*. Indo-Iranian Journal 10.103-125.
-----. 1991. Aryans in the Rigveda. (Leiden Studies in Indo-European.) Amsterdam/Atlanta, GA: Rodopi.
LEWIS, G. L. 1953. (Teach yourself) Turkish. Kent: Hodder and Stoughton.
-----. 1988. Turkish grammar. Revised edition. Oxford/New York: Oxford University Press.
MACDONELL, A. 1916. Vedic grammar for students. London: Oxford University Press. (Repr. 1955, Bombay, Calcutta, Madras: Oxford University Press.)
MASICA, Colin P. 1976. Defining a linguistic area: South Asia. Chicago: University of Chicago Press.
MAYRHOFER, Manfred. 1992. Etymologisches Wörterbuch des Altindoarischen, 2:11. Heidelberg: Winter.
MEID, Wolfgang. 1968. Zum Dvandva-Kompositum im Irischen. Studien zur Sprachwissenschaft und Kulturkunde: Gedenkschrift für Wilhelm Brandenstein ..., ed. by M. Mayrhofer et al., 107-108. (Innsbrucker Beiträge zur Kulturwissenschaft, 14.) Innsbruck.
MENGES, Hermann. 1953. Repetitorium der lateinischen Syntax und Stilistik, rev. by A. Thierfelder. Repr. 1973, Darmstadt: Wissenschaftliche Buchgesellschaft.
RAMANUJAN, A. K., & Colin Masica. 1969. Toward a phonological typology of the Indian linguistic area. Current trends in linguistics, ed. by T. Sebeok, 5.543-577. The Hague: Mouton.

SCHÄUFELE, Steven. 1990. Free word order syntax: The challenge from Vedic Sanskrit to contemporary formal syntactic theory. Urbana: University of Illinois Ph.D. dissertation in Linguistics.

SCHWYZER, Eduard. 1939. Griechische Grammatik, 1. München: Beck.

-----. 1949. Griechische Grammatik, 2, completed by A. Debrunner. München: Beck.

SENN, Alfred. 1966. Handbuch der litauischen Sprache, 1. Heidelberg: Winter.

STEEVER, Sanford B. 1988. The serial verb formation in the Dravidian languages. (MLBD Series in Linguistics, 4.) Delhi: Motilal Banarsidass.

THOMASON, Sarah Grey, & Terrence Kaufman. 1988. Language contact, creolization, and genetic linguistics. Berkeley & Los Angeles: University of California Press.

THORDARSON, Fridrik. 1989. Ossetic. Compendium linguarum iranicarum, ed. by R. Schmitt, 456-479. Wiesbaden: Reichert.

TIKKANEN, Bertil. 1987. The Sanskrit gerund: A synchronic, diachronic, and typological analysis. (Studia Orientalia, 62.) Helsinki: Finnish Oriental Society.

WACKERNAGEL, Jacob. 1905. Altindische Grammatik, 2:1. Göttingen: Vandenhoeck & Ruprecht.

-----. 1929-30. Altindische Grammatik, 3. Göttingen: Vandenhoeck & Ruprecht.

-----, & Albert Debrunner. 1957. Altindische Grammatik, 2:1. [Second edition of Wackernagel 1905.] Göttingen: Vandenhoeck & Ruprecht.

Studies in the Linguistic Sciences
Volume 23, Number 1, Spring 1993

REVIEW

Narindar K. Aggarwal. **Studies on Nepali language and linguistics: A bibliography. (Subject Bibliography Series, 15.)** Gurgaon (India): Indian Documentation Service, 1991; pp. xiv, 93.

Mithilesh K. Mishra

Studies on Nepali language and linguistics (*SNLL*) is undoubtedly the most comprehensive and useful collection of resource materials (both primary and secondary) on Nepali or Gurkhali and Newari that has been published to this date. The author has rendered the researchers working in various areas related to Nepali and Newari an extremely valuable service by compiling at one place various source materials on these languages that remain scattered over the three continents, especially in the U.S.A., South Asia (Nepal and India), and the U.K. *SNLL* has 726 unannotated entries which include books, doctoral dissertations, master theses, and published and unpublished papers on various aspects of Nepali and Newari languages and linguistics (written/published in English, Hindi, Nepali and Newari). The entries are classified under eleven headings: (1) Bibliographies, dictionaries and glossaries (pp.1-8), (2) Studies on Nepali and Newari (pp. 9-13), (3) Nepali and Newari grammars (Traditional) (pp.14-22), (4) Descriptions of Nepali (syntax/semantics and morphology) (pp. 23-29), (5) Descriptions of Newari (syntax/semantics and morphology) (pp. 30-35), (6) Phonetics and phonology (pp. 36-41), (7) Lexicography and lexicology (pp. 42-47), (8) Applied Nepali linguistics: stylistics, socio-linguistics (pp. 48-51), (9) Pedagogical material (pp. 52-58), (10) Historical, comparative and dialectological studies (pp. 59-76), and (11) Nepali and language planning (pp. 77-79). Also included is an author index that provides cross references to the entries. Aggarwal's categorization of various entries is very systematic, exhaustive, and useful. Especially for teachers and students of Nepali and Newari languages, this monograph brings to light a rich source of pedagogical material both in English and in Nepali and Newari.

The bibliography is highly commendable for a variety of reasons. First of all, by publishing this monograph, the author has rectified an overly outdated practice of putting together widely divergent pieces of bibliographic information under one title such as 'A Bibliography of Nepal' (Scarecrow Press, 1973), or 'Bibliography of Nepal' (Royal Nepal Akademy, 1975). In the present age of information the need for a comprehensive and yet very focussed and user-oriented body of information cannot be overemphasized. *SNLL* meets these criteria of an excellent bibliography in exemplary fashion. Secondly, this bibliography, by consolidating information on scholarly work on Nepali and Newari languages and linguistics, has established an

academic bridge between the traditional scholars of Nepal and India and the western scholars in general. Third, the author of *SNLL*, by mapping out the entire field of scholarship on Nepali and Newari languages and linguistics, not only reminds us of the existence of this potentially fruitful but utterly neglected area of research on both sides of the Atlantic, but also challenges us to explore the entire field anew.

Because of its comprehensive coverage of different types of resource materials, *SNLL* is definitely going to be a part of the 'tool-kit' of every brand of linguist working on Nepali and Newari languages in particular and of scholars of Nepali and Newari in general.

Studies in the Linguistic Sciences
Volume 23, Number 1, Spring 1993

REVIEW

John Baldwin & Peter French. **Forensic phonetics.** London & New York: Pinter, 1990; pp. viii, 141. $47.50.

José Ignacio Hualde

Even though this book has two authors, it is not a coauthored book in the sense that the two authors claim joint responsibility for the text. Rather, we are told in the Foreword that chapter 3 was written by French, and all other chapters by Baldwin. The reader who skips the Foreword will experience some puzzlement since the book is not only written in first person singular, but the two authors disagree as to the preferable way to practice forensic phonetics.

The book is intended for the interested layman. It contains six chapters. In chapter 1, 'Introduction', and chapter 4, 'Aspects of forensic phonetics', we are given information on the work of forensic phoneticians, and, in particular, forensic phoneticians working within the English judiciary system.

Forensic phoneticians are those professionals whose job it is to serve as experts in legal cases in the identification of speakers by analyzing and comparing samples of their speech. Typically, a forensic phonetician is consulted when the police are in the possession of an incriminating recording and either have a suspect whose speech they want to have compared to that on tape, or want to obtain information on the features of the taped speech that would allow them to narrow the range of possible suspects. A forensic phonetician may also be consulted by the defense in order to disprove that a recorded voice is that of the defendant.

Chapter 2, 'Phonetics', provides a very brief and rather superficial overview of this field, but intercalating a number of anecdotes which make it quite readable. Chapter 3, 'Acoustic phonetics', deals with those aspects of speaker identification where instrumental analysis offers advantages over a purely auditory examination in the opinion of the author (French). The chapter includes narrations of a number of actual cases where acoustic analysis was crucial in the identification of the speaker. Precisely the main difference of opinion between the two authors is that, whereas Baldwin has faith in identification by auditory means, French favors an acoustic/auditory methodology which combines identification by ear with instrumental analysis.

Chapter 5, 'Cases', is the most entertaining one of the book. In it, Baldwin relates some of the most interesting cases in which he was consulted as an expert. As a linguist, the case which I found of greatest interest is one whose complicating factor was that it involved accent-switching on the part of the suspect. Baldwin distin-

guishes accent-switching from the behavior of someone who puts on an accent different from his or her own in order to deceive. Accent-switchers are in command of two accents, both of them equally real and natural to them, and can voluntarily switch from one accent to the other. These are, thus, bilingual speakers whose two linguistic codes happen to differ solely or mostly on phonetic characteristics. The subject in the case was a native of Northern Ireland who had full command of two varieties of English spoken in that region, which are referred to as Scots Irish and Irish Irish. Baldwin offers the following comparison of the two varieties:

> Some salient features of [Scots Irish], which distinguish it from Irish Irish, are the widespread use of the glottal stop, even to reinforce voiced consonants as in 'Radley' as opposed to its non-occurrence; the use of a back rounded diphthong in words like 'cold', 'go', etc., as opposed to the Irish Irish front rounded diphthong; the correct pronunciation of dental fricatives, as opposed to their realization as dental stops, and so on. [p.105]

Apparently the subject had total control over these features and could switch from one accent to the other depending on who the interlocutor was.

In Chapter 6, 'The future', Baldwin discusses two topics. One is the feasibility of computer-generated 'voice-prints' in a not too distant future. He believes the technology will be available. The second topic is some reforms that he would like to see implemented in the English judiciary system. Baldwin has a rather negative opinion of the adversarial system, as it is practiced in England, which opinion he voices in several places throughout the book. His strongest complaint is that many people involved in such a legal system are purely interested in winning, and not in determining the truth.

Forensic phoneticians who, like other experts, work for one of the two parties in the adversarial system (the defense or the prosecution), are not always able to present all the evidence that they have laboriously obtained and which could be crucial for finding the truth in the case. For instance, if they are working for the defense and their results are against the interest of the defendant, they will be suppressed. Baldwin firmly believes that forensic phoneticians should be impartial experts and proposes that experts should work for the court, not for one of the two parties.

The book is very readable and offers interesting insights in the work of phoneticians working in the judiciary system. Although, its interest for academic linguists is only tangential, it could be a very interesting book for students considering careers involving phonetics.

Studies in the Linguistics Sciences
Volume 23, Number 1, Spring 1993

RECENT BOOKS

Studies in the Linguistic Sciences gratefully accepts review copies of recent publications and tries to find reviewers for them. In this endeavor, however, it does not always succeed. Volumes for which no reviewers have been found so far are publicized in this section, with brief indications of contents or interest. Prices are indicated where known. (Unless otherwise indicated, the brief descriptions below are by the editor.)

Carol Lord. **Historical change in serial verb constructions**. (Typological Studies in Language, 26.) Amsterdam/Philadelphia: John Benjamins Publishing Company, 1993, pp. x, 273.

The term 'serial verb' has been used in a variety of different, often conflicting meanings. Lord (L) concentrates on a type of 'constructions in which one of the verbs is defective in some respect, such as phonological assimilation, failure to take the usual verb inflections or negation affixes, or showing unexpected syntactic properties, for example with respect to movement.' (3) She claims that verbs of this type exhibit strong crosslinguistic similarities both in their original meanings 'and in the functions they come to mark'. (3) L posits a 'directional' change through 'bleaching' or 'desemanticization', such as verb → preposition → case-marking affix and similar developments leading to the development of auxiliaries, adverbs, and conjunction-like elements (complementizers, quotative markers, adverbial subordinators, etc.). The major focus of the book is on West African languages. The wealth of data from these languages, much of which is based on L's own research, is supplemented by evidence of other languages, especially Caribbean creoles and Mandarin Chinese.

Chapter 1 presents a general discussion on 'Serial verbs'. Chapters 2 - 8 are devoted to specific developments: Locative verbs → prepositions (2); verbs → recipient/benefactive marking (3); comitative verbs → prepositions and conjunctions (4); verbs → object markers (5); developments of special prepositions (such as 'except' or 'more than') (6); developments of complementizers and subordinating conjunctions (7); and verbs → adverbs and auxiliaries. Chapter 9 ('Pragmatics, typology, and teleology') addresses the implications of L's findings, as well as 'The "why" question'.

Concerning the latter question, L opts for the view that grammaticalization is attributable to two factors: 'the process of phonological erosion of grammatical forms, and the human imperative to be expressive resulting in the utilization of metaphor.' (249) What is missing in this account is the element of 'bleaching' or 'desemanticization' which L appeals to in the preface. One suspects that it is this semantic process of bleaching which leads to accentual 'downgrading' (including cliticization) and that this downgrading, in turn, is responsible for the fact that these verbs undergo much more radical 'phonological erosion' than non-grammaticalized verbs.

Mortéza Mahmoudian. **Modern theories of language: The empirical challenge**. (Sound and Meaning: The Roman Jakobson Series in Linguistics and Poetics.) Durham & London: Duke University Press, 1993, pp.xvii, 231.

The goal of this book is to critically examine twentieth-century theories of language and to argue for the need to confront them with empirical evi-

dence. Mahmoudian (M) claims that differences in terminology mask principles shared by most linguists. These are the distinction between unit and system; concerns with taxonomy, rules (or other types of generalization), and explanatory adequacy; the issue of linguistic universals; the distinction between 'signifié' and 'signifiant'.

The scope of M's coverage can perhaps most effectively be gauged by examining his references and citations: The four most recent references are from 1989 and 1988. All of these and 17 out of some 24 other citations since 1980 can be broadly classified as being semiotic and/or as adhering to Martinet's functionalism. Although Chomsky is referred to most often (on nine occasions vs. seven each for Martinet and Hjemslev), his most recent publication referred to by M is the 1977 French version of his presentation of the 'Extended Standard Theory'. The most recent publications by US linguists are a 1978 book by Ebeling in the Taxonomic framework and Greenberg's 1978 series on 'Universals of human language'.

P. A. Messelaar. **La confection du dictionnaire général bilingue**. Leuven: Peeters, 1990, pp. 109.

This book is intended as an aide-mémoire for the director of a lexicographical project ad as a manual for the beginning bilingual lexicographer. The introductory section discusses background issues, including the general difficulties associated with translation. The second section outlines the history of bilingual dictionary making. Sections 3 through 17 address specific issues confronting the compiler of a bilingual dictionary: The general organization of the dictionary (3); the required 'note to the reader' (4); the selection of entries (5); the nature of the lexicographical entry (6); the structure of the entry (7); the phonetic transcription (8); the typography to be employed (9); grammatical information (10); information on usage (11); semantic information and additional information helpful in translation (12); synonymy (13); semantic information required to assure translational equivalency (14); issues connected with idiomatic usage and restrictions on collocability (15); the extent to which non-lexical (primarily syntactic) information needs to be included (16); stylistic issues (17). Except for M's own work, three monolingual dictionaries, and one bilingual one, the bibliographical references predate 1980.

Kenneth L. Pike. **Talk, thought, and thing: the emic road toward conscious knowledge**. Dallas: Summer Institute of Linguistics, 1993, pp. xii, 85.

'This book is written for a small number of people unknown to me, who are disillusioned in a changing world' (vii). Pike (P) suggests that a source for hope may lie in 'looking into the nature of language', since language serves as the basic means of communication of all people (viii). Although addressing these issues as a linguist (within the Taxonomic framework), P emphasizes the central role of the 'person' and professes to 'have been helped ... by several philosophers', including Quine, Mavrodes, and Reeder (ix). He concludes that 'A person grows in knowledge through the intersection of networks of patterns of patterns of ... phonological, grammatical, and referential hierarchies' and that 'A person needs his language to help him know himself in relation to his physical, social, aesthetic, and philosophical environment' (78).

V. Prakasham ad S. V. Parasher, eds. **Linguistics at large: Papers in general and applied linguistics**. Hyderabad: Booklinks Corporation, 1991, pp. xvii, 191.

Linguistics at large is a collection of papers in honor of the sixtieth birthday of Shivendra K. Verma, Director of the Central Institute of English

and Foreign Languages (CIEFL), Hyderabad (India). In addition to the editors' preface and a list of Verma's publications, the volume contains sixteen articles, half of which are by members of the CIEFL.

Eight papers address issues directly or indirectly connected with the CIEFL's primary focus on language pedagogy, especially as regards English. Especially interesting and provocative among these is K. Annamalai's 'When an adult learns his mother tongue', an article examining the special difficulties encountered by adults who have 'lost' — or even never learned — the language of their community in trying to learn that language as adults, as part of a kind of ethnic revival.

The remaining papers address issues of the nature of 'Modern [generative] grammar' (K. A. Jayaseelan); 'Contemporary syntactic theories and traditional Indian syntax' (B. N. Patnaik); 'Length as a formative prosody in Telugu' (V. Prakasam); 'Schwa fronting in Hindi' as a variable-rule phenomenon (P. K. Pandey); 'Addressee markers on verbs' in colloquial Tamil as a problem for theories of agreement (R. Amritavalli); 'Bhartṛhari [the eminent indigenous philosopher of language] on lexical meaning' (K. Kapoor); 'The sociolinguistics of code-switching' in Hindi and English (A. Kumar); and 'Modernization of Sanskrit', mainly a study of lexical borrowing throughout the history of Sanskrit (H. S. Anantanarayana). Some of the papers have appeared elsewhere.

The volume is well produced, except for a fair number of self-correcting misprints and the more serious fact that diacritics and special symbols frequently are missing or misprinted.

F. de Saussure. **Troisième cours de lingüistique générale (1910-1911) d'après les cahiers d'Émile Constantin — Saussure's Third course of lectures on general linguistics (1910-1911): From the notebooks of Émile Constantin.** French text edited by Eisuke Komatsu; English translation by Roy Harris. Oxford, New York, Seoul, Tokyo: Pergamon Press, 1993, pp. xxiii, 173 (+143).

The notes on de Saussure's third — and last — course in general linguistics (TCLG) which his student Émile Constantin took in 1910-1911 were not available to the editors of the famous set of de Saussure's lecture notes, *Cours de lingüistique générale* (1916). Their existence became known only in 1958. (As noted in the foreword of TCLG, the editors of the *Cours* claimed to have drawn on de Saussure's third course, but the notes that they consulted were much less complete than Constantin's.)

The importance of Constantin's notes was realized almost immediately after becoming available: They seemed to offer a much more reliable and direct access to de Saussure's mature thinking on general linguistics than the recast and transformation of his lecture notes at the hands of the editors of the *Cours*, Bally and Sechehay. Especially noteworthy are a classification of speech sounds in terms of six degrees of buccal opening covering both consonants and vowels (reminiscent of classifications by the indigenous phoneticians of ancient India as well as recent work by Clements) and what in the Foreword of TCLG is referred to as de Saussure's 'most important theoretical pronouncements about *la langue*'. Among the latter are such statements as 'The language [i.e. *la langue*, as distinct from *langage*] is located only in the brain' (69a) and 'Strictly speaking there are no signs but differences between signs' (141a), as well as an account of synchronic laws as distinct from diachronic ones and remarks on the often indirect relation between the two types of laws (115a-118a).

Because of their significance, Constantin's notes were drawn on extensively by Rudolf Engler in his 1968 critical edition of the *Cours*; but the structure of that edition remained that of Bally and Sechehay's. TCLG presents the notes in their original structure and sequence and thus permits the reader to judge them in their own right. A Foreword, a note on 'The Constantin Notebooks', and a 'Translator's Preface' are followed by a bilingual presentation of the text, with an edited version of the French original on the left (pp. 1-143) and the English translation of that version on the right (pp. 1a-143a). The book concludes with a 'Selective index of French terminology' (145-173).

CONTENTS OF *STUDIES IN THE LINGUISTICS SCIENCES*, VOLUMES 18 - 22

Volume 18:1 (Spring 1988)
Papers in General Linguistics
(Edited by Michael J. Kenstowicz)

Carreira, Maria. The representation of diphthongs in Spanish (1-24)

Downing, Laura J. Tonology of noun-modifier phrases in Jita (25-60)

Irshied, Omar, & Peter Whelan. Exploring the dictionary: On teaching foreign learners of Arabic to use the Arabic-English dictionary (61-75)

Kenstowicz, Michael J., Emmanuel Nikiema, & Meterwa Ourso. Tonal polarity in two Gur languages (77-103)

Moshi, Lioba. A functional typology of *ni* in Kivunjo (Chaga) (105-134)

Patterson, Trudi A. Some morphological and phonological interactions in Lakhota (135-149)

Wong-opasi, Uthaiwan. On deriving specifiers in Spanish: Morpho-phono-syntax interactions (151-177)

Volume 18:2 (Fall 1988)
Papers in General Linguistics
(Edited by Hans Henrich Hock)

Abu-Salim, Issam M., & Hassan R. Abd-el-Jawad. Syllable patterns in Levantine Arabic (1-22)

Cervin, Richard. On the notion of 'second position' in Greek (23-39)

Gerdemann, Dale, & Erhard W. Hinrichs. UNICORN: A unification parser for attribute-value grammars (41-86)

Kamwangamalu, Nkonko Mudipanu. 'C-command' and the phonology-syntax interface in Ciluba (87-109)

Ourso, Meterwa A. Root control, underspecification, and ATR harmony (111-127)

Schäufele, Steven. Where's my NP? Non-transformational analyses of Vedic pronominal fronting (129-162)

Tsiang, Sarah. The discourse function of the absolutive in the Pañcatantra (163-181)

Zhou, Xinping. On the head movement constraint (183-210)

Hock, Hans Henrich. Finiteness in Dravidian (Review article): Sanford B. Steever (1988), *The serial verb formation in the Dravidian languages* (211-233)

Volume 19:1 (Spring 1989)
Papers in General Linguistics
(Edited by Hans Henrich Hock)

Arora, Harbir, & K. V. Subbarao. Convergence and syntactic reanalysis: The case of *so* in Dakkhini (1-18)

Bundrick, Camille. An inference-based account of restrictive relative *which* and *that* (19-31)

Cassimjee, Farida, & Charles W. Kisseberth. Shingazidja nominal accent (33-61)

Chung, Raung-fu. On the representation of Kejia diphthongs (63-80)

Gerdemann, Dale. Restriction as a means of optimizing unification parsing (81-92)

Hock, Hans Henrich. Conjoined we stand: Theoretical implications of Sanskrit relative structures (93-126)

Kachru, Braj B. World Englishes and applied linguistics (127-152)

Kachru, Yamuna. Corpus planning for modernization: Sanskritization and Englishization of Hindi (153-164)
Bhatt, Rakesh Mohan. Good mixes and odd mixes: Implications for the bilingual's grammar. (Squib) (165-168)
Aitchison, Jean. Review of Rama Kant Agnihotri (1987), *Crisis of identity: Sikhs in England* (169-171)
Pandharipande, Rajeshwari. Review of Tej K. Bhatia (1987), *A history of the Hindi grammatical tradition* (173-179)
Markee, Numa. Review of Dick Chamberlain & Robert Baumgardner (eds.) (1988), *ESP in the classroom: Practice and evaluation* (181-185)
Zgusta, Ladislav. Review of Sumitra Katre (1987), *Aṣṭādhyāyī of Pāṇini* (187-193)

Volume 19:2 (Fall 1989)

The contribution of African linguistics to linguistic theory, Vol. 1
(Edited by Eyamba G. Bokamba, Associate Editors: Rick Treece and Dorothy E. Evans)

Introduction(vii-xi)
Clements, G(eorge) N. African linguistics and its contributions to linguistic theory (3-39)
Biloa, Edmond. Tuki gaps: Null resumptive pronouns or variables? (43-54)
Childs, G. Tucker. Where do ideophones come from? (55-76)
Ottenheimer, Harriet, & Heather Primrose. Current research on ShiNzwani ideophones (77-87)
Downing, Laura J. Tone in Jita questions (91-113)
Hyman, Larry M. Accent in Bantu: An appraisal (115-134)
Timmons, Claude, & Christian Dunn. La sélection morphophonologique des classes en kpokolo (135-151)
Ali, Mohammed. Trends in Oromo lexicon and lexicography (155-168)
Botne, Robert. Reconstruction of a grammaticalized auxiliary in Bantu (169-186)
Clamons, Cynthia Robb. Modification of the gender system in the Wollegan dialect of Oromo (187-195)
Appendices to the Proceedings
A. History of the Annual Conference on African Linguistics (199-201)
B. Research and publications in African linguistics by students, alumni, and faculty of the University of Illinois (203-225)

Volume 20:1 (Spring 1990)

The contribution of African linguistics to linguistic theory, Vol. 2
(Edited by Eyamba G. Bokamba, Associate Editors: Amy C. Cheatham, Dorothy E. Evans, and Rick Treece)

Preface (v-vi)
Introduction (vii-ix)
Bokamba, Eyamba G. African languages and sociolinguistic theories (3-34)
Bresnan, Joan. African languages and syntactic theories (35-48)
Goldsmith, John. Phonological theory and African language phonology (49-62)
Mufwene, Salikoko. African languages, African linguistics, and linguistic theory: A commentary on the plenary session papers (63-67)
Broselow, Ellen, & Alice Niyondagara. Feature geometry of Kirundi palatalization (71-88)

Clements, G(eorge) N., & Remi Sonaiya. Underlying feature representation in Yoruba (89-103)
Ka, Omar. Reduplication and prosodic constituents in Wolof (105-121)
Noske, Manuela. Vowel Harmony in Turkana (123-134)
Ourso, Meterwa A., & Charles H. Ulrich. Sonorant-strengthening in Lama (135-147)
Kisseberth, Charles, & Sheila Mmusi. The tonology of the object prefix in Setswana (151-161)
Mutaka, Ngessimo. The tone bearing unit in Kinande (163-172)
Kapanga, Andre Mwamba. Language variation and language attitudes: A case study from Shaba Swahili (175-188)
Wade-Lewis, Margaret. The contribution of Lorenzo Dow Turner to African linguistics (189-204)

Volume 20:2 (Fall 1990)

Linguistics for the Nineties: Papers from a lecture series in celebration of the Department's twenty-fifth anniversary

(Editor: Hans Henrich Hock, Editorial Assistant: Lynne Murphy)

Antonsen, Elmer H. Introduction (ix-xiv)
Kahane, Henry. The establishment of Linguistics at Illinois (1-2)
Langacker, Ronald W. Cognitive Grammar: The symbolic alternative (3-30)
Sadock, Jerrold M. A trimodular account of Yiddish syntax (31-50)
Newmeyer, Frederick J. Some issues in language origins and evolution (51-68)
Odden, David. Phonology and its interaction with syntax and morphology (69-108)
Menn, Lise. Aphasic language under discourse pressure: Functional syntax vs. psycho-linguistic function (109-122)
Lowenberg, Peter H. Standards and norms for World Englishes: Issues and attitudes (123-137)
Hermon, Gabriella. Syntactic theory and language acquisition: A case against parameters (139-163)
Sridhar, S. N. What are applied linguistics? (165-176)
Index to *Studies in the Linguistic Sciences*, Volumes 1 - 19
A. Author index (177-195)
B. Title index (197-214)

Volume 20:3 (Spring 1990)

Meeting handbook, Thirteenth South Asian Languages Analysis Roundtable [Abstracts]

(Edited by Hans Henrich Hock, Editorial Assistant: Lynne Murphy)

Ahmed, Mariam. Convent English: Structure and attitudes (17)
Anderson, Lloyd B. Script Manager software for Indic scripts on the Macintosh (18)
Bagchi, Tista. Conditionals and emphasizers in Bangla: Some pragmatic effects of their interaction (19-20)
Bains, Gurprit. Focus movement in Hindi-Urdu (21)
Bhatt, Rakesh M. An essay on Kashmiri stress (22-23)
Boolchandani, Pushpa. On binding reflexives in Sindhi (24-25)
Bubenik, Vit, & C. Paranjpe. Some observations on the development of West-Indo-Aryan pronominal systems from Apabhraṁśa (26)
Butt, Miriam J., & Tracy Holloway King. Semantic case in Urdu (27-28)

Chakraborty, Jayshree. Perfectivity and the resultative state in Hindi (29-30)
Chandrasekhar, S., & S. N. Sridhar. Case markers and prepositions in Kannada (31)
Cole, Jennifer. Alliteration in Sindhi poetry: Evidence for phonological structure (32-33)
Davison, Alice. Finiteness and case in Hindi-Urdu complements (34-35)
Deshpande, Madhav. Sociolinguistic parameters of Pāṇini's Sanskrit (36)
Genetti, Carol. On the loss of gender distinctions in Nepali (37-38)
Gnanam, M. Religious cum linguistic problems in modern India (39)
Hamp, Eric P. The sources of a passive (40-41)
Herring, Susan. From aspect to tense in Old Tamil: Evidence from narrative discourse (42-43)
Hock, Hans Henrich. Syntax or Phonological Form? Reconsidering some allegedly syntactic phenomena of Vedic Sanskrit (44-45)
Hook, Peter E., & Omkar Nath Koul. Kashmiri causals: Evidence for a transformational approach (46-47)
Jamison, Stephanie W. Demonstratives with non-third persons in Vedic Sanskrit (48)
Jayasuriya, Wilfrid. The web of the spider: Language and power in Sri Lanka (49)
Joseph, Brian. Sibilant confusion in early Indic: Sanskrit prādúr (50-51)
Kamwangamalu, Nkonko M. Advancement in some Asian and African languages (52)
-----. Multilingualism and social identity: The case of Singapore (53)
Kapoor, Kapil. Analogy as argument in Ādi Śaṅkara (54)
Kissock, Madelyn J. Reflexive pronouns in Vedic (55)
Loud, John A. Issues in translating the Puranas (56)
Mahajan, Anoop. Against wh-movement in Hindi (57-58)
Mahajan, Gyanam. Sanskrit reduplication: A templatic approach (59-60)
Marlow, Patrick E. Meet me in the Bazaar: A historical perspective on the origin of a North Indian koine (61)
Mehrotra, Raja Ram. Sociolinguistics of verbal abuse in Hindi (62)
Menon, A. G. Tamil verb formation (63-64)
Moag, Rodney F. The associative case in Malayalam: Making sense of a catch-all category (65-66)
Mohanty, Gopabandhu. Compound verbs in Oriya (67-68)
Nadahalli, Jayashree. Pronouns in Kannada: Sociolinguistic implications (69)
Nadkarni, Mangesh V. On liberating English to be a world language: An Indian perspective (70-71)
Nihalani, Paroo. Articulatory and acoustic properties of apical and laminal stop consonants: A cross-language study (72)
Pandit, P. N. A socio-cognitive approach to designing a self-instructional multi-media course in English communicative skills (73)
Paolillo, John C. Functional articulation: Analyzing diglossic variation (74-76)
Pelletier, Rosanne. Telugu negatives and non-capabilitives: Morphological structure and syntactic structure (77-79)
Rai, Alok. Sammelani Hindi and Malviya Hindi: Language and politics in India between 1875 and 1930 (80)
Ramchand, Gillian. The category of nominals in Bangla (81-83)
Rau, Nalini. Coordination and word order (84)
Sadanand, Kamlesh. The pure vowels of Punjabi (85)
Sadanand, Suchitra. Malayalam syllabication (86)
Satyanath, T. S. On change and variation of (l) in Kannada (87)

Scharf, Peter M. Assessing Śabara's arguments for the conclusion that a generic term denotes just a class property (88)
Schäufele, Steven. The Vedic clause-initial string and universal grammar (89)
Sharma, Krishna K. Semio-linguistic aspect of dhvani siddhānta (90-91)
Sharma, Rama Natha. Naming and expressing objects in Pāṇini (92)
Singh, Atamjit. The aesthetics of play in Punjabi folkloric tradition (93)
Singh, Mona. A situation-type analysis of compound verbs (94-95)
Sreedhar, M. V. Drastic modernization of the curricula of the teacher training courses (96-97)
Sridhar, S. N., & Mark Aronoff. A lexicalist analysis of participle compounds in Kannada (98-99)
-----, & Indira Ayyar. Aspects of the syntax of spoken Indian English (100)
Srivastav, Veneeta. Pair-list answers in Hindi indirect questions (101-103)
Subbarao, K. V., & Harbir Arora. Convergence and syntactic change: The case of the negative participles in Dakkhini (104-105)
-----, & Lalitha M. The INFL nodes in non-finite clauses in Dravidian and Tibeto-Burman languages (106)
Tickoo, Asha. New dimensions of word order freedom in verb-final languages (107)
Tsiang, Sarah. Clausal vs. non-clausal subordination in Sanskrit narratives (108-109)
Vijayakrishnan, K. G. The mental dictionary: Its role in linguistic theroy (110-111)
Winters, Clyde A. The Harappan script: The most ancient form of Dravidian (112)
Yatabe, Shûichi. Verbal compounds in Malayalam (113-114)
Zakharyin, Boris A. Ergativity in the Indo-European languages of South Asia: Diachronic and synchronic processes (115-116)
-----, & L. V. Khokhlova. The development of ergativity in Indo-European languages of Western India in the fifteenth through twentieth centuries (117-118)
Zide, Norman. A sketchy history of cliticization and verb stem noun incorporation in Munda (119)
Pandharipande, Rajeshwari. A grammar of politeness in Marathi (125)
Mishra, Mithilesh K. Towards an ethnography of politeness in Maithili (126)
Bhatia, Tej K. Directives in Panjabi and Lahanda (127)
Verma, Manindra K. Linguistic conventions of politeness in Bhojpuri and Magahi (128)
D'Souza, Jean. Recreating South Asian speech acts in English: A study in linguistic transfer (131)
Kachru, Yamuna. Speech act in the mother tongue and the other tongue (132)
Nelson, Cecil L. On creating speech acts: The creativity of Indian English writers (133-134)
Valentine, Tamara. Language and female identity in India (135)
Ahmed, Mariam. A house divided: Conflict and rivalry in two varieties of a language (139)
Bhatia, Tej K. Transplanted languages and ethnic identity (140)
Sridhar, Kamal K. Language minorities: Issues of identity in a global perspective (141)
Bhatt, Rakesh M. Identity, conflict, and convergence: South Asia as a sociolinguistic area (142-143)
Pandharipande, Rajeshwari. The question of defining the language of religion (147)
Hock, Hans Henrich. *Vaṣaṭ, śrauṣaṭ*, and other ritual particles: Their origin and use in Vedic ritualistic literature (148-149)
Mishra, Mithilesh K. The role of deixis in defining ordinary vs. religious language (150)
Anushivarani, Ali. Rabindranath Tagore's Nobel Prize: What does it mean? (153)

Wu, Yongan. Chinese responses to Tagore: Pin Hsin's poetry (154)
Tikku, Girdhari. Aldous Huxley's *The island* (155)
Harada, Hiroko. Coleridge and Bashō: The legacy of Indian monism (156)

Volume 21:1 (Spring 1991)
Papers in General Linguistics
(Edited by Hans Henrich Hock, Editorial Assistant: Amy C. Cheatham)

Branstine, Zoann. Stop/spirant alternations in Spanish: On the representation of contrast (1-22)
Darzi, Ali. Compensatory lengthening in modern colloquial Tehrani Farsi (23-37)
Hancin, Barbara J. On the phonology-morphology interaction in Brazilian Portuguese vowel harmony (39-54)
Kang, Seok Keun. Moraic phonology and /l/-irregular predicates in Korean (55-66)
Kang, Yongsoon. Coronal: Transparent or opaque? (67-79)
Lee, Han-Gyu. Plural marker copying in Korean (81-105)
Markee, Numa. Toward an integrated approach to language planning (107-123)
Mtenje, Al. On Autosegmental feature-spreading in phonology: Evidence from Chiyao (125-145)
Pandey, Pramod Kumar. Schwa fronting in Hindi (147-159)
Hock, Hans Henrich. Review of R. N. Aralikatti (1989), *Spoken Sanskrit in India: A study of sentence patterns* (161-165)
Yoon, James H. Review of Mark R. Baltin & Anthony S. Kroch, eds. (1989), *Alternative conceptions of phrase structure* (167-178)
Conefrey, Theresa. Review of Deborah Tannen (1990), *You just don't understand* (179-181)
Recent Books (183-188)

Volume 21:2 (Fall 1991)
Illinois studies in Korean linguistics, 2
(Edited by Chin-W. Kim, Jerry L. Morgan, and James H-S. Yoon,
Editorial Assistant: Amy C. Cheatham)

Preface (v)
List of dissertations on Korean between 1987-92 at the University of Illinois at Urbana-Champaign (vi)
Ahn, Sang-Cheol. Vowel deletion and epenthesis: The vowel *ï* (1-18)
Cho, Euiyon. Notes on some tests for subjecthood in Korean (19-29)
Cho, Seikyung. The acquisition of English reflexives by Korean ESL learners (31-67)
Choi, Yeon Hee. Discourse reference in written Korean folk tales (69-87)
Kang, Seok Keun. Moraic representation of ambisyllabicity: Evidence from Korean (89-100)
Kang, Yongsoon. The Locality Condition of tonal systems: With special reference to North Kyungsang dialect in Korean (101-112)
Kim, Chin W., & Hyoung-Youb Kim. The *character* of Korean glides (113-125)
Kim, Hyoung-Youb. Prosodic phonology of Korean (127-142)
Lee, Han-gyu. The pragmatics of the pragmatic morpheme *com* 'a little' in Korean (143-166)
McClanahan, Virginia K. The pragmatics of negation in Korean (167-178)
Yoon, James Hye Suk. Inflectional structures in Korean and headedness (179-192)

Volume 22:1 (Spring 1992)

Papers in General Linguistics

(Edited by Hans Henrich Hock, Editorial Assistant: Amy C. Cheatham)

Preface (v)
Endangered languages: An appeal for publications (vii-viii)
Alho, Irja H. Distinguishing kind and set in Finnish (1-16)
Bhatt, Rakesh Mohan. Language identity, conflict, and convergence in South Asia (17-37)
Hock, Hans Henrich. What's a nice word like you doing in a place like this? Syntax vs. phonological form (39-87)
Kraska, Iwona. From verb to clitic and nominal suffix: The Somali *-e*, *-o* nouns (89-106)
Kuo, Pinmin. On the use and function of Chinese *keshi*: An explanation based on the notion 'inference system' (107-122)
Mmusi, Sheila Onkaetse. OCP violations in Setswana: Evidence for redefining the OCP? (123-142)
Prieto, Pilar. Truncation processes in Spanish (143-158)
Hualde, José Ignacio, & Gorka Elordietta. On the lexical/post-lexical distinction: Vowel assimilation in Lekeitio Basque (Squib) (159-164)
Iwasaki, Yasufumi. Review of Nanette Twine (1991), *Language and the modern state: The reform of written Japanese* (165-168)
Murphy, M. Lynne. Review of Marina Yaguello (1991), *Lunatic lovers of language: Imaginary languages and their inventors* (169-172)
Book notices
- Goldlap, Christel (1991). Lokale Relationen im Yukatekischen: Eine onomasiologische Studie (Hans Henrich Hock) (173-174)
- Kachru, Branj B. (1992). The other tongue: English across cultures (Hans Henrich Hock) (174-176)

Recent Books (177-182)
Contents of volumes 17-21 (183-189)

Volume 22:2 (Fall 1992)

Twenty-five years of linguistic research and teaching at the University of Illinois at Urbana-Champaign

(Edited by Braj B. Kachru, with the assistance of Amy C. Cheatham and Frances Vavrus)

Antonsen, Elmer H. Foreword (vii-xii)
Preface (xiii-xiv)
Acknowledgements (xv)
Part I: Perspectives on linguistics in the Midwest and at Illinois
- Kahane, Henry, & Braj B. Kachru. Introduction (3-5)
- Kachru, Braj B. Linguistics in the Midwestern region: Beginnings to 1973 (7-35)
- Kahane, Henry. History of the Department of Linguistics at the University of Illinois at Urbana-Champaign (37-39)
- Kahane, Henry. The European emigree (40-41)
- Osgood, Charles E. The tale of an eager then lonely then contented dinosaur (42-58)
- Lees, Robert B. How to find the right tree to bark up (59-64)
- Kachru, Braj B. Three linguistic reincarnations of a Kashmiri Pandit (65-74)

Kisseberth, Charles W. A sense of perspective (75-80)
Part II: Memorial tributes to a builder: Henry R. Kahane
Introduction (85)
Antonsen, Elmer H. (87-89)
Garner, Roberta Kahane (89-91)
Kahane, Charles (91-92)
Weir, Morton W. (92-94)
Faulkner, Larry R. (94-95)
Zgusta, Ladislav (95-97)
Kachru, Braj B. (97-100)
Part III: Graduate student research 1964-1992
Introduction (103)
Ph.D. dissertation abstracts (105-237)
Master's thesis abstracts (239-257)
Research in progress up to August 1992 (259-261)
Author index (263-265)
Language index (267-271)
Regional index (273-276)
Area of concentration (277-279)
Index of advisors (281-283)

All checks and money orders must be in U.S. Dollars, drawn on a U.S. Bank, and made payable to University of Illinois. Sales Tax must be included as follows (except tax exempt organizations): **IL** 7¼%, **IN** 5%, **MI** 4%, **MN** 6%, **OH** 5%, **WI** 5%. Please include **$2.00 postage** per book.

While quantities last, Volumes 7:2 through 17:2 are being offered at half-price.

Vol. 7:2	Fall 1977	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 8:1	Spring 1978	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 8:2	Fall 1978	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 9:1	Spring 1979	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 9:2	Fall 1979	Relational Grammar and Semantics (Editor: Jerry L. Morgan)	$2.50	~~$5.00~~
Vol. 10:1	Spring 1980	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 10:2	Fall 1980	Studies in Arabic Linguistics (Editor: Michael J. Kenstowicz)	$2.50	~~$5.00~~
Vol. 11:1	Spring 1981	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 11:2	Fall 1981	Dimensions of South Asian Linguistics (Editor: Yamuna Kachru)	$2.50	~~$5.00~~
Vol. 12:1	Spring 1982	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 12:2	Fall 1982	Papers on Diachronic Syntax: Six Case Studies (Editor: Hans Henrich Hock)	$2.50	~~$5.00~~
Vol. 13:1	Spring 1983	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 13:2	Fall 1983	Studies in Language Variation: Nonwestern Case Studies (Editor: Braj B. Kachru)	$2.50	~~$5.00~~
Vol. 14:1	Spring 1984	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 14:2	Fall 1984	Language in African Culture and Society (Editor: Eyamba G. Bokamba)	$2.50	~~$5.00~~
Vol. 15:1	Spring 1985	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 15:2	Fall 1985	Linguistic Studies in Memory of Theodore M. Lightner (Editor: Michael J. Kenstowicz)	$2.50	~~$5.00~~
Vol. 16:1	Spring 1986	Papers in General Linguistics	$3.00	~~$6.00~~
Vol. 16:2	Fall 1986	Illinois Studies in Korean Linguistics (Editor: Chin-W. Kim)	$3.00	~~$6.00~~
Vol. 17:1	Spring 1987	Papers from the 1986 South Asian Languages Analysis Roundtable (Editor: Hans Henrich Hock)	$3.00	~~$6.00~~
Vol. 17:2	Fall 1987	Papers in General Linguistics	$3.00	~~$6.00~~

STUDIES IN THE LINGUISTIC SCIENCES

The following issues are available

(For contents of volumes 18 - 22, see pp. 201-208)

Vol. 18:1	Spring 1988	Papers in General Linguistics	$6.00
Vol. 18:2	Fall 1988	Papers in General Linguistics	$6.00
Vol. 19:1	Spring 1989	Papers in General Linguistics	$7.50
Vol. 19:2	Fall 1989	The Contribution of African Linguistics to Linguistic Theory, Vol. 1 (Editor: Eyamba G. Bokamba)	$7.50
Vol. 20:1	Spring 1990	The Contribution of African Linguistics to Linguistic Theory, Vol. 2 (Editor: Eyamba G. Bokamba)	$7.50
Vol. 20:2	Fall 1990	Linguistics for the Nineties: Papers from a Lecture Series in Celebration of the Department's Twenty-Fifth Anniversary (Editor: Hans Henrich Hock)	$7.50
Vol. 20:3	Spring 1991	Meeting Handbook: Thirteenth South Asian Languages Analysis Roundtable, 25 - 27 May 1991, University of Illinois at Urbana-Champaign (Editor: Hans Henrich Hock)	$5.00
Vol. 21:1	Spring 1991	Papers in General Linguistics	$7.50
Vol. 21:2	Fall 1991	Illinois Studies in Korean Linguistics, 2 (Editor: Chin-W. Kim)	$7.50
Vol. 22:1	Spring 1992	Papers in General Linguistics	$7.50
Vol. 22:2	Fall 1992	Twenty-five Years of Linguistic Research and Teaching at the University of Illinois at Urbana-Champaign (Editor: Braj B. Kachru)	$10.00

FOR EARLIER ISSUES AND ORDER INFORMATION SEE INSIDE COVER

Orders should be sent to:
SLS Subscriptions, Department of Linguistics
University of Illinois
4088 Foreign Languages Building
707 S. Mathews
Urbana, Illinois 61801

Studies in The Linguistic Sciences

VOLUME 23, NUMBER 2
FALL 1993
[Published October 1996]

DEPARTMENT OF LINGUISTICS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

STUDIES IN THE LINGUISTIC SCIENCES

PUBLICATION OF THE DEPARTMENT OF LINGUISTICS IN THE
COLLEGE OF LIBERAL ARTS AND SCIENCES OF THE
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

GENERAL EDITOR: Elmer H. Antonsen
REVIEW EDITOR: James H. Yoon
EDITORIAL ASSISTANTS: SARA MICHAEL, MARK HONEGGER

EDITORIAL BOARD: Eyamba G. Bokamba, Chin-chuan Cheng, Jennifer S. Cole, Georgia M. Green, Hans Henrich Hock, Braj B. Kachru, Yamuna Kachru, Chin-W. Kim, Charles W. Kisseberth, Peter Lasersohn, Howard Maclay, Jerry L. Morgan, Rajeshwari Pandharipande, James H. Yoon, and Ladislav Zgusta.

AIM: *SLS* is intended as a forum for the presentation of the latest original research by the faculty and students of the Department of Linguistics, University of Illinois, Urbana-Champaign. Scholars outside the Department and from other institutions are also cordially invited to submit original linguistic research for consideration. In all cases, articles submitted for publication will be reviewed by a panel of at least two experts in the appropriate field to determine suitability for publication. Copyright remains with the individual authors. Authors will receive 10 offprints of their individual contributions.

SPECIAL ISSUES: Since its inception *SLS* has devoted one issue each year to restricted, specialized topics. A complete list of such special issues is given on the back cover.

BOOKS FOR REVIEW: Review copies of books may be sent to:
Editor, *Studies in the Linguistic Sciences*,
Department of Linguistics, 4088 For. Lang. Bldg.,
University of Illinois
707 S. Mathews,
Urbana, IL 61801, USA

SUBSCRIPTION: Normally, there are two issues per year. Requests for subscriptions should be addressed to *SLS* Subscriptions, Department of Linguistics, University of Illinois, 4088 Foreign Languages Building, 707 South Mathews, Urbana, Illinois 61801.

Price: $8.00 per single issue.

STUDIES IN THE LINGUISTIC SCIENCES

Papers in General Linguistics

EDITOR

Elmer H. Antonsen

EDITORIAL ASSISTANTS

Sara Michael
Mark Honegger

VOLUME 23, NUMBER 2
FALL 1993

DEPARTMENT OF LINGUISTICS, UNIVERSITY OF ILLINOIS
URBANA, ILLINOIS 61801

EDITOR'S NOTE

Although this issue bears the volume and issue designation *23:2 Fall 1993*, filling a gap in the journal's twice-a-year scheduled appearance, because of unavoidable circumstances it could not be published until October 1996. We have therefore indicated the actual publication date in square brackets where appropriate.

CORRECTION

The Editors wish to correct an unfortunate inaccuracy in the following, previously published article:

Paul Agdebor: Verb serialization in Ewe, *Studies in the Linguistic Sciences* 23:1.21-42, Spring 1993.

The designation of the voiceless bilabial fricative, <ɸ>, was apparently misinterpreted as the IPA symbol for the voiceless alveopalatal fricative [ʃ], which was then converted to the phonological equivalent /š/ producing an inaccurate phonological rendition of examples 7, 8, 9, 10, 12, 31, 34, 50, and 57. In all of these examples, /š/ should be read as /ɸ/. We greatly regret this error and apologize to the author.

In the future, authors will be given the opportunity to proof-read their contributions before they go to press.

CONTENTS

Editor's Note and Correction iv

SAE-YOUN CHO: Auxiliary verb constructions in Korean 1

CASSANDRE CRESWELL: Criticizing with a question 25

ABDUL AZIZ DIOP: Vowel deletion in Pulaar: Rime and nuclear mergers and the issue of the syntax-phonology interface 33

HANS HENRICH HOCK: Subversion or convergence? The issue of pre-Vedic retroflexion reexamined 73

JONI KAY HURLEY: Request formation in Ecuadorian Quichua 117

NKONKO M. KAMWANGAMALU: Advancement in some Asian and African languages 137

EMMANUEL KWEKU OSAM: Animacy distinctions in Akan grammar 153

UTHAIWAN WONG-OPASI: The interplay between tone, stress, and syllabification in Thai 165

MARY A. WU: Adjectival amd determinate measure phrases and NP interpretations in Mandarin Chinese 193

Studies in the Linguistic Sciences
Volume 23, Number 2, Fall 1993 [publ. October 1996]

AUXILIARY VERB CONSTRUCTIONS IN KOREAN*

Sae-Youn Cho

Korean auxiliary verb constructions have led to much controversy concerning how they can be analyzed. In analyzing this construction there have been at least two approaches: one is a syntactic approach, including the bi-clausal analysis and the VP analysis. The other is a lexical approach, including the compound verb analyses. Recently many papers on this auxiliary construction have taken either the bi-clausal analysis or the compound verb analyses without specifying why other alternatives cannot be good candidates. This paper presents the VP hypothesis to account for auxiliary constructions and argue that this analysis provides a simpler explanation of various phenomena related to this construction than the bi-clausal or the compound verb analyses.

1. Introduction

This paper provides an analysis of Korean auxiliary verb constructions[1] under the Head-Driven Phrase Structure Grammar (hereafter, HPSG) framework.[2] The data in (1) show various types of auxiliary constructions in Korean where the first one of the bold strings in the data is either a verb (e.g. *ilk-e* in (1a)) or an adjective (e.g. *yeypp-e* in (1b)), and the second one is an auxiliary verb (e.g. *po-ass-ta* in (1a)).[3]

(1) a. Mary-ka chayk-ul **ilk-e** **po-ass-ta.**[4]
-N book-A read-Comp try-P-Dec
'Mary tried to read a book.'

b. Mary-ka **yeyppu-e** **ci-ass-ta.**
-N pretty-Comp become-P-Dec
'Mary became pretty.'

c. Mary-ka **yeyppu-e** **poi-ass-ta.**
-N pretty-Comp seem-P-Dec
'Mary seemed to be pretty.'

d. Mary-ka chayk-ul **ilk-na** **po-ta.**
-N book-A read-Comp seem-Dec
'Mary seems to read a book.'

e. Mary-ka chayk-ul **ilk-eya** **ha-n-ta.**
-N book-A read-Comp must-Pres-Dec
'Mary must read a book.'

f. Mary-ka chayk-ul **ilk-ko** **iss-ta.**
-N book-A read-Comp be-Pres-Dec
'Mary is reading a book.'

There has been much controversy concerning the constructions in (1), centering on two problems. One is the question of which phrasal categories, such as VP or S, each auxiliary verb subcategorizes for. The other is how to handle the morphological requirements for the subcategorized elements by each auxiliary verb (AUX). For example, *po-ass-ta* in (1a) always requires a preceding verb with the suffix *-e*. If the preceding verb has a different suffix such as *-eya* in (1e), the sentence is ungrammatical, as in **Mary-ka chayk-ul ilk-eya po-ass-ta*. Recently many papers on Korean linguistic phenomena in HPSG, including Yoo 1993, assume that the two italicized strings in (1) are a compound verb, where the AUX is only a part of the compound verb, without specifying any reason why AUX does not constitute an independent category.

This paper will show that if AUX is assumed to be an independent category which subcategorizes for a VP and a NP (=Subject), the AUX constructions in (1) can be sufficiently explained in the HPSG framework. In addition, this analysis can deal with the morphological requirements for the subcategorized elements, such as the restrictions on the occurrence of tense suffix and the suffix form (Comp).

The arguments of this paper are organized in three main sections. In section 2, three competing hypotheses, the compound verb hypothesis, the bi-clausal hypothesis and the VP hypothesis, are presented. The primary claim of the VP hypothesis, that the AUXs in (1) are an independent category, will be motivated by arguments regarding the scope interpretation in coordination structures in section 2.1.1, the distributional properties of *kuray*+verb constructions in section 2.1.2, and verbal fronting (V^2) in section 2.1.3.[5]

Section 2.2 argues for AUX as an independent category by presenting arguments showing that AUX subcategorizes for a VP rather than a S in various environments. Section 2.2.1 demonstrates that negative polarity item (NPI) requirements do not provide evidence that AUX would subcategorize for a S. Section 2.2.2 also demonstrates that the reflexive *caki-ka* + Verb in the AUX constructions does not necessarily constitute a S, thus posing no problems for the AUX subcategorization proposal. Rather, the arguments can be used as evidence that AUX subcategorizes for a VP.

Section 3 identifies two required verbal suffixes, Comp and tense, for the subcategorized VP, which will be integrated into a formalized account of AUX subcategorization in section 4. Consequently, if an AUX subcategorizes for a VP and a NP, a unified and intuitive explanation for the AUX constructions is possible.

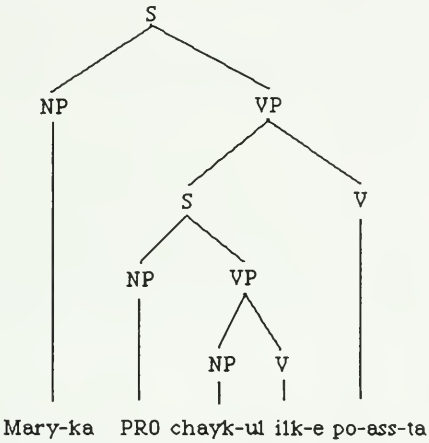
2. Constituency tests

There have been at least three analyses of the AUX constructions in (1): bi-clausal analysis by many early transformationalists, VP analysis by Park 1990 and compound verb analyses by Cho 1988, Sells 1991 and Yoo 1993.[6] For each analysis described below, there is a representation of (1a), which demonstrates that analysis's structural claims.

The bi-clausal analysis regards AUX as a category subcategorizing for a S and a NP (Subject) where the S has a trace or PRO depending on the AUXs; if the AUX is an equi-verb, the gap is a PRO but if it is a raising verb, then it is a trace. The analysis treats the verb *ilk-e* as the verb of the embedded clause [*ø chayk-ul ilk-e*], whereas the AUX *po-ass-ta* is treated as the verb of the main clause, as shown in (2).

(1a) Mary-ka chayk-ul **ilk-e** **po-ass-ta.**
-N book-A read-Comp try-P-Dec
'Mary tried to read a book.'

(2) Bi-clausal analysis



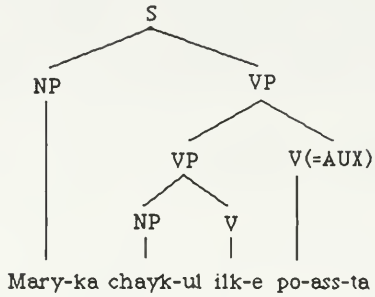
In this analysis the correct surface structure can be derived in terms of Equi-NP Deletion, since the AUX *po-ass-ta* is an equi-verb.[7] On the other hand, if the AUX in a sentence is a raising verb like *iss-ta* the surface structure can be derived by Subject-to-Subject Raising. This analysis is called bi-clausal because (1a) has two sentences where one is a main clause and the other is an embedded one, as shown in (2).

The VP analysis regards the AUX as a category subcategorizing for a VP and a NP. This analysis treats the VP *chayk-ul ilk-e* as a complement of the AUX *po-ass-ta,* so that the constructions like (1a) have no embedded sentence as shown in (3) on the next page.

Unlike the bi-clausal analysis this does not postulate on empty category in the constructions. The difference between equi and raising AUXs can be distinguished by the semantic CONTENT of AUXs.

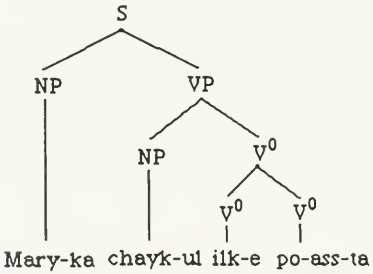
Finally, there are two different Compound analyses: Cho 1988 and Sells 1991. Sells 1991 claims that the AUX and the preceding verb syntactically form a compound verb so that the compound verb subcategorizes for two NPs to account for the example in (1a). In this analysis *ilk-e* and *po-ass-ta* are each members

(3) VP analysis



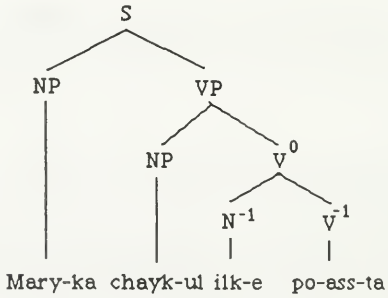
of a lexical category but syntactically form a compound verb *ilk-e po-ass-ta* as shown in (4).

(4) Complex verb analysis by Sells 1991



Similarly to Sells 1991, Cho's 1988 analysis also regards two verbs as a compound verb, where the compound verb subcategorizes for two NPs to explain the example in (1a). However, Cho claims that the suffixed element *ilk-e* is a gerundive nominal and that this gerundive nominal and the verb *po-ass-ta* morphologically form a compound verb *ilk-e-po-ass-ta* as in (5).

(5) Compound verb analysis by Cho 1988



Both compound verb analyses differently predict the possibility of the occurrence of adverbs between the AUX and the preceding verb since they have dif-

ferent compound formations. Under Sells' 1991 analysis any adverb modifying the AUX can occur between the two because each part of the compound verb is a bar-level 0 category. But Cho's 1988 analysis predicts that no adverb modifying the AUX can occur between the two because the compound verb is morphologically formed so that nothing can be placed in front of the AUX.

Despite their differences in category assignment and compound formation, they make similar claims about constituency of the AUX constructions. Therefore, they will be grouped together as the Compound Verb analysis.

The following sections will argue that the VP analysis is more plausible than the compound verb or bi-clausal analysis in accounting for the AUX constructions in (1).

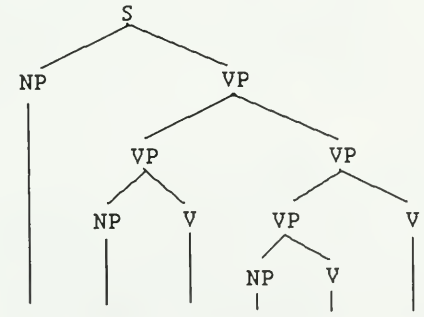
2.1. Evidence against compound verb analysis

2.1.1. Coordination and scope problems

The VP analysis provides a simpler analysis for the ambiguity of sentences with VP coordination than does the compound verb analysis. VP coordinations with AUX are possible in Korean as shown in (6). A sentence with VP coordination like (6) can have two different interpretations, (7) and (8). To have the correct readings the AUX in (7) must have scope over the VP *pap-ul mek-e* while the AUX in (8) must have scope over the whole conjoined VP *chayk-ul ilk-ko pap-ul mek-e*.

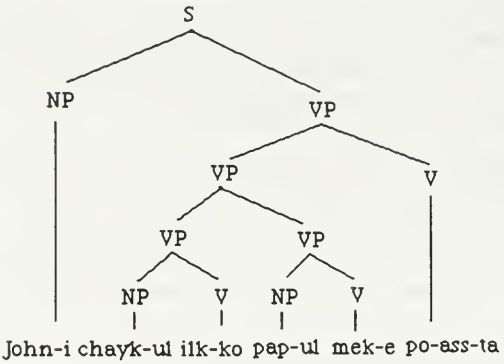
(6) John-i chayk-ul ilk-ko pap-ul mek-e po-ass-ta.
-N book-A read-and rice-A eat-Comp try-P-Dec
'John read a book and tried to have a meal.' OR
'John tried to read a book and have a meal.'

(7) John-i [[chayk-ul ilk-ko]$_{VP}$ [pap-ul mek-e po-ass-ta]$_{VP}$]$_{VP}$
-N book(s)-A read-and rice-A eat-Comp try-P-Dec
'John read a book and tried to have a meal.'



John-i chayk-ul ilk-ko pap-ul mek-e po-ass-ta

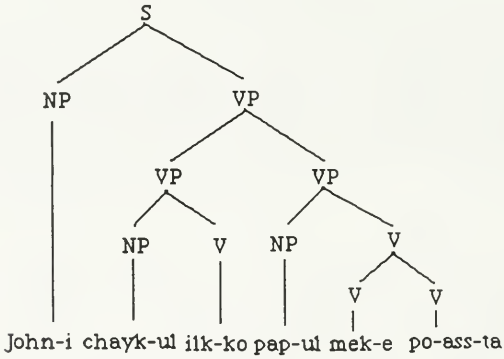
(8) John-i [[chayk-ul ilk-ko]$_{VP}$ [pap-ul- mek-e]$_{VP}$ [po-ass-ta]$_{V}$]$_{VP}$
-N book-A read-and rice-A eat-Comp try-P-Dec
'John tried to read a book and have a meal.'



If *po-ass-ta* in (6) is an AUX which subcategorizes for a VP, the structures and their interpretations like (7) and (8) can both be derived from (6). If the object of the AUX *po-ass-ta* in (6) is the VP *pap-ul mek-e*, then its interpretation must be *John read a book and tried to have a meal,* like (7), whereas if the AUX takes the whole conjoined VP *chak-ul ilk-ko pap-ul mek-e* as its object its interpretation must be *John tried to read a book and have a meal,* as in (8).

However, the compound verb analysis cannot predict that the sentence in (6) can have two interpretations, since it provides a representation as shown in (9). Under the compound verb analysis *po-ass-ta* is a part of the verb *mek-e po-ass-ta* as shown in (9), not an independent constituent.

(9)



Thus, the structure and meaning in (8) cannot be derived from (6). If the compound verb analysis is taken as correct, an additional explanation for why (7) is possible while (8) is not must be provided.[8] Such an explanation will not be needed under the VP analysis to deal with VP coordination with AUX.

2.1.2. *Kuray* substitution

The VP analysis predicts the possibility of substituting a VP for the word *kuray+suffix* in Korean while the compound verb analysis does not. An interroga-

tive like (10a) may be answered with a sentence like (10b), where a VP like *pap-ul mek-ess-ni* may be replaced with the word *kuray+suffix*. Thus the word *kuray-ess-e* in (10b) as an answer for the interrogative sentence (10a) can be used instead of the VP *pap-ul mek-ess-e*, which is like *do so* constructions in English.

(10) a. Mary-ka [pap-ul mek-ess-ni]$_{VP}$?
-N rice-A eat-P-Q
'Did Mary have a meal?'
b. ung, Mary-ka **kuray-ess-e.**
yes, -N do so-P-Dec
'Yes, Mary did so.'

Similarly, VPs with AUXs can also be replaced by the word *kuray+suffix* as shown in (11b-c), while AUXs alone cannot be replaced by it, as in (11d).

(11) a. Mary-ka [[pap-ul mek-e]$_{VP}$ cwu-ko]$_{VP}$ iss-ni?
-N rice-A eat-Comp give-a-favor-of be-Q
'Is Mary giving a favor of having a meal?'
b. ung, Mary-ka **kule-ko** iss-e.[9]
yes, -N do-so be-Dec
'Yes, Mary does so.'
c. ung, Mary-ka **kuray** cwu-ko iss-e.
yes, -N do so give-a-favor-of be-Dec
'Yes, Mary gives a favor of doing so.'
d. *ung, Mary-ka pap-ul mek-e **kule-ko** iss-e.
yes, -N rice-A eat do-so be-Dec
'Yes, Mary does so of having a meal.'

As answers for the interrogative sentence containing a raising AUX *iss-* in (11a), the sentence (11b) has the word *kule-ko* replacing the VP *pap-ul mek-e cwu-ko* in (11a) and the sentence (11c) has the word *kuray* replacing the VP *pap-ul mek-e* in (11a). On the other hand, the AUX *cwu-ko* in (11a) cannot be replaced by the word *kule* in (11d) since it is not a VP. Thus the distributional behavior of *kuray* can be accounted for if the VP analysis is chosen.

In addition, the fact that sentences (12b-c) are acceptable answers for (12a) but (12d) is not also shows that the word *kuray+ suffix* can replace only VPs, not AUXs alone.

(12) a. Mary-ka [[pap-ul mek-e]$_{VP}$ po-ass-ni]?
-N rice-A eat-Comp try-P-Q
'Did Mary try to have a meal?'
b. ung, Mary-ka **kuray** po-ass-e.
yes, -N do-so try-P-Dec
'Yes, Mary did try to do so.'
c. ung, Mary-ka **kuray-ass-e.**
yes, -N do so-P-Dec
'Yes, Mary did so.'

d. *ung, Mary-ka pap-ul **kuray** po-ass-e.
yes, -N rice-A do-so try-P-Dec
'Yes, Mary did so of having a meal.'

As answers for the interrogative sentence containing an equi AUX *po-ass-ta* in (12a), the sentence (12b) has the word *kuray* replacing the VP *pap-ul mek-e* in (12a) and the sentence (12c) has the word *kuray-ass-e* replacing the VP *pap-ul mek-e po-* in (12a). But the AUX *mek-e* in (12a) cannot be replaced by the word *kuray* in (12d) because it is not a VP. Again, the distributional behavior of *kuray* can be sufficiently explained if the VP analysis is taken as the correct analysis.

However, if the compound verb analysis is chosen, an explanation must be provided for how a part of a verb can be replaced by the word *kuray* as in (11b-c) and (12b-c) and why the verb *cwu-ko* or *mek-e* cannot be replaced by *kule* or *kuray* in (11d) and (12d).

Furthermore, there is another compound verb, *cap-e mek-ess-ni* as in (13a), whose constituents cannot be replaced by the word *kuray* as seen in (13b-c).

(13) a. John-i thokki-lul $_{V}$[cap-e mek-ess-ni?].
-N rabbit-A catch-(Comp) eat-P-Q
'Did John catch and eat a rabbit?'
b. *ung, John-i **kuray** mek-ess-e.
yes, -N do so eat-P-Dec
'Yes, John did so and ate.'
c. *ung, John-i **cap-e** kule-ess-e.
yes, -N catch do so-P-Dec
'Yes, John caught and did so.'

Under the VP analysis, (13b) and (13c) are not possible answers to the interrogative sentence (13a) because *cap-e mek-ess-ni* in (13a) is not a phrase (VP) but a single word (verb). Thus a part of the verb, like *cap-e* or *mek-ess-e,* cannot be replaced by the word *kuray* as in (13b) or (13c). Again, the compound verb analysis requires additional restrictions in order to account for why a part of a compound verb in (13a) cannot be replaced by the word *kuray* as in (13b) and (13c).

Therefore, the distributional restrictions of the word *kuray* under the VP analysis follow from the generalization that VPs can be replaced by the word *kuray*.

2.1.3. V^2 fronting

The VP analysis also provides a simpler analysis of verb phrase fronting than the compound verb analysis. Verbal phrases like a VP or S can be fronted when the gap is filled with the verb *ha-* as in (14), whereas the lexical category V^0, *mek-ki-nun*, cannot be fronted as shown in (15).

(14) a. Mary-ka pap-ul mek-nun-ta.
-N rice-A eat-Pres-Dec
'Mary has a meal'

b. [pap-ul mek-ki-nun]$_{VP}$ Mary-ka ø **han-ta.**[10]
rice-A eat-NM-T .....-N do-Dec
'It is Mary who has a meal.'

c. [Mary-ka pap-ul mek-ki-nun]$_S$ ø **han-ta.**
-N rice-A eat-NM-T do-Dec
'Mary has a meal.'

(15) *[mek-ki-nun]$_V$ Mary-ka pap-ul ø **han-ta.**
eat-NM-T -N rice-A do-Dec
'Mary has a meal.'

Since the VP *pap-ul mek-nun-ta* in (14a) is fronted, (14b) is acceptable. In the same way, (14c) is also acceptable because the S *Mary-ka pap-ul mek-nun-ta* in (14a) is fronted. However, (15) is unacceptable since the lexical category (V^0) *mek-nun-ta* in (14a) cannot be fronted.

Either a VP with an AUX as in (16c) or a VP without an AUX as in (16b) can be fronted. The possibility of VP fronting without AUX as in (16b) or with AUX in (16c) follows from a generalization that all V^2 categories can be fronted in Korean under the VP hypothesis.

(16) a. Mary-ka pap-ul **mek-eya** **han-ta.**
-N rice-A eat-VForm must-Dec
'Mary must have a meal.'

b. [pap-ul mek-ki-nun] Mary-ka ha-eya han-ta.
-A eat-NM-T -N must-Comp do-Dec
'It is Mary who must have a meal.'

c. [pap-ul mek-eya ha-ki-nun] Mary-ka han-ta.
-A eat-Comp must-NM-T -N do-Dec
'It is Mary who must have a meal.'

But under the compound verb analysis an additional restriction to explain why a compound verb like *mek-eya han-ta* in (16b) can be divided into two parts to be fronted is necessary.

Cho 1988 argues that the compound verb analysis is preferable because of the fact that scrambling of the VP subcategorized for by AUX is impossible and adverbs immediately preceding AUX are impossible. The HPSG analysis of this problem is in section 4.1.

2.2. Evidence against bi-clausal analysis

Once the compound verb hypothesis is found incorrect, it must be determined if the AUX subcategorizes for a S or a VP. Specifically, the problem here is whether the phrase *chayk-ul ilk-e* in (17=1a) is a S or a VP.

(17=1a) Mary-ka [X **chayk-ul** **ilk-e**] po-ass-ta.
-N book-A read-Comp try-P-Dec
'Mary tried to read a book.'

To show that the phrasal category that each AUX subcategorizes for is a S, the bi-clausal analysis provides two arguments. The first is based on negative po-

larity items. The second concerns the occurrences of *caki-ka* in the X position of (17). However, the VP analysis is more plausible than the bi-clausal analysis because these two arguments actually support the VP analysis rather than the bi-clausal analysis.

2.2.1. Negative polarity items (NPIs)

The VP analysis provides a simpler analysis for the distributional behaviour of NPIs than the bi-clausal analysis. Negative polarity items such as *amwukes-to* must occur with a negative element (Neg) like *anh-* within the same clause, as shown in (18) and (19). The sentences in (18a) and (19a) are acceptable because the NPI *amwukesto* and the Neg *anh-ass-ta* co-occur in the same clause. But (18b) and (19b) are unacceptable because (18b) has no Neg in the sentence and (19b) has no Neg in the embedded clause with the NPI.

(18) a. [Mary-ka **amwukesto** mek-ci **anh-ass-ta**].
-N anything eat Neg-P-Dec
'Mary ate nothing.'

b. *[Mary-ka **amwukesto** mek-ass-ta].
-N anything eat-P-Dec
'Mary ate nothing.'

(19) a. [Mary-ka **amwukesto** mek-ci **anh-ass-ta-ko**]
-N anything eat Neg-P-Dec-Comp
John-i (Sue-lul) seltukha-ess-ta.
-N -A persuade-P-Dec
'John persuaded Sue that Mary ate nothing.'

b. *[Mary-ka **amwukesto** mek-ass-ta-ko]
-N anything eat-P-Dec-Comp
John-i (Sue-lul) seltukha-ci **anh-ass-ta**.
-N tell-P-Comp Neg-P-Dec
'John persuaded Sue that Mary ate nothing.'

Examples (18) and (19) show that the clause-mate constraint, which specifies that a NPI must occur with a Neg in the same clause, is needed to explain the constructions.

The AUX constructions with a NPI as in (20) are also possible. Under the VP analysis (20a) and (20b) are predicted as grammatical because the phrases *amwukesto mek-ci anh-a* in (20a) and *amwukesto mek-e* in (20b) are VPs where the NPI in each VP observes the clause-mate constraint.

(20) a. Mary-ka **amwukesto** **mek-ci** **anh-a** po-ass-ta.
-N anything eat Neg try-P-Dec
'Mary tried not to eat anything.'

b. Mary-ka **amwukesto** **mek-e** po-ci **anh-ass-ta**.
-N anything eat try Neg-P-Dec
'Mary didn't try to eat anything.'

Thus, under the VP analysis, the AUX constructions with a NPI follow from the generalization that a NPI must occur with a Neg in the same clause.

On the other hand, the bi-clausal analysis predicts that the italicized words *amwukesto mek-ci anh-a* in (20a) are a S because the AUX *po-ass-ta* subcategorizes for a S. If the strings are a S and the NPI observes the clause-mate constraint, the grammaticality of (20a) can be accounted for. Still, the analysis wrongly predicts that (20b) is ungrammatical, because the NPI in the S *amwukesto mek-e* violates the clause-mate constraint. There are two possible solutions to this problem. One is that the Tensed S Condition (TSC) is invoked to explain the behaviour of NPIs, instead of the clause-mate constraint. Because the embedded S *amwukesto mek-e* in (20b) has no Tense and TSC restricts the co-occurrence of the NPI and the Neg to a sentence with Tense, TSC can correctly predict that (20b) is grammatical. The other solution is to treat sentences like (20b) as scrambled constructions so that they still seem to observe the clause-mate constraint on NPIs. If the NPI *amwukesto* in (20b) is adjoined to the embedded S by Scrambling, the NPI can occur with the Neg in the same clause. This solution makes the correct prediction for the grammaticality of (20b). However, these alternatives are not preferable.

If the bi-clausal analysis takes Tensed S Condition as the proper constraint to deal with the AUX constructions with a NPI, instead of the clause-mate constraint, the differences in acceptability between (18b & 19b) and (20b) can be accounted for as follows. Even though all three sentences with the NPI *amwukesto* do not have a Neg in the same clause, (20b) is possible because the S *amwukesto mek-e* has no Tense, thereby not violating TSC, whereas (18b) and (19b) are impossible because each clause with a NPI has a Tense and thus violates TSC.

However, examples like (20b) show that although there is a tensed clause with a NPI but no Neg within the same clause, the sentence can be grammatical. The bi-clausal analysis in conjunction with TSC predicts that (21a) is acceptable because the NPI in the embedded S subcategorized for by the AUX *ha-* in (21a) does not violate TSC. But it wrongly predicts that (21b) is unacceptable because the NPI in the embedded S violates TSC.

(21) a. Mary-ka [amwukesto mek-ci anh-ass-eya] ha-ess-ta.
-N anything eat-Comp Neg-P-Comp must-P-Dec
'Mary must not have eaten anything.'

b. Mary-ka [amwukesto mek-ess-eya] ha-ci anh-ass-ta.
-N anything eat-P-Comp must-Comp Neg-P-Dec
'Mary didn't have to eat something.'

Therefore, this solution is not a good candidate to account for these constructions.

With the second solution of the bi-clausal analysis the clause-mate constraint is regarded as correct, but the structure for (20b) is considered scrambled. The NPI *amwukesto* as the object of the verb *mek-e* in (20b) can be analyzed like (22), where it does not move at all, or it may be treated as a scrambled structure like (23), where the NPI moves to the sister of the embedded S.

(20b) Mary-ka **amwukesto** **mek-e** po-ci anh-ass-ta.
-N anything eat try Neg-P-Dec
'Mary didn't try to eat anything.'

(22) Mary-ka$_i$ [PRO$_i$ [amwukesto mek-e]$_{VP}$]$_S$ po-ci anh-ass-ta.

(23) Mary-ka$_i$ [amwukesto$_j$ [PRO$_i$ [t$_j$ mek-e]$_{VP}$]$_S$]$_S$ po-ci anh-ass-ta.

If (20b) has a structure like (23), in which the NPI *amwukesto* is adjoined to the embedded S in terms of scrambling, the NPI belongs to the higher S so that it observes the clause-mate constraint. Thus the bi-clausal analysis does not need any modification for these constructions.

However, there are examples which show that even though we treat some sentences as scrambled to make a NPI occur with a Neg within the same clause, these sentences cannot be grammatical. The sentence in (24) is unacceptable, even if the NPI is scrambled to occur with a Neg within the same clause. Under the bi-clausal analysis, in conjunction with the Scrambling solution, (19b) is unacceptable because the NPI in the embedded sentence violates the clause-mate constraint. On the other hand, (24) should be acceptable because when the NPI in the embedded sentence is adjoined to the S in terms of Scrambling, and it belongs to the higher S which has a Neg, it does not violate the clause-mate constraint. Yet (24) is still unacceptable.

(19b) *[Mary-ka amwukesto mek-ass-ta-ko]
-N anything eat-P-Dec-Comp
John-i (Sue-lul) seltukha-ci anh-ass-ta.
-N tell-P-Comp Neg-P-Dec
'John persuaded Sue that Mary ate nothing.'

(24) *amwukesto$_j$, John-i (Sue-lul) [Mary-ka
anything -N -A -N
ø$_i$ mek-ass-ta-ko]$_S$ seltukha-ci anh-ass-ta.[11]
eat-P-Dec-Comp persuade-Comp Neg-P-Dec

The Scrambling solution, thus, is also not a good candidate to account for these constructions.

Again, if the bi-clausal analysis is taken to explain the AUX constructions, NPI restrictions are necessary, but under the VP analysis such restrictions are not necessary.

2.2.2. The occurrence of *caki-ka*

The VP analysis can correctly predict the occurrence of reflexive *caki-ka* in Korean while the bi-clausal analysis cannot. For the bi-clausal analysis claims that, as in example (25), the reflexive *caki-ka* is the subject of the embedded sentence, where only PRO and the reflexive *caki-ka* can occur in the X position of (17=1a).[12] Conversely, the VP analysis claims that the reflexive is not the subject of the embedded sentence [*X chayk-ul ilk-e*] but an adjunct to modify the subject *Mary-ka* in (17=1a).

(17=1a) Mary-ka [X **chayk-ul** **ilk-e**] po-ass-ta.
-N book-A read-Comp try-P-Dec
'Mary tried to read a book.'

(25) Mary-ka [X **chayk-ul** **ilk-e**] po-ass-ta.
a. ***John-i**
b. ***kunye-ka**(=she)
c. **caki-k*a*(=self-N)**
d. PRO

On the basis of Sells' 1993 claim that the reflexive *caki-ka* and PRO can occur in the X position as in (25c-d) but a R-expression and Pronominal cannot as in (25a-b), the Bi-clausal analysis can claim that the reflexive *caki-ka* is the subject of the embedded sentence so that the phrase [X *chayk-ul ilk-e*] constitutes a S as in (26). If this is true, the occurrence of *caki-ka* in the subject position can support the Bi-clausal analysis, and the VP analysis, assigning (17) a structure like (27), must explain why the VP [*chayk-ul ilk-e*] in the AUX constructions can have a reflexive subject.

(26) Mary-ka [*caki-ka* *chayk-ul* *ilk-e*]$_S$ po-ass-ta.
-N self-N book-A read-Comp try-P-Dec
'Mary herself tried to read a book.'

(27) Mary-ka *caki-ka* [*chayk-ul* *ilk-e*]$_{VP}$ po-ass-ta.
-N self-N book-A read-Comp try-P-Dec
'Mary herself tried to read a book.'

However, there are examples showing that the reflexive with a subject case marker *caki-ka* can freely occur as an adjunct in a sentence. The examples (28a-b), where the reflexive *caki-ka* occurs as an emphatic expression modifying the subject, show that the reflexive with a subject marker need not always be regarded as a subject. The subject of *cohaha-ess-ta* ('liked') in (28a) is *Mary-ka,* and in (28b) the subject of the embedded sentence is *Mary-ka* and that of the higher S is *John-i*. The reflexive *caki-ka* in (28) is an adjunct which modifies *Mary-ka* in (28a) and *John-i/un* in (28b).

(28) a. [Mary-ka *caki-ka* John-ul cohaha-ess-ta]$_S$.
-N self-N -A like-P-Dec
'Mary herself liked John.'
b. John-un/i *caki-ka* [Mary-ka can-ta-ko]$_S$ malha-ess-ta.
-T/N self-N -N sleep-Dec-Comp tell-P-Dec
'John himself said that Mary slept.'

To deal with the emphatic reflexive the Bi-clausal analysis must allow both a structure like (29), for (25c) where the reflexive *caki-ka* is the subject of the embedded clause, and a structure like (30) for (28a). Since the subject of the sentence (30) is not *caki-ka* but *Mary-ka*, the structure in (30), where *caki-ka* is an adjunct, is necessary.

(29) [Mary-ka [*caki-ka chayk-ul ilk-e*]$_S$ po-ass-ta]$_S$.

(30) [Mary-ka *caki-ka* [John-ul cohaha-ess-ta]$_{VP}$]$_S$.

On the other hand, the VP analysis needs only one structure like (27) for (25c) and (28), because this analysis regards the reflexive *caki-ka* in (25c) and (28) as an adjunct.

The difference in grammaticality between (31a) and (31b) shows that the VP analysis predicts the correct structure for the AUX constructions with the emphatic reflexive *caki-ka*. The fact that the multiple occurrences of the emphatic reflexive *caki-ka* in a sentence with AUX are not possible, as in (31a), shows that the string *pap-ul mek-e* in (31a) cannot be a S. However, the sentences with one emphatic reflexive and one reflexive as the subject of the embedded sentence are grammatical as in (31b).

(31) a. *Mary-ka/nun *caki-ka* *caki-ka* pap-ul mek-e po-ass-ta.[13]
-N/-T self-N self-N rice-A eat-Comp try-P-Dec
'Mary herself tried to have a meal.'

b. Mary-nun/ka *caki-ka* *caki-ka* (kacang) yeypputa-ko
-T/-N self-N self-N the most pretty-Comp
malha-n-ta.[14]
tell-Pres-Dec
'Mary herself says that she is pretty.'

Under the VP analysis, the structure of (31a) is regarded as (32a) and the structure of (31b) must be (32b). So this analysis correctly predicts that (32a) is unacceptable but (32b) is acceptable. The sentence (32a) is unacceptable because both reflexives *caki-ka* in (32a) are adjuncts modifying the subject *Mary,* where one of them is redundant. But the sentence (32b) is acceptable because the first *caki-ka* in the higher S is an adjunct and the second is the subject of the embedded sentences subcategorized by the verb *malha-* ('say').

(32) a. *[Mary-ka/nun *caki-ka* *caki-ka* [pap-ul mek-e po-ass-ta]$_{VP}$]$_S$.
-N/-T self self rice eat-Comp try-P-Dec
'Mary herself tried to have meal.'

b. [Mary-nun/ka *caki-ka* [*caki-ka* (kacang) yeyppu-ta-ko]$_S$
-T/-N self-N self-N the most pretty-Comp
malha-n-ta]$_S$.
say-Pres-Dec
'Mary herself says that she is pretty.'

The Bi-clausal analysis wrongly predicts that both (31a) and (31b) are acceptable, because this analysis predicts that the first *caki-ka* in (31a) and (31b) is an adjunct and the second is the subject of the embedded S, as shown in (33). Thus the Bi-clausal analysis must specify additional restrictions to explain why (31a) is ungrammatical and (31b) is grammatical.

(33a)=(31a) *[Mary-ka *caki-ka* [*caki-ka* pap-ul mek-e po-ass-ta]$_S$]$_S$.
(33b)=(31b) [Mary-ka *caki-ka* [*caki-ka* (kacang) yeyppe-ta-ko]$_S$ malha-n-ta]$_S$.

The VP analysis needs no such restrictions. Therefore, the VP analysis makes correct predictions and is preferable, whereas the Bi-clausal analysis does not make correct predictions.

3. Morphological Requirements for the VP

When an AUX subcategorizes for a VP, the VP has at least two restrictions on the suffix form: a restriction on the existence of tense, and a restriction on the suffix form for the Comp.

First of all, the fact that only a VP with the correct suffix for the Comp can be grammatical shows that each AUX subcategorizes for the VP with a specific suffix for the Comp. The sentence (34a) is acceptable because the AUX *po-ass-ta* requires a VP with the Comp *-e* and the verb *ilk-* within the VP has the Comp *-e*.; (34b-d) are excluded because the requirement for the Comp is not satisfied. For example, (34b) is ungrammatical because the verb *ilk-* has the wrong Comp *-eya*. Similarly, the sentence (34'a) is acceptable because the AUX *ha-n-ta* requires a VP with the Comp *-eya* and the verb *ilk-* within the VP has the Comp *-eya*. But (34'b-d) are unacceptable because the requirement for the Comp is not satisfied.

(34) a. Mary-ka chayk-ul *ilk-e* *po-ass-ta.*
-N book-A read-Comp try-P-Dec
'Mary tried to read a book.'
b. *Mary-ka chayk-ul *ilk-eya* *po-ass-ta.*
c. *Mary-ka chayk-ul *ilk-ci* *po-ass-ta.*
d. *Mary-ka chayk-ul *ilk-key* *po-ass-ta.*

(34') a. Mary-ka chayk-ul *ilk-ø-eya* *ha-n-ta.*
-N book-A read-Pres-Comp must-Pres-Dec
'Mary must read the book.'
b. *Mary-ka chayk-ul *ilk-e* *ha-n-ta.*
c. *Mary-ka chayk-ul *ilk-key* *ha-n-ta.*
d. *Mary-ka chayk-ul *ilk-ci* *ha-n-ta.*

Secondly, the fact that some AUXs require a tensed verb while some do not shows that each AUX subcategorizes for a VP but the existence of the tense suffix within the VP depends on the AUX. The sentence (35a) is acceptable because the verb *ilk-e* does not have a Tense suffix, whereas (35b) is unacceptable because the verb *ilk-e* has a tense suffix. On the other hand, if the AUX is *hanta* the verb *mek-eya* in the VP must have a tense suffix as in (36).

(35) a. Mary-ka chayk-ul *ilk-e* *po-ass-ta.*
-N book-A read-Comp try-P-Dec
'Mary tried to read a book.'
b. *Mary-ka chayk-ul *ilk-ess-e* *po-ass-ta.*
read-P-Comp

(36) a. Mary-ka pap-ul *mek-ø-eya* *ha-n-ta.*
-N rice-A eat-Pres-Comp must-Pres-Dec
'Mary must have a meal.'
b. Mary-ka pap-ul *mek-ess-eya* *ha-n-ta.*
eat-P-Comp

To account for these morphological problems there are at least two approaches: the lexical approach and the syntactic approach. The syntactic approach assumes that there are several syntactic nodes, such as INFL Phrase (IP), where the AUX may or may not subcategorize for a IP, instead of a VP or a S, in order to account for the requirement for the tense morpheme. The lexical approach treats the Tense requirement in terms of features in the syntax, where the feature and its value can play an important role in selecting the correct morpheme from the lexicon. The next section argues for the lexical approach to solve the inflectional problem (e.g. tense).[15]

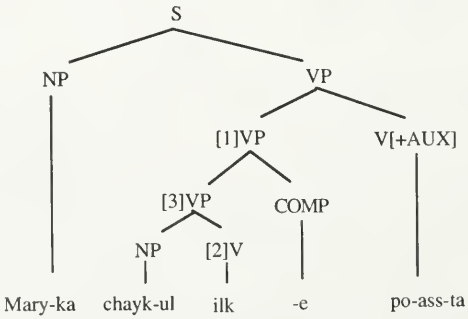
3.2. Analysis of Tense and Comp in HPSG

If the AUX *po-ass-ta* subcategorizes for a VP and a NP, we can specify the subcategorization information in the SUBCAT of the lexical representation of AUX in HPSG, as shown in (37).

(37) |SUBCAT <NP[1], [2]VP |HEAD [3] |> |
| |MARKING [4] | |
| |SUBCAT<NP[1]> | |

However, we need to indicate that the verb within the VP must have specific suffixes with respect to AUX. One way to ensure the matching of the morphological information between the VP and the verb is to treat tense information as a HEAD feature and to treat the information about Comp as a marker in HPSG. For example, when the AUX *po-ass-ta* takes a VP as an argument, the base verb form *ilk-* in the VP must have Comp *-e* but cannot have tense in grammatical sentences like (38).[16]

(38)



In this case we can specify the information about the tense and the Comp for the VP in the subcategorization (SUBCAT) of (37). Once the morphological information is specified for [1]VP in (38), the information about tense must be identical

with [2]V in (38). The HEAD Feature Principle (HFP) in HPSG, as shown in (39), specifies this condition.[17]

(39) The HEAD Feature Principle.
In a headed phrase, the values of HEAD and HEAD-Daughter's HEAD are token-identical.

The information about Comp is also specified on the VP in the SUBCAT in (37) as the MARKING feature and its value. If the value of MARKING is specified on [1]VP in the tree (38), the value triggers a sort *head-marker-structure* in terms of schema 4 in (40) and the value (Comp) is realized as a marker daughter.

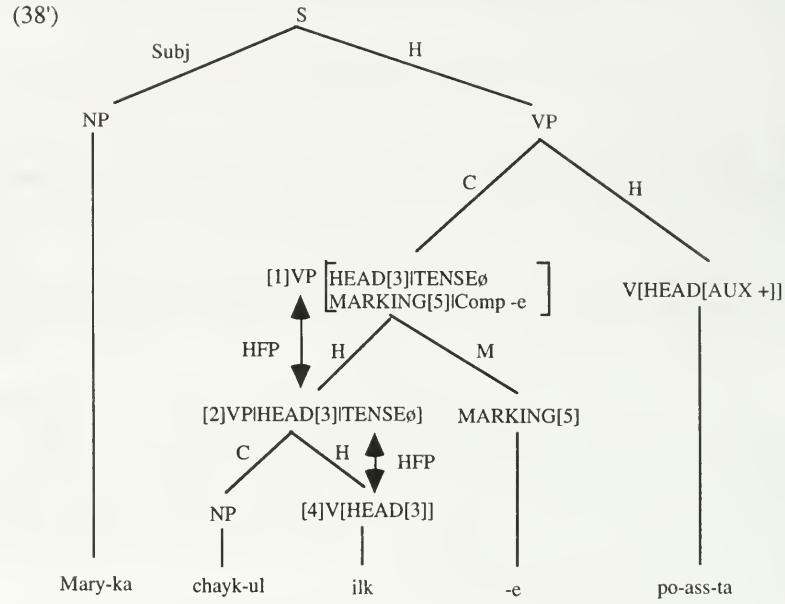
(40) Schema 4: a phrase with DTRS value of sort head-marker-structure whose marker daughter is a marker whose SPEC value is structure-shared with the SYNSEM value of the head daughter, and whose MARKING value is structure-shared with that of the mother.

The sort hierarchy and feature declarations related to the information about tense and Comp, defined using the HFP and schema 4, as shown in (41-42). If the sort *head* as a value of the attribute HEAD in (42a) is a sort *verb* as a subsort of *substantive*, the sort *verb* must have a *tense* value like *P(ast)* for the attribute TENSE as in (42b). Similarly, if the sort *marking* as a value of the attribute MARKING in (42a) is a sort *complementizer* as a subsort of the sort *marked*, the sort *complementizer* must have a subsort like *-e*.

(41) a. partition of *head*: *substantive, functional*
b. partition of *substantive:* *noun, verb, adjective, ...*
c. partition of *tense:* *ø, P(ast), Pre(sent), ...*
d. partition of *functional:* *marker, determiner*
e. partition of *marking:* *unmarked, marked*
f. partition of marked: *complementizer (Comp), conjunction*
g. partition of *complementizer: -e, -key, -ci, -ko, ...*

(42) a. *category*: |HEAD *head* |
|SUBCAT *list(synsem)* |
|MARKING *marking* |
b. *verb*: [TENSE *tense*]

Under the HPSG analysis, including the HFP in (39), Schema 4 in (40), the sort hierarchy in (41) and the feature declarations in (42), the new tree diagram (38') replaces (38). The tree (38') shows that [1]VP in the head-marker structure has a tag [3] as the value of the attribute HEAD, including the tense information which is also the value for [2]VP and [4]V, indicating that the value of HEAD of the three categories is token-identical. Thus, the structure satisfies the HFP, and the morphological requirements for tense can be dealt with in terms of the HFP. The morphological requirement for the Comp can be dealt with in terms of the head-marker schema in (40). If the AUX *po-ass-ta* subcategorizes for a VP with the Comp *-e*, the information for the Comp is specified as the MARKING feature and its value as shown in [1]VP of the tree (38') where the value of MARKING is structure-shared with that of mother, [1]VP, by the definition (40).



(Subj = Subject daughter, C = Complement daughter,
H = Head daughter and M = Marker daughter.)

4. SUBCAT for AUX in HPSG

On the basis of the VP hypothesis and the morphological information for the VP, (43) shows the complete SUBCAT equi (type 1) and raising (type 2) AUXs in HPSG.

(43) AUX: TYPE1 (Equi)
ex: *po-* ('try'), ...
TYPE2 (Raising)
ex: *ci-* ('become'), *po-* ('seem'),[18] *ha-* ('must'), ...

The TYPE1 AUX *po-* ('try') as an equi-verb has a SUBCAT and a (semantic) CONTENT as shown in (44). The Attribute Value Matrix (AVM) (44a) specifies that the AUX needs a NP and a VP to be saturated and the values of COMP and TENSE for the VP are *-e* and *0*, respectively. In addition, the SUBCAT in the VP indicates that the INDEX of the subject of the AUX and the VP must be identical. The AVM (44b) defines the CONTENT of the AUX *po-* where its RELATION *try* has two arguments. One is the COMMITTOR whose INDEX value is [1] and the other is the SOA-ARG whose value is a proposition.

(44) a. |SUBCAT <NP[1], [2]VP |HEAD |TENSE *ø* | |> |
| | |MARKING |COMP *-e* | | |
| | |SUBCAT <NP[1]> | | |

b. |CONTENT | RELATION *po-* ('try') |
| | COMMITTOR [1] |
| | SOA-ARG [2] |

When the AUX is in a sentence like *Mary-ka pap-ul mek-e po-ass-ta*, NP[1] represents the NP *Mary* and the value of SOA-ARG is (45), where the INDEX [1] as a value of the argument role EATER refers to the INDEX of *Mary* and the INDEX [3] as a value of EATEN refers to that of *pap-ul* ('rice').

(45) SOA-ARG: | RELATION *mek-* ('eat') |
[2] | EATER [1] |
| EATEN [3] |

The TYPE 2 AUX *ci-* ('become') has SUBCAT and CONTENT as shown in (46). The AVM (46a) specifies that the AUX *ci-* needs a NP and a AP[+PRD] to be saturated, and that the value of COMP for the AP is *-e* and that of TENSE must be *0*. The SUBCAT of the AP also indicates that the INDEX of the subject of the AUX and that of the AP must be identical. The AVM (46b) defines the CONTENT of the AUX *ci-* where its RELATION *become* has only one argument, SOA-ARG.

(46) a. |SUBCAT <NP[1], [2]AP |HEAD |TENSE *0* | |> |
		PRD+		
		SUBCAT <NP[1]>	
		MARKING	COMP-*e*	

b. |CONTENT | RELATION *ci-* ('become') |
| | SOA-ARG [2] | |

When the AUX is in a sentence like *Mary-ka yeyppu-e ci-ass-ta* ('Mary became pretty.'), the INDEX value of the subject NP of the AUX in SUBCAT refers to that of the NP *Mary* and the SOA-ARG in (46b) is as represented in (47). For the SOA-ARG in (47), the INDEX value of the argument role INSTANCE and that of the NP *Mary* in SUBCAT are identical.

(47) SOA-ARG: [2] | RELATION *yeypp-* ('pretty') |
| INSTANCE [1] |

The AUX *po-* ('seem'), as a member of TYPE2, has a different SUBCAT than other members of TYPE 2 like *ci-* ('become'). Its SUBCAT takes a VP as one of its arguments while the AUX *ci-* ('become') needs a AP to be saturated. The lexical representation for the AUX *po-* ('seem') is as represented in (48). The SUBCAT in (48a) specifies that the AUX needs a NP and a VP to be saturated, the value of COMP for the VP is *-na* and the value of TENSE can be *Pres(ent)* or *P(ast)*. Again, the AVM (48b) defines the CONTENT of the AUX *po-* where its RELATION *seem* has only one argument, SOA-ARG.

(48) a. |SUBCAT <NP[1], [2]VP |HEAD [TENSE Pres V P] |> |
| |SUBCAT <NP[1]> | |
| |MARKING|COMP *-na* | | |

b. |CONTENT | RELATION *po-* ('seem') | |
| | SOA-ARG [2] | |

When the AUX exists in a sentence like *Mary-ka chayk-ul ilk-na po-ta* ('Mary seems to read a book'), the SOA-ARG in (48b) can be represented as in (49).

(49) SOA-ARG: | RELATION *ilk-* ('read') |
[2] | READER [1] |
| READED [3] |

4.1. Evidence against Phrasal analyses

Cho 1988 proposes an argument against Phrasal analyses on the basis of the impossibility of Scrambling of the VP subcategorized for by AUX and the impossibility of adverbs occurring in front of AUX.[19] The following will show how these phenomena can be accounted for in HPSG.

Cho argues that (50a) can be scrambled in Korean, with the result (50b), whereas (51a) cannot be scrambled as in (51b). According to her explanation, (51b) is unacceptable because *mek-e po-ass-ta* in (51a) is a compound verb and the NP *pap-ul* and the verb *mek-e* cannot be a constituent, so that the string *pap-ul mek-e* cannot undergo Scrambling.

(50) a. John-un [Suni-ka ka-ass-ta-ko]$_S$ sayngkakha-ess-ta.
John-top Suni-N go-P-Dec-Comp think-P-Dec
'John thought Suni went away.'
b. [Suni-ka ka-ass-ta-ko]$_S$ John-un sayngkakha-ess-ta

(51) a. John-un [pap-ul mek-e] po-ass-ta
John-top rice-A eat-Comp try-P-Dec
'John tried to have a meal.'
b. *[pap-ul mek-e] John-un po-ass-ta

The evidence that the S *Suni-ka ka-ass-ta-ko* in (50) can be scrambled but the S *pap-ul mek-e* in (51) cannot is a crucial argument against the Bi-clausal analysis, because the difference in acceptability between (50b) and (51b) must be stipulated.

However, the VP analysis predicts that (50b) is possible while (51b) is impossible. The fact that the sentences, where predicative categories such as VP or AP[+PRD] or NP[+PRD] undergo Scrambling, are unacceptable shows that the VP subcategorized for by AUX also cannot undergo Scrambling. The sentence (52), which contains a small clause, is ungrammatical because the NP[+PRD] *papo-lul* is scrambled with the result in (52b). To deal with the scrambling problem in a small clause, Yoo 1993 proposes a Linear Precedence (LP) rule specifying that any predicative category cannot precede its subject, as shown in (53). This LP prevents the NP *Mary-lul* and the NP[+PRD] *papo-lul* in (52a) from Scrambling. The LP rule states that if there is any predicative category, like a VP, which needs only a SUBJECT to be saturated, that predicative category cannot precede the SUBJECT. So the independently motivated LP predicts that (50b) is possible and (51b) is not. In other words, the constituent *Suni-ka ka-ass-ta-ko* in

(50) is a S so that it has no need to observe the LP whereas *pap-ul mek-e* is a VP under the VP analysis which must.

(52) a. John-i Mary-lul *papo-lul/lo* mantul-ess-ta.
-N -A fool-A/-PP make-P-Dec
'John made Mary a fool.'

b. *John-i *papo-lul/lo* Mary-lul mantul-ess-ta.[20]
-N -A/PP -A make-P-Dec
'John made Mary a fool.'

(53) [1] < [VALENCE|SUBJ<[1]>]

Under the VP analysis, the difference between (50b) and (51b) naturally follows from the LP. Therefore, the evidence in (50-51) is not a counter-example to the VP analysis but, rather, supports the claim that AUX subcategorizes for a VP, not a S, in Korean.

Cho 1988 also claims that adverbs like *cacwu* ('often') cannot occur in front of the AUX as shown in (54). On the basis of the fact that the adverb cannot intervene between the verb *mek-e* and the AUX *po-ass-ta*, she claims that the verb and the AUX form a compound verb so that the adverb cannot modify the AUX and argues that Phrasal analyses are implausible since this problem is difficult to solve under these analyses.

(54) *John-i pap-ul mek-e *cacwu* po-ass-ta.
-N -A eat-Comp often try-P-Dec
'John often tried to have a meal.'

However, the fact that the adverb can occur between a verb and some AUX shows that her argument against Phrasal analyses is untenable. The sentence with the causative AUX *ha-ta* ('cause') is acceptable even if an adverb occurs between the two, as shown in (55).

(55) John-i chayk-ul ilk-key *cacwu* ha-n-ta.
John-N book-A read often cause-Pres-Dec
'John often causes someone to read a book.'

If sentence (54) is considered unacceptable but sentence (55) is considered acceptable, even though both sentences have an adverb modifying AUX, the argument for the Compound Verb analysis is not valid.[21]

Therefore, Cho's argument against Phrasal analyses is not tenable for the VP analysis.

5. Conclusion

The fact that, in spite of the Compound Verb analysts' claims, AUXs are an independent category has been shown through constituency tests such as scope interpretations, *kuray* + verb constructions and verbal fronting in sec 2.1. Section 2.2 showed that AUX subcategorizes for a VP rather than a S by demonstrating that against the Bi-clausal analysts' claims, sentences with NPI or the reflexive *caki-ka* can be evidence for the VP analysis. Thus, this supports the claim that the

VP analysis is more plausible than both the Compound Verb analysis and the Biclausal analysis in explaining the AUX constructions. Section 3 claimed that the morphological requirements for the VP can be accounted for in terms of the HFP and schema 4 only if tense and Comp are regarded as a value of HEAD feature and a marker, respectively. On the basis of section 2. and 3. section 4 presented the proposal for two types of AUXs, equi- and raising-AUX, under the HPSG framework.

Consequently, AUX constructions can be sufficiently accounted for by the VP analysis. Furthermore, the Compound Verb analysis must not be assumed to be the only hypothesis to explain these constructions in HPSG. Rather, if my analysis is chosen, the restrictions for the AUX constructions can be regarded as a subcase of the restrictions for the VP.

NOTES

* I am grateful to professors Georgia Green, Jerry Morgan and James Yoon for their valuable comments.

[1] Following Choi 1971 I will call an auxiliary verb the last one of the italicized two sequencing verbs in the data. For example, *po-ass-ta* will be an auxiliary verb in (1a).

[2] In this paper I refer to Pollard & Sag 1994 as HPSG 1994 and Pollard & Sag 1987 as HPSG 1987.

[3] In describing Korean sentences, I will use the Yale Romanization System.

[4] N stands for Nominative, A Accusative, Comp Complementizer, P Past, Pres Present, Dec Declarative, Neg Negation, Q Question, and T Topic.

[5] Both VP and S are called V^2 in this paper.

[6] Yoo 1993 simply adopts Sells' analysis in studying subcategorization in Korean.

[7] My analysis will cover both raising and equi auxiliary verbs, even though in this paper I mainly deal with the AUX in (1a).

[8] Sells 1991 proposes a semantic analysis of the scope problem. His analysis is based on the classification of event types of each AUX. However, it is not clear that his analysis can predict the scope problem in coordination.

[9] *Kule* is a variant form for *kuray*.

[10] NM stands for Nominalizing Suffix.

[11] For me, (24) is unacceptable. Regardless of bridge verbs like *sayngkakha-* ('think') or non-bridge verbs like *seltukha-* in (24), the interpretation where the NPI negates the embedded sentence is almost impossible. However, I think we can get the interpretation where the NPI only negates the higher sentence. For example, (24) can have a interpretation like *John didn't persuade Sue anything about the fact that Mary ate (something).*

[12] Sells 1993 claims that a Japanese phrase like [*X chayk-ul ilk-e*] in (17=1a) is a sub-clause, showing the distributional behaviour of NP for the X position in (25).

[13] The sentence (31a) may improve a little bit with a pause between two *caki-ka*. Still, it is unacceptable for me. On the other hand, (31b) is better in acceptability.

[14] The verb *malha-* takes a NP, a S and an optional PP as its arguments.

[15] I call my approach the lexical approach in that the information about the required suffixes is specified as features.

[16] This tree is only an abbreviation for Attribute Value Matrix.

[17] The definition of HFP in (39) is a simplified one to enhance the readability. The original definition of the HFP is as follows:

The HFP.

In a headed phrase, the values of SYNSEM| LOCAL| CATEGORY| HEAD and DAUGHTERS| HEAD-DAUGHTER| SYNSEM| LOCAL| CATEGORY| HEAD are token-identical. cf. HPSG (1994:491)

[18] Even though the raising-AUX *po-*('seem') has the same phonological base form as the equi-AUX *po-* ('try'), the two AUXs are different words.

[19] Cho 1988 uses the term *Phrasal analyses* to refer to both the VP analysis and the Bi-clausal analysis.

[20] The PP stands for Postposition.

[21] The question why (55) is good but (54) is bad is open to further study.

REFERENCES

CHO, J. O. 1988. Suffixed verb forms and compound verb constructions. In Eung-Jin Baek(ed.). Papers from the Sixth International Conference on Korean Linguistics. International Circle of Korean Linguistics 6: 77-106.

CHOI, H. B. 1971. Korean grammar. Seoul: Chengumsa.

CHOMSKY, N. 1981. Lectures on government and binding. Dordrecht: Foris Publications.

PARK, B. S. 1990. Contrastive analysis of lexical, morphological and syntactic structures for English-Korean machine translation. Kyung Hee Language Institute.

POLLARD, C. and I. Sag. 1987. Information-based syntax and semantics, Vol. 1. CSLI, Stanford University.

——. 1994. Head-driven phrase structure grammar. University of Chicago Press and CSLI Publications.

SELLS, P. 1991. Complex verbs and argument structures in Korean. Harvard Studies in Korean Linguistics 4: 395-406.

——. 1991. Korean and Japanese morphology from a lexical perspective. MS., Stanford University.

——. 1993. The projection of phrase structure and argument structure in Japanese. MS., Stanford University.

YOO, E. J. 1993. Subcategorization and case marking in Korean. In A. Kathol and C. Pollard eds. Ohio-State University Working Papers in Linguistics 42: 178-198.

YOON, J. H.-S. 1987. Reconciling lexical integrity with affixation in syntax. In J. McDonough and B. Plunkett eds. Proceedings of North Eastern Linguistic Society 17: 663-683.

——. 1993. The headedness of inflectional structures in Korean. Seoul International Conference on Linguistics 12: 826-838.

Studies in the Linguistic Sciences
Volume 23, Number 2, Fall 1993 [publ. October 1996]

CRITICIZING WITH A QUESTION

Cassandre Creswell

People in intimate, informal relationships frequently use questions as an indirect method to criticize an undesirable situation, particularly one that is already in existence and contrary to what the speaker wishes, but in accordance with what the speaker believes the addressee wants, as in example (1):

(1) Are we going this way? (*implicating* 'This is not a desirable way to go.')

The purpose of this paper is to explain how and why questions of this kind can be used to implicate criticisms.

1. Introduction

The first section discusses how these questions are distinguished from sincere questions with no implicature of criticism. I claim the crucial criteria on which correct interpretation depends are the beliefs of the addressee, and show how these beliefs result in a pattern of inferences that lead to correct or incorrect interpretation of the question as a criticism. The second section identifies why a speaker's use of these questions is a strategy of politeness. Briefly, these questions allow the speaker to refrain from directly performing a face-threatening act in order to satisfy the negative face of the addressee (Brown and Levinson 1987), and they allow the speaker to offer options by giving the addressee more than one possiblity of how to react (Lakoff 1973). Both of these reasons are subsumed by the more fundamental desire of the speaker to preserve an informal, intimate relationship with the addressee. The final section compares rephrasings of the question in alternate forms with a different quality of politeness but with the same ability to criticize some state of affairs. This comparison explains why a speaker chooses to use a question, rather than some other form, to express criticism.

Form of the Question

Although an explanation of how criticism-implying questions function should be universally applicable to all questions that implicate criticism, in this paper I restrict the class of questions considered to positively phrased, yes-no questions that use forms of *be* or *do*. This restriction of class allows a more focused and effective explanation of the logic of my argument.

The criterion of positive phrasing eliminates questions with the *n't* contraction as in (2), but allows questions like (3) where the *not* occurs in a position other than the position immediately following the initial element:

(2) Aren't you going to get a haircut?
(3) Are you not going to get a haircut?

Question (2) can make either an implicature that the speaker believes that the addressee intends to get a haircut or that the addressee intends to refrain from getting a haircut and in either case, the addressee's intention is inconsistent with another assumption of the speaker. Question (3), however, can only implicate that the speaker believes the addressee intends to refrain from getting a haircut. This double implicature of (2) with its conflicting possible interpretations for exactly what state of affairs is being criticized interferes with a clear characterization of the necessary inferences the addressee makes. Therefore, including negatively-phrased questions in this analysis would unnecessarily complicate the characterization.

No restriction on tense of verb or person of the subject is motivated. Earlier examples—(3), using second person, and (1), using first person plural—and further examples (4-6) using third person and a variety of tenses can all be used to implicate criticism:

(4) Does that shirt have to be washed?
(5) Did that shirt have to be washed?
(6) Will that shirt have to be washed?

Variations in person and tense do not affect the usefulness of a question for implicating criticism. Neither do they complicate the description of the inferences an addressee makes, and so they are not restricted in this characterization.

2. Distinguishing Criteria: Beliefs and Inferences of the Addressee

Although criticizing with a question is an act of a speaker, the characterization of criticism-implicating questions is explained from an addressee's, rather than a speaker's, perspective. Questions can only be successfully used to criticize if they are interpreted as implicating criticism. Correct interpretation depends on the beliefs of the addressee about the speaker's intentions, not the speaker's intentions themselves. The addressee must be able to distinguish a critical question from a sincere request for information. The difference can only be perceived if the addressee holds certain beliefs. The absence of these beliefs will prevent the addressee from making the inferences necessary for a correct interpretation of the question. Therefore, characterizing the beliefs and inferences of the addressee accounts for both successful and unsuccessful interpretation of questions that imply criticism.

The addressee must hold two beliefs for interpretation of criticism-implying questions. First, the addressee must believe that the speaker believes she[1] knows the answer to the question, and second, the addressee must believe that the speaker intends for the addressee to believe that the speaker knows the answer. In the absence of these beliefs, misinterpretation occurs. In order to illustrate how different interpretations can be generated, different beliefs of the addressee will be matched with the use of an example question in a given situation. In the situation in (7), John is holding a really ugly shirt. Mary asks John:

(7) Are you going to wear that shirt?

First, if John does not believe that Mary believes he does intend to wear that shirt, then he cannot correctly interpret the question as a criticism. One possibility for a non-criticism interpretation occurs if John believes that Mary believes that he does *not* intend to wear the shirt. The utterance will be almost nonsensical; John will have a difficult time thinking of any interpretation of it.[2] A second possibility for non-criticism interpretation occurs if John believes that Mary holds no belief about whether or not he intends to wear the shirt. In this case then, the utterance must be a sincere question because Mary wants information, i.e. she wants to know what John intends to do.

Correct interpretation as implying criticism can only occur if, as stated above, the addressee believes the speaker believes she knows the answer. This inference is made through assessing the relevance of the question. When a speaker asks a question with an answer she already knows, she is apparently violating Grice's Cooperative Principle that conversational contributions must follow the accepted purpose or direction of the exchange (Grice 1975). A question with a known answer makes no readily recognized contribution to a conversation. No obvious reason exists for the speaker to ask something she already knows. As a rational human being engaged in conversation, the addressee assumes that the speaker is following the Cooperative Principle and, therefore, a reason does exist for her utterance. He then constructs a reason for asking such a question, and infers what implicature the speaker desired to make through her use of the question.

The chain of reasoning that leads to an implicature of criticism can best be outlined in combination with the use of an example, such as (7) in the above situation. This chain begins with John's beliefs: one, that Mary knows the answer to (7), that is she already thinks 'Yes, he is going to wear the shirt,' and two, that Mary intends for him to believe that she knows this answer. Holding these two beliefs, John cannot regard the question as a sincere request for information because Mary apparently already possesses the information. Nonetheless, if Mary is asking about the wearing of the shirt, it must have some relevance.[3] Questioning a state of affairs, the wearing of a certain shirt in this situation, regarded as definitely true could hardly be rational if the speaker is in full support of such a state of affairs. On the other hand if the speaker is unhappy with a state of affairs, calling attention to the situation through asking an obvious question is perfectly rational because, if the addressee is aware of the speaker's unhappiness with a situation, he may try to rectify the situation in accordance with the speaker's wishes. So, the next inference the addressee makes is that the relevance of the question lies in its ability to call attention to a state of affairs and make the speaker's unhappiness known. Making one's unhappiness about a state of affairs known is a very simple definition of criticism.

The example situation can demonstrate this final part of the chain of inference too. John must interpret the apparent irrelevance of Mary's question about his wearing of the shirt as Mary calling attention to his plan to wear the ugly shirt in order to implicate her unhappiness with this state of affairs and her wish to make this unhappiness known. If John does follow this chain of inferences, he can cor-

rectly interpret Mary's question as an implicature of criticism, specifically the implicature in (8):

(8) 'I don't think you should wear that shirt.'

The chain of inferences made by the addressee in this particular situation can be applied in any other situation in which questions are used to implicate criticisms. The following example situation will further demonstrate the process of interpretation. In this case, John has been repeatedly cracking his knuckles for the last five minutes. Mary asks John (9).

(9) Do you have to do that?

The steps of reasoning John needs to follow to interpret (9) as criticism-implicating are very similar to the ones in the situation explicated above. First, John must believe that Mary already knows the answer to her question; she thinks 'No, there is no compelling reason for John to be cracking his knuckles.' He must also believe that Mary intends for him to believe that she knows this answer. Because he thinks she knows the answer already, the question cannot be a sincere request for information, and he must construct an alternate explanation for Mary's asking it. Because the answer to her question is obvious, John must interpret her asking it as a way to call attention to the state of affairs she is questioning. While questioning a situation that Mary is in full support of would not be regarded as rational, questioning a situation that she is displeased with seems reasonable. So, John interprets Mary's questioning of the necessity of his cracking his knuckles as implying that it bothers her, more directly stated as the assertion in (10):

(10) 'You don't have to crack your knuckles, and it bothers me that you are doing so.'

The chain of reasoning used in interpreting criticism-implying questions in terms of the beliefs of the addressee can be generalized and applied to different uses of this type of question in order to explain their implicatures. In addition, this chain supports the characterization of this type by explaining how the beliefs of the addressee distinguish criticism-implicating questions from sincere information-requesting ones.

3. Motivations for Use

If criticizing with a question is a politeness strategy, it must be consistent with the principles of a general theory of politeness. Politeness in discourse can be regarded in a very general way as a means to maintain or change interpersonal relations (Green 1989). Because the act of criticizing may disrupt a given level of interpersonal relations, in order to be polite a speaker will try to minimize this disruption. Speakers want to maintain and change relations even within their most informal and intimate relationships, the kind in which criticism-implying questions are frequently used.

This desire to preserve the speaker's informal, intimate relationship with the addressee motivates two considerations for her. First, she must refrain from directly threatening the addressee's negative face by showing respect for the

addressee's self-image and desire for freedom of action. Second, she must offer the addressee options by allowing the addressee more than one possibility of how to react to the criticism. One means of acting in accordance with these two considerations is the criticism-implying question.

Threatening the Addressee

Sincere questions pose a threat to the addressee's face only in that they expect him to use his time to answer them and to know what the answer is. Criticism-implying questions are much more threatening because they express the speaker's doubts and displeasure about a state of affairs the addressee is presumed to be responsible for or able to rectify. Because it is non-threatening, a sincere question does not require the speaker to apologize when the addressee offers an unanticipated response, although she could offer an apology for imposing upon the addressee, as in the dialog in (11):

(11) Sue: Sorry to bother you, but did you let the cat out?
Matt: No, I didn't.
Sue: Oh, okay. Just wondering.

In contrast, if the question is to implicate criticism, an unexpected answer will merit an apology, as in (12), where Sue has discovered the cat outside:

(12) Sue: Did you let the cat out? (implying 'The cat ought not to have been let out.')
Matt: No, I didn't.
Sue: Oh, sorry.

The mistaken criticism results in an apology by the speaker for making an unnecessary threat. An initial apology like that in (11) seems incongruent when matched with a question the speaker is using to criticize, as in example (13):

(13) I don't mean to bother you, but are you really going out of the house in that hideous shirt for the third day in a row?

The combination of a statement that mitigates threat, *I don't mean to bother you*, with a question that strongly implicates criticism, is self-contradictory and will probably result in a conscious attempt by the addressee to assess the speaker's reason for using such a combination. The assessment he generates may be that the speaker's use of the first is entirely insincere and used for a sarcastic effect or, along opposite lines, that she is in fact sincerely concerned and desires more information about actions she regards as unusual.

As shown above, the criticism-implying question does threaten the addressee's face to a greater degree than its information-requesting counterpart does. The speaker mitigates the threatening aspect of criticism by implicating rather than asserting it. Because more than one communicative intention can be inferred from a criticism-implicating question, it is done "off-record", a strategy of negative politeness behavior, as characterized in Brown and Levinson 1987. An off-record strategy does not commit a speaker to a face-threatening act as strongly as one done

on-record, and so the speaker can use this type of question to criticize with less risk of disrupting her relationship with the addressee.

Offering the Addressee Options

The criticism-implicating question's surface resemblance to an information-requesting question means it offers the addressee options. The speaker's implicature can be ignored if the addressee disregards his own beliefs about the speaker's beliefs and responds as if the question was a request for information. When the addressee takes this option, it results in a discourse like (14):

(14) Mary: Are you going to wear that shirt? (attempting to implicate 'That's an ugly shirt and I don't think you should wear it.')
John: Yes, I am.

With an affirmative response here, Mary must assume either that John did not understand what she was attempting to imply with her question or that John ignored her implicature deliberately. In either case, in order to attempt to remedy John's misunderstanding or to emphasize her unhappiness with the state of affairs, Mary might respond with a more direct statement of what she intended to implicate, as in (15):

(15) Well, I don't think you should. It's an ugly shirt.

The appropriateness of responding to the exchange in (14) with (15) supports the claim that although the implicature of (14) can be ignored, intentionally or unintentionally, its existence can be reaffirmed if the speaker asserts it directly. As discussed in the previous section, by offering the option of ignoring its criticism the criticism-implying question can function as a politeness strategy, allowing the speaker to preserve a relationship and still voice a potentially threatening criticism.

4. Other Forms that Criticize

A speaker uses a criticism-implicating question as a politeness strategy to mitigate threatening criticism. A criticism that differs in form may also differ in its politeness. Criticism in question form is more polite, i.e., less threatening of others' beliefs and values, than other forms that do not offer the same kinds of options in interpretation. As explained above, a question gives the addressee the option of ignoring the implicature of criticism and interpreting the question as a sincere request for information because the criticism is only implied, not directly expressed. When a speaker wants her utterance to be less polite, she will use more directness in stating the criticism, as in examples (16-20). The example decrease in directness from (16) to (20):

(16) Inviting John to the reception is wrong and it makes me unhappy.
(17) I can't believe John is invited to the reception.
(18) John is invited to the reception?!
(19) Is John invited to the reception?
(20) I see John is invited to the reception.

Example (16), because it directly states the criticism, allows the addressee no options in interpretation. Example (17) could be interpreted literally as a statement of the speaker's disbelief, but the high degree of conventionalization of the implicature of *I can't believe* as 'I am surprised and/or dismayed that such a state of affairs exists' makes the literal interpretation unlikely. Although (18), as an exclamation, only implicates the criticism, it cannot be treated as a sincere request for information, and as a question, it will be interpreted by way of the same reasoning as any other criticism-implicating question. In contrast, when uttered with the typical rising intonation of a question, (19) cannot be treated as an exclamation, only as a question, making it less directly critical and more polite. The statement of fact, (20), can implicate the criticism in much the same way as the question (19); by calling attention to the obvious, it will lead to a very similar chain of implicature. These rephrasings of a single criticism in multiple ways, one of which is in the form of the criticism-implicating question, differ in politeness because they differ in directness. The fact that each form can express the same criticism provides proof that the questions being characterized can be used to criticize. The fact that each expresses a different degree of politeness provides a reason for a speaker to choose to use one form, such as the question, instead of one of the others.

5. Conclusion

This paper has demonstrated how positively-phrased, yes-no questions that begin with a form of *be* or *do* can be used to implicate criticism. First, the specific beliefs and inferences of the addressee necessary for interpreting this kind of question as implying criticism rather than requesting information were made explicit. Second, the use of this type of question as a politeness strategy was explained. Finally, rephrasings of different degrees of politeness were compared with criticism-implicating questions. Together these three sections provide a thorough characterization of how and why a speaker criticizes with a question.

NOTES

[1] Throughout the paper the speaker will be referred to as female and the addressee as male, in accordance with the sex of the participants in the examples used.

[2] Nonetheless, the natural inclination to interpret speech and in fact human behavior in general as rational, i.e., done with a reason, discussed in Green 1993 means that he will still attempt to construct an interpretation and a reason for her apparently nonsensical utterance.

[3] This claim seems quite similar to that of Sperber and Wilson 1987, in which they assert that "a speaker who asks a question ... indicates that some relevant completion of the incomplete thought represented by her utterance is relevant." Disregarding the controversial definition of relevance outlined therein and instead thinking of relevance merely as "conforming to Grice's Maxim of

Relevance," this assertion is a good characterization of the chain of reasoning behind criticism interpretation.

REFERENCES

BROWN, Penelope, and Stephen Levinson. 1987. Politeness: Some universals in language use. Cambridge University Press.

GREEN, Georgia. 1989. Pragmatics and natural language understanding. Hillsdale, NJ: Erlbaum.

——. 1993. Rationality and Gricean inference. Cognitive Science Technical Report, Beckman Institute, University of Illinois at Urbana-Champaign.

GRICE, H. Paul. 1975. Logic and conversation. Syntax and semantics, vol. 3: Speech acts, ed. by Peter Cole & Jerry L. Morgan, 41-58. New York: Academic Press.

LAKOFF, Robin. 1973. The logic of politeness; Or, minding your P's and Q's. Papers from the Ninth Regional Meeting of the Chicago Linguistic Society, ed. by C. Corum, T. C. Smith-Stark, & A. Weiser, 292-305. Chicago: Chicago Linguistic Society.

SPERBER, Dan, and Wilson, Deirdre. 1987. Précis of Relevance: Communication and cognition. Behavioral and Brain Sciences 10: 697-710.

Studies in the Linguistic Sciences
Volume 23, Number 2, Fall 1993 [publ. October 1996]

VOWEL DELETION IN PULAAR: RIME AND NUCLEAR MERGERS AND THE ISSUE OF THE SYNTAX-PHONOLOGY INTERFACE*

Abdul Aziz Diop

In this paper I am going to analyze vowel deletion (VD, henceforth) in Pulaar, the Mauritanian dialect of Fula. The analysis has two parts. First, after some background discussion of Pulaar vowels and Pulaar syllables I present some data and suggest a phonological analysis in the form of rime and nucleus mergers. Second, I present data that suggest that in establishing its domain of application, VD is sensitive to syntactic information. In the literature on syntax-phonology interface we have two main approaches: the direct-access (Clements 1978; Kaisse 1987; Odden 1990), and the indirect-access (Hyman 1990; Selkirk 1986, 1987, 1990). In this paper I will demonstrate that the generalization about the Pulaar data is not consistent with the basic tenets of either approach and that an additional statement in the formulation of the rule, along with either approach, will account for the data. There is a third approach in the interface between syntax and phonology that space would not permit to go into, however. It is raised in Kenstowicz (1987:229) and has to do with whether the application or blockage of (phrasal) phonological rules 'can tell us something about the surface syntactic structure — in particular something that we did not know already'.

1. Introduction

The aim of this paper is to provide an account for all the conditions under which vowel deletion and vowel spreading take place in Pulaar. The paper is organized as follows: first, after a brief introduction of the vowel system I present the phenomenon from a descriptive standpoint in order to show which vowel deletes, and what the output of deletion is; then I give some background information pertinent to the topic of the paper. Second, I present my analysis of VD in Pulaar. Third, I present two sets of data that seem to be problematic for my analysis in the sense of the rule failing to apply. For the first set I show that failure of the rule to apply has to do with prosodic information whereas with the second set I propose a solution within the syntax-phonology approach to domain definition. I conclude by demonstrating that the two major approaches to this theory of domain definition, in their current formulation, cannot handle the Pulaar data.

2. Pulaar vowel deletion: a description

2.1. The vowels of Pulaar

First, an inventory of Pulaar vowels. Pulaar has five phonemic and seven phonetic vowels illustrated below in (1a) and (1b), respectively.

(1) Pulaar vowel inventory
a. Phonemic: /i/, /u/, / ɛ/, / ɔ/, and /a/
b. Phonetic: /i/, /u/, /e/, / ɛ/, /o/, / ɔ/, and /a/

Each of these vowels has its long counterpart but the distinction between a long vowel and short vowel is unpredictable, as illustrated in (2) where the meaning of two otherwise similar words only differs because they have one vowel realized as short in one word and long in the other member of the pair.

(2) Contrastive vowel length in Pulaar

ñol-de	'to be rotten'	ñool-de	'to win'
luɓ-de	'to lend'	luuɓ-de	'to smell bad'
hir-de	'to be jealous'	hiir-de	'to be late (sp.)'
nan-e	'left-plural'	naane	'earlier'
fere	'expense'	fee-re	'manner (spec.)'

In Pulaar the only environment in which a long vowel is predictable is when /h/ or the glottal stop (/ʔ/) is deleted (in coda position) causing the preceding vowel to lengthen. (3a-b) illustrate this. There are no complex onsets or codas in the language. In (3a) we have nominal roots followed by consonant-initial noun class agreement markers. The /h/ or the glottal stop /ʔ/ deletes and its mora is assigned to the preceding short vowel, making it realized as a long vowel. In (3b) the same roots are used either with consonant-initial noun class agreement markers whose initial consonants have been deleted (cf. Paradis (1986, 1992) for a discussion and an analysis of such initial deletion) (I call these vowel-initial markers for expository purposes), or with (vowel-initial) aspectual markers. The /h/ or /ʔ/ are then syllabified as onsets, not as codas. They do not delete in this position; therefore the vowel that precedes them does not lengthen.

(3) Predictable vowel length in Pulaar
a. Root + consonant-initial markers

/wah-re/	[waa-re]	'beard'
/yah-re/	[yaa-re]	'scorpion'
/mah-de/	[maa-de]	'to build'
/sah-de/	[saa-de]	'to fry'
/faʔ-de/	[faa-de]	'to be headed for'
/hoʔ-re/	[hoo-re]	'head'
/fiʔ-de/	[fii-de]	'to beat up'

b. Root + 'vowel'-initial markers

/wah-e/	[ba-he]	'beard'-pl.
/mah-i/	[ma-hi]	'build'-past
/sah-i/	[sa-hi]	'fry'-past
/faʔ-i/	[fa-ʔi]	'be headed for'-past
/fiʔ-i/	[fi-ʔi]	'beat up'-past

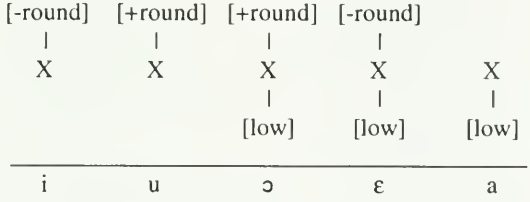
Long/short vowels can occur freely in the word, as shown by (4). The representation we give for Pulaar vowels is as in (5) (Goldsmith 1990) where (5a) is the underlying phonological representation and (5b) the redundancy rule that captures the fact that the feature [back] is predictable for Pulaar vowels.

(4) Distribution

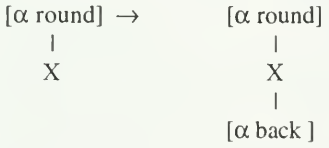
Beginning	**Middle**	**End**
ɛkkaadɛ	dɛlɛp	harɛ
ɛɛwnaadɛ	faliima	kataa

(5) Representation of vowels (Goldsmith 1990)

a. Representation



b. Redundancy rule



Having gone through the vowel inventory of Pulaar I now present vowel deletion in its descriptive form and discuss some of the syllable-related issues that are central to the analysis given in this paper. (For the remainder of this paper I will not distinguish between [+ATR] and [-ATR] vowels orthographically because they are irrelevant for the present purposes.)

In Pulaar, vowel deletion is observed only across morpheme boundaries. It does not take place within the morpheme itself because intramorphemically one never finds sequences of different vowels. In a configuration where two morphemes are 'adjacent'; given that the first morpheme ends in a vowel and the second one starts with a vowel the final vowel of the first morpheme deletes provided the latter is not a major lexical category, i.e. a verb, a noun, or an adjective. Typically, the morphemes in second position are: subject pronouns, conjunctions, prepositions, and the vowel-initial possessive pronoun /am/ 'my'. When the final vowel of the first morpheme deletes (across morpheme boundaries) it is the features of the initial vowel of the second vowel that are 'kept'. This is the reason why I refer to the phenomenon as deletion rather than coalescence. However, I will not dwell on this point as the data will illustrate it even better. In cases where the first morpheme (alpha) is consonant-final its final consonant becomes the onset of the initial vowel of the second morpheme (beta). In section 2.2., (6)-(8) below I de-

scribe the output of vowel deletion and the relevant data (the target vowels are in bold).

2.2. The output of vowel deletion in Pulaar

A long vowel is produced either when the (short) vowels of the two morphemes involved are identical, when the vowel of the first morpheme is high (or high and long), or when final /o/ of the first morpheme is deleted in front of /a/ of the second morpheme, as shown by (6a-f).

(6) a. hannd**e-e**-janngo → hanndee janngo
today-conj-tomorrow
'today and tomorrow'

b. Mali**-e**-Moritani → Maalee Moritani
Mali-conj-Mauritania
'Mali and Mauritania'

c. ɓayri**-o**-yim-at → bayroo yimat
since-3sg-sing-Asp.
'since he sings'

d. sab**u-a**-yim-at → sabaa yimat
because-3sg-sing-Asp
'because you sing'

e. o-yah-**ii-e**-meere → a yahee meere
3sg-go-Asp-for-nothing
'he went for nothing'

f. o-wii-k**o-a**-yah-ii → o wii kaa yahii
3sg-said-that-2sg-go-Asp
'he said that you went/left'

A short vowel obtains when the vowel of the second morpheme is part of a closed syllable, as shown by (7a-b).

(7) a. o-wii-k**o-on**-njah-ii → o wii kon njahii
3sg-said-that-2pl-go-asp *koon
'he said that you went/left'

b. o-wii-k**o-en**-njah-ii → o wii ken njahii
3sg-said-that-1pl-go-asp *keen
'he said that we went/left' *koon

We get a vowel-glide sequence when we have the following (mid) vowel combinations: /a/ + /ɛ/ = *ay*; /ɔ/ + /ɛ/ = *ɔy*; /a/ + /ɔ/ = *aw*; /ɛ/ + /ɔ/ = *ɛw*, as illustrated by (8a-d). From these combinations the generalization to be drawn is that we get glide formation in Pulaar if a non-high non-low vowel is preceded by an non-identical non-high vowel. This generalization is formalized in (8e) within an SPE (Chomsky & Halle 1968) type of framework. What (8e) says is that a low vowel will become non-consonantal and non-vocalic (therefore, a glide) in the environment where it is preceded by another low vowel which differs in its rounding specification. In case the first morpheme is consonant-final the vowel of the second morpheme is syllabified with the final consonant of the first word for

which it forms a nucleus, as illustrated in (8f) below (a dot indicates a syllable break).

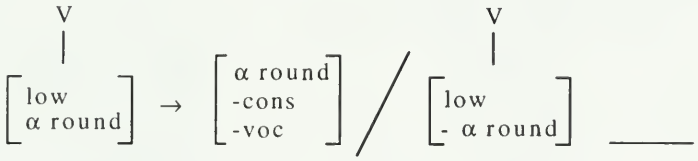
(8) a. haala-**e**-kawr-al → haalay kawral
talk-conj-agreement-NCagr
'talk and agreement'

b. wuro-**e**-Fuuta → wuroy Fuuta
city-prep-Fuuta
'a city in Fuuta (Region)'

c. ma-**o**-yah → maw yaa
future-3sg-go
's/he will go'

d. nd**e**-**o**-yah-i → ndew yahi
when-3sg-go-asp
'when s/he went/left'

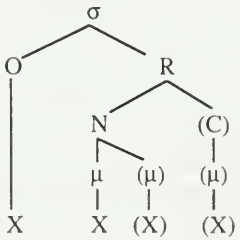
e.



f. Kan-**e**-sehil-mum → ka.ne sehilmum
Kan-conj-friend-3sgposs.
'Kan and his friend'

Having shown the various patterns attested so far, I am going to discuss some of the syllable-related issues that are pertinent to the analysis that would be proposed for the data in (6)-(8). First, the syllable types. In Pulaar, on the surface, the following syllable types are found: CV, CVV, CVC, and CVVC. The syllable template that I assume in this paper is as proposed in Diop (1993) and shown in (9) below where aspects of both moraic phonology and X-slot theory are used (cf. Diop 1993 for further discussion of this model).

(9) Pulaar syllable template



The first syllable-related issue I wish to address here has to do with weight. Following Hayes (1987; 1988; 1989) I represent vowels (and geminates) with an underlying mora whereas single consonants (in coda position) acquire a mora by

virtue of the Weight by Position (WBP) Rule (Hayes 1987; 1988; 1989) which, on a language-specific basis assigns a mora to a consonant in coda position if such mora assignment does not violate the upper-bound limit on the number of moras per syllable imposed by the language in question. The motivation for saying that WBP applies in Pulaar is as follows. There is a rule in Pulaar that shortens long vowels (in certain positions) when they are followed by a heavy syllable (cf. Diop (1993); Paradis (1986, 1992); Prunet & Tellier (1984)). This rule applies when the syllable following the long vowel is either CVV(C) or CVC, and never before CV. This is an indication that with respect to vowel shortening CVC counts as heavy (bimoraic) in the same way CVV or CVVC does. However, while CVV occurs rather freely, CVVC is more restricted in that in final position its coda consonant is either /n/ or /r/, as shown in (10a) (where the dots indicate syllable breaks). When closed by another consonant in final position a vowel is suffixed to that consonant, as shown by (10b). Furthermore, CVVC is never derived by any process, phonological or otherwise.

(10) CVVC in word-final position

a.	mbin. daan	'maid/butler'
	Hal. waar	'name of a village'
	mon. toor	'watch'
	dii. waan	'place/area'
	mi. soor	'headscarf'
	ti. su. baar	'prayer time'
	kaf. taan	'type of dress'
b.	tuu. baak-o	'European'
	European-Noun class agr.	
	ca. paat-o	'Moor/Arab'
	moor-Noun class agr.	

Syllable weight and the fact that CVVC is never derived is relevant to VD in that in a configuration like that shown in (7a-b) above where the vowel of the second morpheme is closed by a consonant (that consonant is moraic by virtue of the WBP rule that I mentioned earlier) the rule of deletion is going to avoid deriving a long vowel because this will result in a CVVC where the last /C/ is moraic; therefore a trimoraic structure, violating thus the upper limit of two moras per syllable in Pulaar. This is not to say that CVVC is always trimoraic. In fact it is my claim that non-derived CVVC is not trimoraic and here is the evidence. If CVVC were to be treated as trimoraic then when its coda consonant is deleted one would expect compensatory lengthening to take place and derive a triply long vowel. This does not happen. For instance, the Arabic word /Wallaah/ has been borrowed into Pulaar; but when the /h/ got deleted (because of the restriction I discussed above against coda /h/) what we have is /wallaayi/ instead of */wallaaa/, which is expected if CVVC syllables were treated as trimoraic by Pulaar speakers (Cf. Diop (1993) for more discussion).

The next issue is that of glottal insertion. In Pulaar, vowel-initial morphemes trigger a glottal insertion. Major lexical category items (nouns, adjectives, adverbs, and verbs) systematically do so. For example, in the data in (11a) below the

words are underlyingly vowel initial. However, each of them is pronounced with a glottal stop before the initial vowel, as illustrated by (11b).

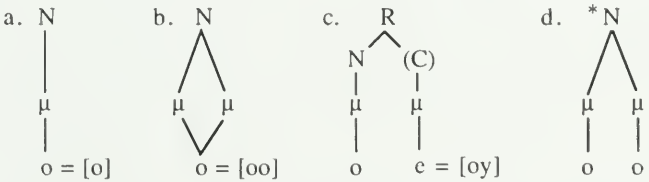
(11) Vowel-initial words in Pulaar

a.	ekk-o	'apprenticeship'
	am-re	'turtle'
	in-nde	'name/naming ceremony'
	oto	'car'
	uur-de	'to smell good'
b.	ʔekko	
	ʔamre	
	ʔinnde	
	ʔoto	
	ʔuurde	

What is then the relevance of this to VD? The answer is a question. If vowel-initial words are realized with a glottal stop on the surface, why then would the vowels of morphemes like /e/, /on/, /a/, /en/, etc. in (6)-(8) above be syllabified with the final syllable of the preceding morpheme? Therefore, one would have to say that glottal insertion affects only major lexical category items such as nouns, verbs, and adjectives. That is not quite adequate because the morphemes whose (initial) vowel gets resyllabified with the previous morpheme, when they appear in sentence-initial position are also pronounced with a glottal stop. So, /a/ (you-singular) in the sentence /a-jal-ii/ 'you-laugh-ed' is pronounced /ʔa/. Glottal insertion will then be permitted even with lexical categories other than nouns, verbs, and adjectives if those are in sentence-initial position.

The next and last issue has to do with a proposal (Diop (1993), cf. same reference for more discussion) that moraic phonology have a rimal tier. Thus, (12a-c) show, respectively, the representation of a short vowel (/o/), a long vowel (/oo/), and a vowel-glide sequence (/oy/) within the different syllable-internal nodes assumed in this paper (only the relevant part is shown for each representation). (12c) illustrates the fact that I treat the vowel-glide sequence as being the direct 'product' of syllabification whereby the first vowel is syllabified in the nucleus whereas the second vowel is syllabified outside it. (12d) will be ruled out as a possible representation for /oo/ because it is a violation of OCP (McCarthy 1986; Hayes 1986; Odden 1986).

(12) Syllabification of vowels and vowel-glide sequences



In the foregoing I have discussed some background information pertinent to the analysis of VD in Pulaar. Here is, again, a summary of the patterns observed so

far in (6)-(8) (overlooking the rather unproblematic case of (8e)). We had a pattern deriving a long vowel from two short identical vowels. A long vowel was also derived when we have a vowel preceded by a high (or high and long) vowel. Another case where the rule yielded a long vowel is when /a/ was preceded by /o/.

The second pattern was when the vowel of the second morpheme was closed by a consonant; syllabifying that vowel with the preceding morpheme produced a short vowel (and a consonant) rather than a long one.

The third pattern was that in which a vowel-glide sequence was created. In that pattern, for the /VW/ sequence we had the combinations in (13a) whereas for the /VY/ sequence the combinations are illustrated by (13b) below.

(13) Vowel-glide sequences in Pulaar: a reminder

a. V-W sequences

/a/	+	/o/	= /aw/
/e/	+	/o/	= /ew/

b. V-Y sequences

/a/	+	/e/	= /ay/
/o/	+	/e/	= /oy/

The challenge presented by the data for any analysis of vowel deletion would to predict each of the different outcomes just outlined. Furthermore, in (6e) above we have a case where a short vowel (monomoraic) spreads to a position previously occupied by a long one (bimoraic). Yet we do not get a trimoraic syllable. What we derive is a long (bimoraic) vowel. One has to account for what happened to the third mora. This case is somewhat similar to that in (7a-b) where a trimoraic structure is also reduced to a bimoraic one. But, as we shall see, they are treated differently. Another challenge for any analysis of VD would also be to predict the occurrence of vowel-glide sequences. In the section to follow I am going to give the solution that I propose for VD.

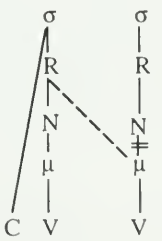
3. A phonological analysis of vowel deletion in Pulaar

In (5) above I gave a representation of Pulaar vowels and a redundancy rule that captures the fact that the feature [back] is predictable in Pulaar. Given the Pulaar syllable facts I outlined at the beginning of the paper I also take it to be the case that syllable structure is assigned lexically and post-lexically. Thus, following the operation of certain phonological or morphological rules the output of such rules can feed syllabification. Syllabification can take place across morpheme boundaries. So, in (8f) for instance, the conjunction /e/ 'and' is pronounced with the final consonant of the previous morpheme (i.e. /n/) as its onset; yet these two segments came from two different morphemes. My analysis, however, focuses more on the cases where two vowels (instead of a consonant plus vowel) merge. In such cases VD takes the form of a rime merger that is illustrated in (14) below. Referring to the syllable to account for phonological processes involving vowel mergers is not a novel idea. It has been proposed in Schane (1987). Although Schane's framework and the one I propose here have some similarities they differ in ways that I do not intend to discuss in this paper since it is not my intention to

compare the two frameworks. Nevertheless, I will spend some time summarizing Schane's theoretical framework. Schane's framework (which he uses to analyze hiatus in Sanskrit and Chicano Spanish) is one in which the autosegmental representations contain three tiers: a syllable tier, a CV tier, and a segmental tier. The first tier depicts the number of syllables while the second gives information about the quantitative characteristics of phonological units. The quality of the phonological units is determined by the third tier. In addition to these tiers Schane uses the notion of closure (merger) at each of these tiers. In the framework that I propose syllable count is not regarded as crucial. The quantitative characteristics of phonological units is determined by the mora whereas the rime and nucleus node determine the quality of phonological units. So, any phonological unit couched within the nucleus is going to be realized as a vowel whereas anything outside the nucleus is going to have consonantal status. As I said earlier the rule is formulated in terms of a rime and nucleus merger. The first rule (the rime merger rule), responsible for glide formation, because it has a more specific environment, is going to apply first following Kiparsky's (1973b; 1982a) Elsewhere Condition. It is illustrated in (14a) where it takes the initial vowel of the second morpheme and syllabifies it at the rime node level of the preceding syllable, giving a vowel-glide sequence for reasons that I explained earlier. (14c) is a derivation illustrating (14a) whereas in (15) we have the general rule that syllabifies vowels under the nuclear node. It accounts for the more general cases, (6)-(7). The delinking of the association line in the second syllable causes that syllable to collapse, making its rimal content available for the merger. As a reminder, the combinations that yield glide formation in Pulaar are shown in (14b). The generalization was that we get glide formation if a non-high non-low vowel is preceded by a non-identical non-high vowel. When such is the case the rime merger rule in (14a) syllabifies the second vowel within the rime of the first syllable, turning it into a glide. There is no contrast between high and mid glides (at least in Pulaar). Whether we get the labial glide /w/ or the palatal glide /y/ depends on the rounding specification of the second vowel. Where that specification is [+] we get /w/; and we get /y/ where it is [-]. In (14c) we give a derivation for /Sammbay Zeynabu/ < /Samba e Zeynabu/ 'Sammba and Zeynabu'. (14d) shows the surface representation of (14c). (Only the relevant parts are syllabified for /Sammba/.)

(14) Pulaar rime merger: a two-step process

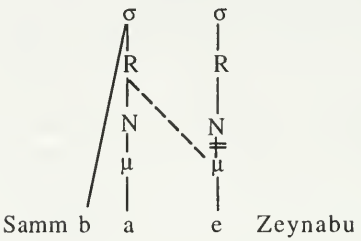
a. The specific rule: glide formation: rime merger



b. Pulaar glide formation

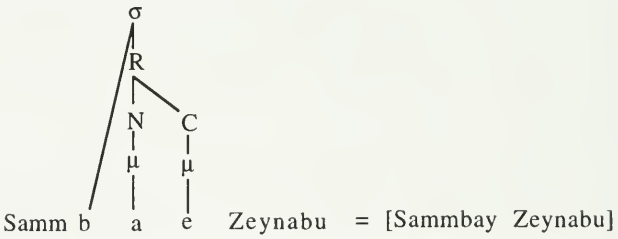
a + ɛ = /ay/	ɔ + ɛ = /ɔy/
a + ɔ = /aw/	ɛ + ɔ = /ɛw/

c. Derivation for [Sammbay Zeynabu]:



Glottal insertion _____
rime merger: specific applies
rime merger: general _____

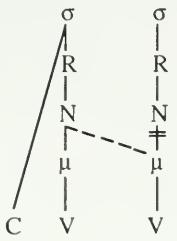
d. Surface representation of (13b) above:



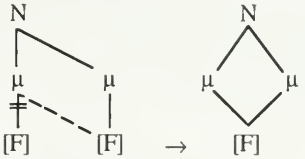
Having illustrated how the glide formation takes place I am now going to account for the general case: short/long vowels. The merger rule in (15a) below (ordered after (14a)) syllabifies the second vowel under the nuclear node. Following that we have a deletion rule (affecting the first vowel) and a root node spreading rule from the second vowel to the position formerly occupied by the first vowel. This second rule affects not only cases where the vowels involved are not identical but also those cases where they are in fact identical (in the latter case the rule may be considered as having the same effect as OCP). As said earlier, a set of vowel features on two (adjacent) moras under the same nuclear node gives a long-vowelled nucleus. The deletion and spreading rules are shown in (15b) below whereas (15c-d) show the derivation for [lewree koode] < /lewru e koode/.

(15) The Elsewhere rule:

a. nucleus merger

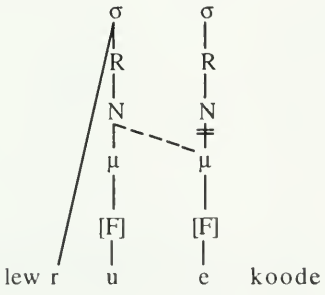


b. Feature deletion and root node spreading (FD and RNS)



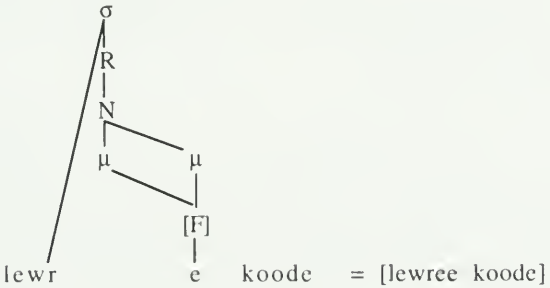
(where [F] is the root node)

c. Derivation for [lewree koode] < /lewru e koode/



Specific rule	_____
General rule	applies
FD and RNS	applies
Syllabification	applies

d. Surface representation of (15c)

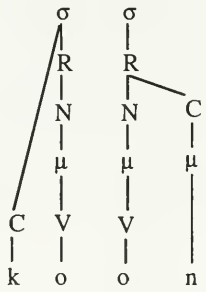


So far I have accounted for two cases: cases where VD yields a long vowel and those which yield a vowel and a glide. I now go on to the third case, i.e. where a short vowel results from VD. The relevant data are in items (7a-b) above

and they are reproduced here as (16a) for convenience (the target vowels are bold). In (16a) we have a monosyllabic (and closed syllable) pronoun /on/. The initial vowel of the second morpheme and the final vowel of the first are identical. Still, a long vowel is not produced. There are at least two ways in which the data in (16a) can be explained. The first tack we can take is as follows. Assuming that there is syllable structure prior to and following VD, also taking into account the fact that the WBP rule applies in Pulaar, the relevant part of (16a) can be represented as in (16b).

(16) a. o-wii-k**o**-**o**n-njah-ii → ʔo wii kon njahii
3sg-say-that-2pl-go-Asp *koon
'he said that you went'

b.



To the second syllable of (16b) we cannot apply glottal insertion because the morpheme is neither sentence-initial nor a major lexical category item. So, it is syllabified with the preceding syllable, causing a trimoraic structure. This structure is reduced to two moras (by virtue of a mora deletion rule that applies whenever a trimoraic structure is created) because Pulaar avoids trimoraic syllables (as said earlier), and never derives them.

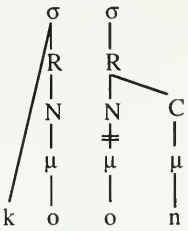
This solution is not all that desirable (though probably unavoidable in some cases; as will be shown later) as it relies on something that does not look quite natural in phonological theory, creating a structure and then erasing it in order to arrive at the correct derivation.

We can do away with such an intermediary stage that creates three moras. The second approach is going to do just that; and for that reason it is the approach we adopt here. In this approach the syllable of the second morpheme /on/ (prior to the application of the rime merger process) collapses following the delinking of the association line from the mora to the nuclear node because a Pulaar syllable can only exist if it has a nuclear mora (i.e. a mora dominating a vowel). The mora of the nasal, naturally, is 'deleted' because at that stage there is no syllable structure and the nasal is no longer in rime position to receive a mora by virtue of the WBP and there are no syllabic nasals in Pulaar. The steps just outlined are illustrated in (17a-b) below. At this point in the derivation we proceed with syllabification. Following the rules that we postulated earlier ((14) above) the vowel of the second

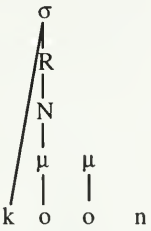
morpheme is syllabified in the nucleus node of the final syllable of the first morpheme. So, it is the general, not the specific one, that applies. Feature deletion and root node spreading apply, as illustrated by (17c) below.

(17) Syllable 'collapse'

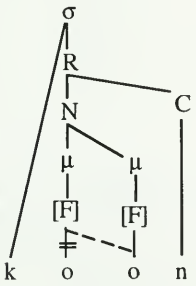
a.



b.



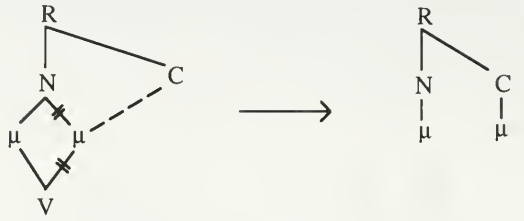
c.



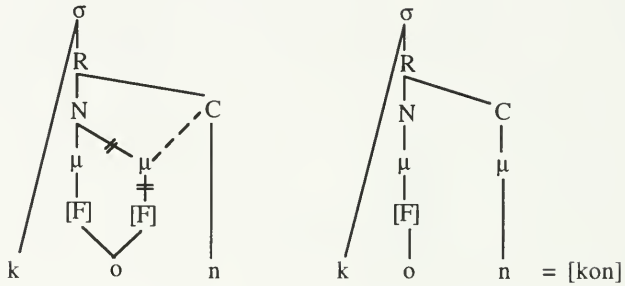
Earlier, I claimed, without elaborating on it much, that in Pulaar a consonant in coda position must be moraic. A piece of evidence for that was that long vowels shorten before CVC syllable in the same way they do before CVV(C) ones. However, WBP can only apply if its application does not violate the upper bound limit of two moras per syllable. Therefore, in (16c) WBP cannot apply. Since WBP cannot apply the second mora is then donated to the rimal consonant leaving the first vowel with one mora (therefore realized as a short vowel). This way we derive a bimoraic single short vowel syllable. In (18a) (next page) we illustrate mora donation whereas (18b-c) (next page) show the last and final steps of the derivation, including mora donation.

I have just given an account of how VD works when the second vowel-initial morpheme involved in the process is monosyllabic. Vowel-initial bisyllabic pronouns show similar behavior to that of monosyllabic pronouns in that when their initial is syllabified with the previous morpheme a short, not a long, vowel is derived. These pronouns, illustrated in (19) next page, are problematic for the analysis given for the monosyllabic pronouns in that while in the data in (16-18) we had single syllable morphemes (pronouns), in (19) we have bisyllabic morphemes as the second morpheme (e.g. /oɗon/). Following the syllabification

(18) a. Mora donation



b. c.



principles (both universal and parametric), the medial consonant in /oɗon/ is syllabified as the onset of the second syllable instead of the coda of the first. Consequently, it is mora-less. Therefore, the first syllable is monomoraic. Merging it with the preceding monomoraic syllable should not lead to any mora deletion or donation of some sort since bimoraic syllables are accepted in Pulaar. The issue, then, is whether the approach used to explain the data in (16-18) can be improved to accommodate the data in (19) or whether a whole new approach is needed for the new set of data. Clearly the latter is not desirable. We do not want to have a multitude of different rules and approaches to account for what might be one and the same phenomenon. We adopt the first suggestion then; i.e. improve the first approach in order for it to explain the data in (19). (Once again, whether the two vowels involved in the process are identical or not is irrelevant because we are always going to have a short vowel.)

(19) Bisyllabic pronouns and VD

a.	o-wii-ko-oɗon-njar-a	=	o wii koɗon njara
	3sg-say-that-2pl-drink-Asp		*kooɗon
	'he said that you drink'		
b.	ɗo-omo-yah-a-fof	=	ɗomo yaha fof
	'pl.Adv-3sg-go-Asp-every'		*ɗoomo
	wherever s/he goes		

In order to put things into perspective a brief survey of Pulaar pronouns is necessary. (20a-c), (21), (22), and (23) illustrate these.

(20) Pulaar subject pronouns: the one-syllable set

a. Preposed

	Singular	Plural
1	mi	en / min
2	a	on
3	o	ɓe

Example

mi-yah-ii 'I went'
1sg-go-Asp

b. Postposed

mi	ɗen / en
ɗaa / aa	ɗon / on
---	---

Example

njah-mi 'I went' (focus construction)
go-1sg

c. The two syllable set: never postposed

miɗo	amin / emin
aɗa	oɗon
omo	eɓe

Example

miɗo-yah-a 'I go' (habitual)
1sg-go-Asp

(21). Object pronouns

kam / mi	men / en; min
ma / maa	mon / on
mo / moo	ɓe

(22). Independent/focus pronouns

miin	minen / enen
aan/maa	onon/mon
kaŋko/makko	kamɓe/maɓɓe

(23). Possessive pronouns

am	amen / men
maa	mon
makko / iiko / mum / um	maɓɓe

In Pulaar /on/, /en/, /o/ and /a/ are subject pronouns. They correspond, respectively, to the second part of the two-syllable pronouns in (20c). The first part in each of these pronouns (i.e. /oɗ/, /eɗ/, /om/, /aɗ/) is neither a pronoun nor an attested (synchronic) prefix of some sort. All things being equal we take it to be the case that in each of the pronouns in (20c) a highly morphologized phonological rule suffixes the pronoun to a closed syllable 'morpheme'. That being said, we also take it to be the case that the rime merger rule will take place before this process of pronoun suffixation. This way the process in (20c) is similar to that in (16)-(18)

in that we have a closed-syllable vowel-initial morpheme (/oƈ/, /eƈ/, /om/, and /aƈ/) merging with a vowel-final syllable and the same solution (not repeated here) can account for both cases.

The solutions proposed so far seem to predict and explain the data presented thus far on VD in Pulaar. However, there is a set of data that seem to be problematic for our analysis. In (24a-g) below we have sentences in which VD could take place, given that the context of (14) above is met, but indeed does not. (A slash between two vowels indicates that the rule fails to apply and that the vowel is pronounced with a glottal stop.) The data are organized in pairs that show a contrast between two cases. On the one hand we have cases where the first morpheme ends in a long vowel that does not delete. The next example will show the same vowel at the end of a morpheme where that vowel is short. In that case the second vowel is syllabified with the first to form a vowel-glide sequence. This pattern is found, respectively, from (24a-f) where (a) contrasts with (b), (c) with (d), so on. (24f) illustrates the fact that long vowels at the end of a morpheme do indeed allow the rule to operate.

(21) Vowel deletion and prosodic information

a. Muusa**a-/-e**-debbo-mum
Moses-coord.-woman-his
Muusaa and his wife
=Muusaa ʔe debbomum *Muusaay

b. Rama**-e**-Abu
coord.
Rama and Abu
= Ramay Abu

c. njol-ƈ**aa-/-e-**oto-makko
enter-2sg-prep-car-his/her
you entered (in) his/her car
= njolƈaa ʔe otomakko *njolƈaay

d. njol-m**i-e**-oto-makko
enter-1sg prep car-his
I entered (in) his car
= njolmee otomakko

e. Bus**oo-/-e-**sehil-mum
coord.-friend-his
Busoo and his friend
= Busoo ʔe sehilmum *Busooy

f. Don**o-e-**miñ-um
coord.-young-sibling-his
Dono and his younger sibling
= Donoy miñum

g. mi-ar-**ii-e**-meere
1sg-come-Asp-prep.for-nothing
I came for nothing
= mi aree meere

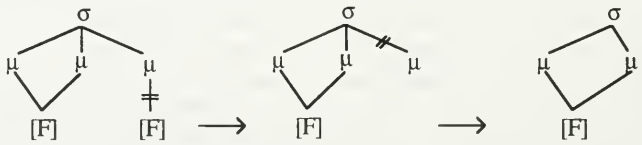
Our hypothesis is that prosodic information is responsible for failure of the rule to apply in these cases. The explanation for the 'apparent' problem in (24) above is as follows. First, notice in Pulaar (as just said) that the initial vowel of a second morpheme can be syllabified with the final syllable of the previous morpheme even when the latter ends in a long vowel (cf. 24g). Second, notice also that in all the cases where VD fails to apply (namely, in examples (24a, c, and e) the long vowel of the first morpheme is either /a/ or /o/. As shown in (8) and (13) above, /a/ or /o/ + /e/ give, respectively, the sequences /ay/ and /oy/, a closed syllable. A sequence of one vowel and a coda consonant is treated as bimoraic in our approach to the Pulaar syllable. In the data in (24), where the rule fails to apply we have long vowels, instead of short vowels as host, except for (24g). So, in (24a) for instance we have /aa/ + /e/; which, in theory, should yield the sequence /aay/ (*Muusaay Abu). However, if this was to be the case we create a (superheavy) trimoraic syllable (viz. /saay/). As said earlier this type of syllable in Pulaar is not desirable (especially 'word'-finally); nor is it derived by any phonological process. The prohibition against trimoraic syllables acts then as a filter on all phonological operations that involve syllabification to the extent that it prevents one from deriving super heavy syllables. So, in (25) we have a case where /e/ could be brought under the rime node subject to being 'rejected' from under it by the moraic filter as its syllabification under that rime will create a trimoraic structure. Once it is 'rejected' from the previous syllable it has to form a syllable on its own, triggering thus the glottal insertion rule (mentioned earlier without much formalism) that provides a default onset to vowel-initial morphemes. Consequently, we have another environment where such glottal insertion takes place. First we said it took place sentence-initially. We have to add to that another environment; namely, after the application of the moraic filter (in phrase-initial position). To end this section I am going to present our analysis of the data in (24g). The problem presented by (24g) is as follows. The first morpheme has a long (therefore bimoraic) vowel. The second morpheme is a single vowel (monomoraic) syllable. Deleting the final (bimoraic) vowel of the first morpheme and spreading the second (monomoraic) one in this instance will inevitably give rise to a trimoraic structure, especially since the mora donation rule in (18a) cannot be applied here as there is no coda consonant to yield the third mora to (unlike the case of (17a) above). It looks, then, that in deriving (24g) we have to go through a stage at which a trimoraic structure is created. However, the difference between this case and the one in (17a) is that in the latter there was no reason to believe that the nasal carried a mora throughout the derivation because when the syllable structure in which it was found collapsed the mora was then lost because consonants only receive a mora by virtue of the WBP, in the theoretical framework that we adopt here. In (24g) this argument cannot be made. Instead, we propose the following solution. First, some (by now familiar) assumptions. In Pulaar, there are no sequences of three vowels (or consonants), identical or not. So, this gap can be expressed in terms of a Pulaar-specific prohibition against ternary branching for vowel (consonant) features. In addition, the restriction against trimoraic structures in Pulaar (*σ+2μ, i.e. no syllable can have more than two moras) acts as a filter on all phonological operations involving syllabification to the effect that it prevents derivation of syllables that

violate the upper bound limit of two moras per syllable. When this limit is violated it triggers a deletion of the third mora. In (24g) the issue will be to decide which mora to delete and how to make it candidate for deletion. (24g) has been repeated here as (25) as a reminder.

(25) mi-ar-ii-e-meere = mi aree meere
1sg-come-Asp-Prep.for-nothing
I came for nothing
* areee

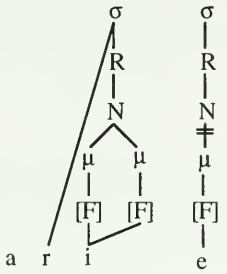
To the vowel /e/ the glottal insertion rule fails to apply, so the rime merger rule (the elsewhere rule) applies syllabifying it under the nuclear node of the syllable containing the long vowel. Following this we apply the feature deletion rule deleting the features of the long vowel /ii/. We have, then, two moras without features. The root node spreading rule applies, spreading the features of the vowel of the second morpheme to one of the moras of alpha given that vowel features can only be binary branching, at the most. This leaves us with one mora that does not dominate any features. Ito (1986) introduced the notion of Prosodic Licensing that says that 'phonological material must be incorporated into the next higher level of prosodic structure'; otherwise, 'it is deleted by Stray Erasure' (Steriade 1982, Harris 1983). As mentioned by Hayes, 'a natural extension of this principle would state that higher-level phonological elements, such as moras, are also subject to Stray Erasure if they fail to dominate any lower-level element' (Hayes 1989:264). One of the moras of the host is exactly in that situation. Its features had been deleted by the feature deletion rule but it is not linked to any feature at the end of the derivation where syllabification takes place again. Therefore, we have a mora that does not dominate any features (vocalic or consonantal); so it is candidate for Stray Erasure. The mora deletion rule is illustrated in (26a) whereas the steps discussed so far are illustrated in (26b) below in a derivation for [mi aree meere] < /mi arii e meere/. Since the analysis here is quite complex (26b) will be done in five steps showing the different stages of the derivation and the surface representation. What (26a) says in essence is that a mora that does not dominate any segmental material is erased (since it would be impossible for it to be realized phonetically).

(26) Trimoraic mergers
a. Mora deletion

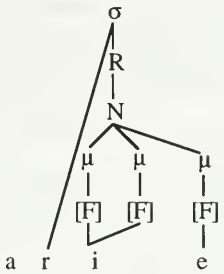


b. Derivation for [mi aree meere] < /mi arii e meere/
(only the relevant parts /arii e/ are shown)

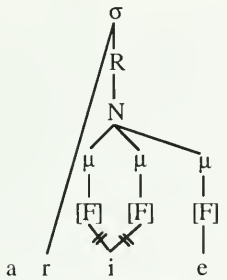
Step 1: syllable collapse



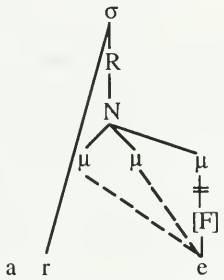
Step 2: nucleus merger



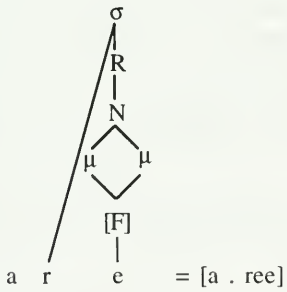
Step 3: feature deletion



Step 4: Root node spreading and mora deletion



Step 5: Surface representation



In the foregoing I have shown that the data in (24), which seemed to be problematic for my analysis turned out to be easily handled by it. I have, thus, accounted for what would have otherwise been considered exceptions to VD.

This is not all, however. We do have further cases where failure of the rule to apply cannot be attributed to prosodic information. I shall demonstrate that syntax is the reason for such failure. VD in Pulaar is sensitive to syntactic information in establishing its domain of application. In this last section we are going to try and

demonstrate that fact. First I present the data, followed by a presentation of basic facts about Pulaar syntax; then I present the syntactic structure in which the first and second morphemes are found; for instance /V-P/ meaning the first morpheme is a verb whereas the second is a preposition. I end this section by testing the data against two major approaches to syntax-phonology interface, the direct approach and the indirect approach, and showing that the generalization about the data escapes the predictions of either approach.

4. Vowel deletion and syntax

First, the data (a slash between two vowels indicates that the rule does not apply and the glottal stop with which the second vowel is pronounced in this case is shown when the example is repeated after the glosses).

(27) Vowel deletion and syntax

a. teew-ng**u,-/-a**-yiƈ-aa-ngu
meat-NC-2sg-like-Neg-Respr.
the meat, you don't like it
= teew ngu, ʔa yiƈaa ngu *ngaa

b. teew-ng**u-a**-yiƈ-aa-ngu
meat-Rel.-2sg-like-Neg-NC
the meat (that) you don't like
= teew ngaa yidaa ngu

c. rawaa-ndu-nd**u-/-e**-joom-um
dog-NCagr-NCDet.-coord.-owner-3sgPoss
dog the and owner its
the dog and its owner
= rawaandu ndu ʔe joomum *ndee

d. ndu-rawaa-nd**u-e**-joom-um
NCDet-dog-NCagr-coord.-owner-3sgPoss
this dog and its owner
=ndu rawaadee joomum

e. nd**u-/-e**-joom-um
NCDet-sing-coord-owner-3sgposs
this (one) and its owner
=ndu ʔe joomum *ndee

f. gor-k-o-mo-calmin-m**i-/-o**
man-ʔ-NCagr-Rel-greet-1sg-NC
man that greet I the
the man I greeted
= gorko mo calmin mi ʔo * moo

g. gor-k-o-mo-calmin-m**i-e**-mon-o
man-ʔ-NCagr-Rel-greet-1sg-Prep.-2pl. NC
man that greet I among you the
the man I greeted among you
= gorko mo calminmee mon o

h.	o-rokk-**i-/-on**-jawdi	
	3sg-give-Asp-2pl-wealth	
	He gave you (pl.) wealth	
	= o rokki ʔon jawdi	*rokkon
i.	**ko-/-enen**	
	foc.-1pl.	
	It's us	
	= ko ʔenen	*kenen
j.	**oto-am**	
	car-1sgPoss	
	my car	
	= otam	
k.	**mo-/-am**	
	ʔof-1sgPoss	
	that (of) my	
	mine	
	= mo ʔam	*mam

The data just introduced in (27) present some interesting and challenging problems. In the next paragraphs to follow I am going to outline what those problems are. First, (27a) and (27b) show that the morpheme /ngu/, a noun class in the first example and a relative clause marker (i.e. head of Comp) in the second, displays two different patterns in the sense of its vowel deleting in (27a) but not in (27b). (27c) and (27e) are other examples that illustrate a case of a noun class (or a determiner) whose final vowel does not delete whereas (27d), the 'mirror' image of (27c), shows that the final vowel of the head noun /rawaandu/ deletes. In (27f) we see that the vowel of the subject pronoun /mi/ does not delete when followed by the (open syllable) noun class (determiner of the head noun in the relative clause) whereas the same subject pronoun used in front of a following preposition /e/ (in (27g)) shows final vowel deletion. In (27h) the final vowel of the verb does not delete when followed by a vowel-initial object pronoun /on/ whereas earlier in (26b) we saw that final vowels of verbs can delete. In (27j) and (27k) we have another interesting alternation. In (27j) the vowel of the first morpheme deletes, allowing /am/ 'my' to be syllabified with it (cf./otam/) whereas in (27k) the rule is blocked between /mo/ and /am/.

The data presented in (27) show the first and the second morphemes (whose vowels are involved in the process of VD), occurring in the following syntactic configurations (28).

(28) Syntactic structures

Synt. Structure	Example number	rule applies?
NP-Conj	6a-b; 8a; 8f; 21b; 24f; 27d	yes
NP-Conj	24; 24e	no
C-NP	6c-d; 6f; 7a-b; 8c-d; 19a-b; 27b	yes
V-P	6e; 24d; 24g; 27g	yes
V-P	24c	no
NP-P	8b	yes

Det-NP	27a	no
Det-conj	27c; 27e	no
V-Det	27f	no
V-NP	27h	no
Focus-NP	27i	no
NP-poss	27j	yes
P-poss.	27k	no

Having presented the data, the nature of the alternations involved, and the syntactic configurations in which the two morphemes involved in vowel deletion occur, I am going to first outline some basic facts about Pulaar syntax before getting into the details of the analyses. In the basic word order of the Pulaar sentence subjects precede their predicate, objects follow verbs, as illustrated by (29) below.

(29) Pulaar basic word order
mi-yar-ii-kos-am-hanki
1sg-drink-Asp-milk-NCagr-yesterday
'I drank milk yesterday'

The Pulaar noun typically has a root and noun class agreement attached to that root that shows what class the noun belongs to (cf. Paradis (1986, 1992); Sylla (1982) for detailed studies of Pulaar noun classes). For instance, in (27) below the word for dog /rawaa-/ belongs to the /ndu/ class in the singular and to the /ɗi/ class in the plural. Consequently, it bears /ndu/ or /ɗi/ on its root. We can also add to it the diminutives /gel/, /kon/, or the augmentative /gal/.

(30) The Pulaar noun

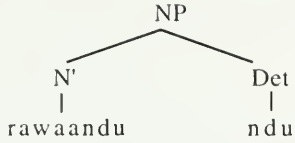
rawaa-ndu	rawaandu	dog
rawaa-ɗi	dawaaɗi	dogs
rawaa-gel	dawangel	small dog
rawaa-kon	ndawakon	small dogs
rawaa-gal	dawangal	big dog

Within the noun phrase the head noun (root + noun class agreement marker) is either in initial position followed by the noun class, as illustrated by (31a-b) below, or in phrase-final position preceded by the determiner (e.g. 'this'). (31c-d) is an illustration. The noun phrase can be null-headed, as illustrated in (31e) where /ndu/ is understood to refer to /rawaandu/ 'dog'. These null-headed NPs are represented in (31f).

(31) The Pulaar noun phrase
a. Full NPs

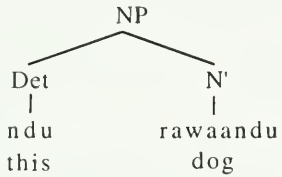
rawaa-ndu	ndu	= rawaandu ndu
dog-NCagr	NC	
dog	the	
the dog		

b. Representation



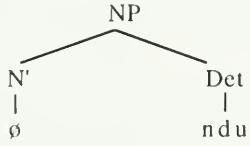
c. ndu-rawaa-ndu = ndu rawaandu
Det-dog-NCagr
dog this
this dog

d. Representation



e. Null-headed NPs
ndu-dog-ii = ndu dogii
NC(Det)-run-Asp
it ran (away)

f. Representation of null-headed NPs



Relative clauses are formed by using the noun class of the head noun (as a relativizer) or by using /mo/ for nouns which belong to the /o/-class (of humans and borrowed words). However, this /mo/ can only appear when we have a negative relative clause, as in (32a-b). Positive relative clauses where the head noun is subject (instead of agent) do not allow use of the relativizing pronoun, as shown in (32c). Instead, the verb shows agreement with the head noun in the sense of bearing an agreement marker that corresponds to the head noun.

(32) Relative clauses in Pulaar

a. gor-k-o-mo-yah-aan-i-o
man-ʔ-NCagr-Rel-go-Neg-Asp-NC(Det)
man who go not the
the man who did not go

b. gor-k-o-mo-rokk-u-mi-ndiy-am-o
man-ʔ-NCagr-Rel-give-Epen.-1sg-water-NCagr-NC(Det)
man who give I water the
the man I gave water to.

c. gor-k-o-ñaam-ɗo-o
man-ʔ-NCagr-eat-Agr-NC(Det)
man eat the
the man who ate
* gorko mo ñaamɗo o

Verb phrases are somewhat complex. The verb precedes the direct object, indirect object, and the prepositional phrase. Verbal complexes may be formed of a root plus a number of extensions that are subject to both ordering and co-occurrence restrictions. There are two object pronouns /moo/ 'him/her' and /maa/ 'you-sing' that are internal to the verbal complex in the sense that they occur before the postposed subject clitic which is also considered part of the verbal complex (cf. Paradis (1986, 1992); Prunet & Tellier (1984); Diop (1993) for further discussion). Other object pronouns are 'outside' the verbal complex in the sense of occurring after the postposed subject clitic. (33) below is an illustration of the different verbal complexes just mentioned.

(33) Pulaar verbal complexes

a. add-ii = ʔaddii
bring-Asp
brought

b. add-ii-jawdi = ʔaddii jawdi
bring-Asp-wealth
brought wealth

c. add-ii-e-jam = addee jam
bring-Asp-Prep.-peace
brought in peace

d. add-an-ii = addanii
bring-Ben.-Asp
brought for

e. add-an-oy-ii = ʔaddanoyii
bring-Ben.-Mvt-Asp
went and brought for

f. add-an-oy-moo-mi = ngaddanoymoomi
bring-Ben.-Mvt-3sgobj-1sg
I went and brought for him/her

g. add-an-oy-mi-on = ngaddanoymi ʔon
bring-Ben.-Mvt-1sg-2pl.obj
I went and brought for you (pl.)

In focus constructions the focus marker /ko/ appears before the focused NP (cf. Sylla (1982) for further discussion). If the NP is a pronoun then it has to be from the set given in (20c) or (22) above. Pronouns from those two sets always precede the verb. They are never postposed. (34) below illustrates focus constructions.

(34) Focused NPs in Pulaar

a. ko aan = ko ʔaan
It's you

b. ko-aan-e-makko = ko aane makko
foc.-2sing-Conj.-3sg
It's you and s/he

c. ko-aan-e-makko-yah-i-e-oto
foc.-1sg-Conj.-3sg-go-Asp-Prep.-car
It's you and s/he went in car
It's you and s/he who went by car
= ko ʔaane makko njahee ʔoto

Having presented the basic picture of Pulaar word order that is relevant to vowel deletion I am now going to discuss the two approaches to syntax-phonology interface. As I discuss each approach I will test it against the data given so far and point to the problematic cases. First, the direct-syntax approach (DSA, henceforth).

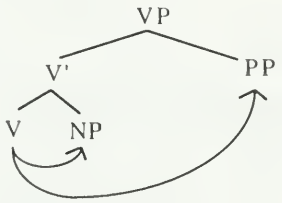
4.1. The direct-syntax approach. (Clements 1978; Kaisse 1985; 1987, 1990; Odden 1987).

In this approach an external sandhi rule applies between a sequence of two words α and β when it is the case that either the two belong to the same X^{max} or if some c-command relation holds between α and β, depending on what version of c-command one adopts. Two versions of c-command prevail. They are presented in (35) below (Sells 1985:39).

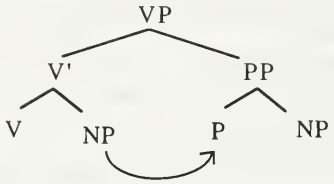
(35) C-command
α c-commands β iff:
a. every branching node dominating α dominates β
b. every XP dominating α dominates β

Under definition (35a) V in (36a) below c-commands NP but not PP whereas under (35b) V c-commands both NP and PP. By the same token NP c-commands P of PP in (36b) only by virtue of (35b) (as illustrated by the arrows).

(36) C-command illustrated
a.



b.

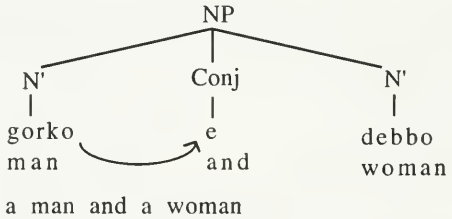


Given these two notions of c-command I am going to discuss what the predictions the direct access approach are going to be with respect to the Pulaar data. First, the simple cases. Assuming the c-command notions in (35) (within the DSA) we can formulate the rule of vowel deletion as in (37).

(37) Vowel deletion in Pulaar: preliminary version
Delete the final vowel of a word when it is followed by a vowel-initial word that it c-commands.

The rule in (37), using either version of c-command, is going to easily predict the data illustrating noun coordination where the first NP is not branching, therefore not necessarily dominated by XP. The same holds for cases where we have a CP whose head (C) is in a c-command relation with the specifier of the following IP. These two cases are schematically illustrated in (38a-b), respectively, whereas (39) gives a representative sample of the data presented earlier. (Throughout the rest of this section of the paper an arrow between two constituents means that the rule operates between them whereas a barred arrow means that the rule is blocked.)

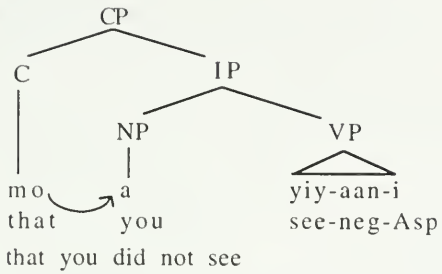
(38) Predictions of the DSA
a. Coordination (Jackendoff 1977:190)



a man and a woman

Data that fit this pattern are presented in (6a-b); (8a, f); (24b, f); and (27d);

b. CPs

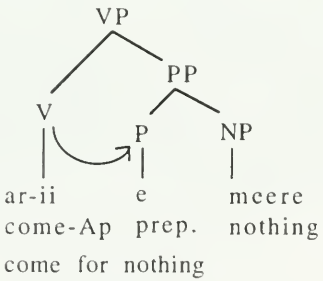


Data that fit this pattern are illustrated in (6c-d, f); (7a-b); (8c-d); (19a-b) and (27b).

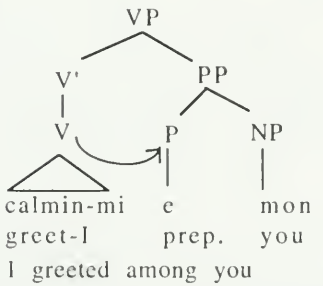
The rule in (37) also predicts (using either version of c-command) the data in which a verb(al complex) is followed by a preposition. The data are in (6e), (24d, g), and (27g) above and the relevant trees are drawn in (39a-b) below where the arrow shows the rule operating. In (39b) the rule operates between the verbal complex (verb root + postposed subject pronoun) and the following pronoun.

(39) Further correct predictions of c-command

a.



b.

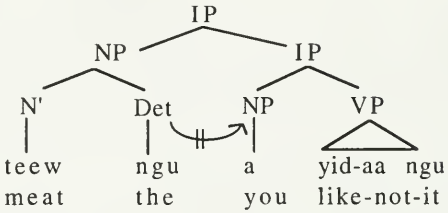


In addition to these data the rule in (37) correctly predicts that deletion (and spreading) do not occur in (27a, c, f). In all these data c-command does not hold

between the two morphemes involved, no matter which of the two versions in (35) above one is adopting. So deletion is blocked because the morphemes are not within its domain. In (40a-c) below I draw trees for (27a, c, f), respectively, to illustrate failure of c-command to hold.

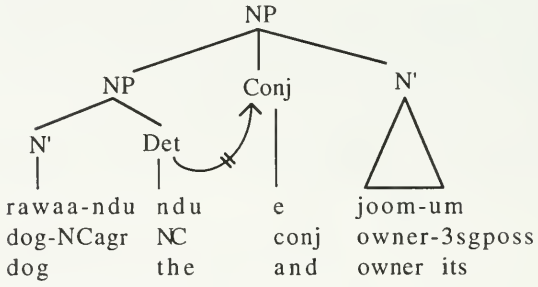
(40) More predictions of the DSA

a.



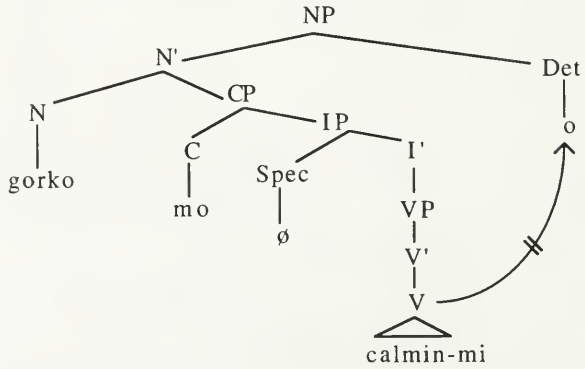
the meat, you don't like it

b.



the dog and its owner

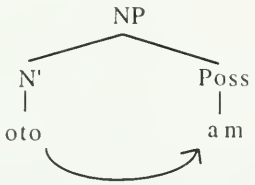
c.



gor-k-o-mo-calmin-mi-o
man-?-NCagr-Rel-greet-1sg-NC
the man that I greeted

As shown by the trees in (40a-c) above c-command does not hold in (40a) between /ngu/ 'the' and /a/ 'you-sing.' since the first element is dominated by a branching maximal projection (NP) that does not dominate the second element /a/. In this case neither version of c-command is satisfied; so the two morphemes fall outside the domain of vowel deletion. The same reasoning holds for (40b) where the determiner /ndu/, daughter of a branching maximal projection, cannot c-command the following conjunction /e/. In (40c) we see that clearly the pronoun /mi/, which is internal to the verbal complex cannot under any of the versions outlined above, c-command the determiner /o/. So, deletion is expected not to take place there either. Another correct prediction that (37) makes is (27j) where we have a noun followed by a possessive pronoun /am/. Both of these are within the same branching maximal projection; so either version of c-command holds for them, as shown in (41) below.

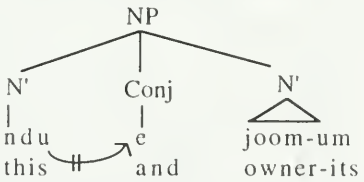
(41)



However, looking further into the data in (27) we can see that the rule in (37) gets in trouble with (27e, h, i, k). In (27e) the final vowel of the first element in an NP coordination does not delete whereas it did in (6b), or it fed glide formation, as in (8a). (42a) below is a possible representation for (27e). In this representation c-command holds between /ndu/ and /e/ assuming either version; yet the rule fails to apply. (42b-d) illustrate trees for cases (24)h, i, and k where the rule also fails to apply despite the fact that c-command holds.

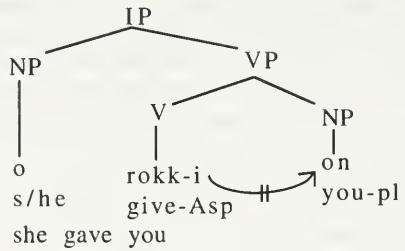
(42) Problematic cases for c-command

a.

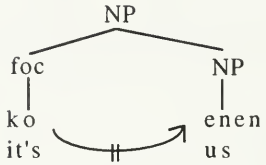


this (one) and its owner

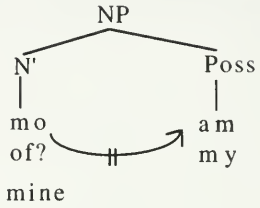
b.



c.

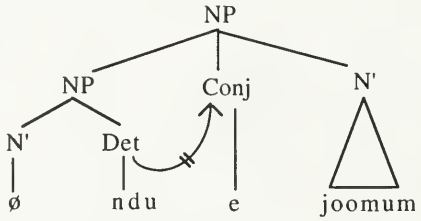


d.



In all these cases c-command holds but the vowel of the first word is not deleted and the initial vowel of the second morpheme is, consequently, realized with an initial glottal stop. However, there is a way in which (37) can account for these data without any additional stipulations to the rule. It is as follows. In (42a) /ndu/, as a determiner, is not the head of the NP. It refers to a noun belonging to the /ndu/ class, e.g. /rawaandu/ 'dog'. The head, in this case, is missing. Therefore, what we actually have in (42a) is a case of a null-headed NP of the kind introduced in (31f) above. This being the fact, the syntactic representation of (27e) is one in which the first member of the coordinate structure is a branching NP that lacks a head, as illustrated in (43) below. In this case c-command does not hold and the data are accounted for by the rule in (37), as was the case for (40b) above.

(43) Further predictions of c-command



The case illustrated in (42b) above can also be accounted for if we assume the following facts. As mentioned earlier, (cf. Prunet & Tellier (1984); Paradis (1986, 1992) for discussion) while postposed subject pronouns are considered part of the verbal complex, object pronouns (with the exception of /maa/ and /moo/) are viewed as being outside it. One of the arguments used for this analysis of Pulaar

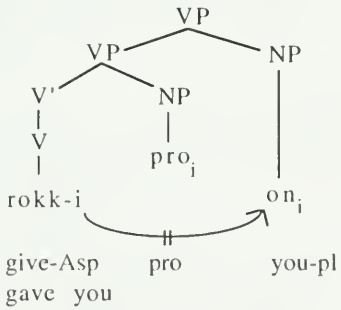
verbal complexes by the above-mentioned authors is the fact that subject pronouns cause the preceding long vowel within the verbal root to shorten whereas object pronouns do not have such effect. Furthermore, they also contend that a high vowel from a subject pronoun will cause preceding mid vowels from the preceding verb root to undergo ATR harmony whereas high vowels from the object pronoun do not cause ATR harmony in verbal roots. So, a verbal complex followed by a subject pronoun can be represented as in (44) below.

(44) Verbal complexes in Pulaar

rokk-u-mi
give-epen-1sgSubj
I gave
= ndokkumi

Since object pronouns do not form a constituent with the preceding verb they cannot be represented as in (44) above. So, my analysis is that they were moved from that position to a position outside V' (by application of the general rule Move-α; cf. Chomsky 1986), as shown in (45) below where pro is coindexed with the NP object pronoun that has been moved. In this configuration V does not c-command the following NP and deletion is not supposed to occur as the two elements /rokki/ and /on/ fall outside the domain of vowel deletion.

(45) Pulaar object pronouns: a syntactic representation

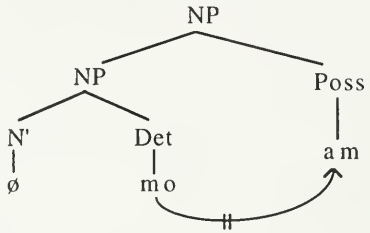


The arguments that lead to (46b) below are similar to those used to arrive at (43) above. Therefore it is not necessary to repeat them all. What is of most relevance here is the fact that in a construction like (42d) above the 'complementizer' that is used for relative clause formation is used before the possessive pronoun despite the fact that a relative meaning is not necessarily implied. That same complementizer

is also used in constructions such as 'of + place name' as in (46a) below. So, /mo/ whose meaning is something close to 'of' is analyzed as a determiner. In this case it is part of a branching NP whose head is null, as illustrated in (46b) below. In this configuration c-command does not hold and the data are predicted by rule (37) above.

(46) a. gor-k-o-mo-Dimmbee Jooro
man-ʔ-NCagr-of-Dimmbee Jooro (name of a city)
the man from Dimmbee Jooro

b.



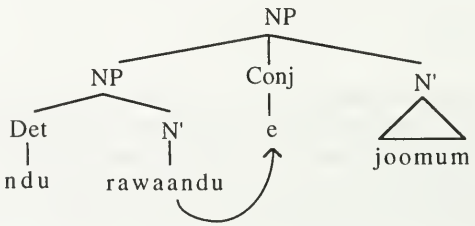
The data in (27i), represented in (42c) are a bit different in that we have a focused NP. In this case the explanation for failure of the rule to apply could be as simple as saying that focused NPs are promoted to the category of major lexical item and, therefore, behave like verbs, adjectives, or nouns whose initial vowel never participates in the process of vowel spreading.

So the rule in (37) seems to be able to account for all the data discussed so far. However, (37) along with the c-command versions assumed in (35) above incorrectly predict vowel deletion to be blocked in (27d) which is reproduced here as (47a). In (47) we have a branching NP (/ndu rawaandu/ 'this dog') the head of which is vowel-final. The rule applies to that head although it does not c-command the following conjunction, under either version in (35), as illustrated by (47b).

(47) C-command from branching structures

a. ndu-rawaa-ndu-e-joom-um
Det-dog-NCagr-Conj-owner-its
this dog and its owner
= ndu rawaandee joomum

b.



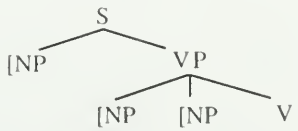
The problem is that Pulaar is full of cases like (47a) above. It is not the case that they are just a handful of exceptions that can be dealt with by adding some diacritic to the rule in (37). In fact it is not easy to imagine a proviso that can be added to (35) or (37), that is consistent with the different theories of c-command in the literature. In (47) there is both a maximal projection and a branching node that intervenes between /rawaandu/ 'dog' and /e/ 'and'. So, neither c-command nor government is supposed to hold between these two constituents. Therefore, although (37) handles all the data presented it cannot account for (27d) and the many instances of similar coordination. In such structures the first head noun has a determiner and is, consequently, daughter of a branching node. Thus, it cannot c-command the following constituent. In the next section I am going to analyze the data within the indirect access (also call the end-based theory) framework with a view to showing that it too cannot 'straightforwardly' render certain facts of VD in Pulaar. First, a look at the basic tenets of the end-based theory.

4.2 VD in Pulaar and the end-based theory (Selkirk 1984, 1986, 1987)

Within this approach to domain definition there are two important notions that play a crucial role in 'the mapping of syntactic representation into that hierarchy of prosodic domains which forms the essential constituency of phonological representation' (Selkirk 1987:152). First, the notion of Designated Category (DC). This notion has to do with the idea that for each prosodic category P_i 'there is a single designated category in the syntax with respect to which phonological representation at level P_i is defined' (Selkirk 1987:152). Selkirk argues that the basic X-bar levels in syntax, along with Government (more specifically L-governed/non-L-governed; cf. Selkirk (1982) for more discussion) determine the different designated category types. So, a designated category could be X^0, X', or X"/XP (i.e. a maximal projection). The second notion for the end-based theory is the End Parameter according to which only one end of a given designated category within the X-bar hierarchy is 'relevant in the formation of a prosodic constituent P_i: a P_i is claimed to extend from one instance of the appropriate end (R/L) of the DCi to the next (or failing that to the limit of the sentence) (Selkirk 1987:152).

Thus, in (48) below, where it is the designated category is XP and the end parameter is left, all three NPs constitute three separate domains on their own whereas V is not a domain (because not an XP) though it may be part of the domain of the NP to its left:

(48) XP/ L



If the right edge were to be the parameter then the V would not be within the domain of any of the NPs in (48). It would form a domain on its own. If we change

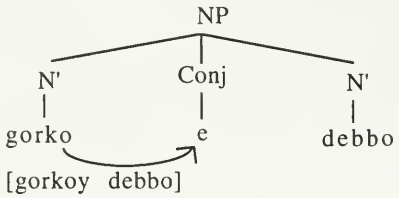
the designated category in (48) above from XP to X and the end parameter to L (left), then V would be in a domain separate from that which contains the two NPs containing it. If the end parameter is R (right) then V and both NPs preceding it would be in the same domain.

The Pulaar data that I have presented so far establish the fact that when two morphemes (the first one vowel-final and the other vowel-initial) are within the domain of VD, deletion of the final vowel of the first morpheme is followed by a rightward spreading of the initial vowel of the second morpheme. Throughout the data one can also establish that the first morpheme within the domain of VD is typically the host in the sense of its final vowel undergoing deletion and in terms of it hosting spreading from the next morpheme's vowel. Let us also assume for the sake of the argument that the prosodic category P_i here is the phonological word. The parameters for VD in Pulaar can be set as in (49a) where the designated category is XP and the end-parameter is L (left). What (49a) says is to insert a bracket to the left of each maximal projection, in which case anything to its right will be included within the same domain (P_i) up to the next maximal projection or to the end of the sentence. To test (49) against the data presented in this paper I start, first, with coordination. Given the representation of coordination that I gave in (38) where neither head is dominated by a maximal projection (49a) correctly predicts that the first head and the conjunction are going to be within the same domain, as shown in (49b) below.

(49) Parameters for VD: the left edge

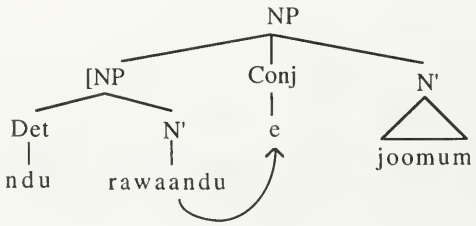
a. XP/L

b.



(49) also correctly predicts cases where the first noun in a coordinate structure is dominated by a (branching) maximal projection, as in (47b) above repeated here as (50) below.

(50)



In this and similar cases the first NP and the following conjunction form the same domain and VD applies. (49) also correctly predicts cases such as (41) where a noun forms a domain with the following possessive pronoun because they are both dominated by the same maximal projection. Furthermore, (49) also correctly predicts VD to be blocked in (40a) because the subject pronoun /a/ is dominated by a maximal projection and is, therefore, the beginning of a domain. Consequently it is within a separate domain from that containing the vowel-final /ndu/. Other data that (49) can easily accommodate are illustrated by (45) where a verb is separated from the following object pronoun by two maximal projections at the left of which a bracket is inserted to start a domain. VD is expected not to apply in this environment. In (42a) (49) also makes the correct prediction since a bracket is going to be inserted to the left of the NP dominating /enen/, putting it in a separate domain from the preceding focus marker.

As one can see (49) can account for a large body of data on VD. However, it is not general enough to handle a considerable amount of cases. In the following I discuss these. In (38b) the subject pronoun /a/ and the preceding complementizer are within the domain of VD because what native speakers say is /maa yiyaani/ < /mo-a-yiy-aan-i/ 'that you see-not-Asp'. The subject pronoun is dominated by NP and is expected to form the beginning of a domain separate from that containing the complementizer; and that is exactly the wrong prediction. Likewise, in (39a-b) the algorithm in (49) wrongly predicts that the preposition /e/ is going to form a separate domain from the preceding verb because the preposition is dominated by a maximal projection (PP) the left of which is a domain break that puts /e/ in a separate domain from the preceding verb. (49) also wrongly predicts that in (40b) /ndu/ (dominated by NP) and the following conjunction are going to be within the domain of VD. Likewise, it also predicts that in (40c) the verbal complex (dominated by VP) is going to group with the following determiner /o/. This is wrong. (49) gets in trouble further with (43) where it wrongly groups /ndu/ (dominated by the left NP) with the following conjunction within the same domain. The same problem arises with (46b). Therefore, if we maintain (49) we are going to have to explain all these exceptions. Instead of doing that let us set the parameter as in (51) where the Edge Parameter is set at R (right); everything else is going to remain the same as in (49).

(51) Parameter for VD: the right edge
XP/R

(51) inserts a bracket to the right of a maximal projection. Everything within that maximal projection falls within the domain of VD up to the next maximal projection or the end of the sentence. I assume the syntactic representations found in (38)-(50) above to be representative of all the data presented in this paper. Given that assumption (51), as will be shown shortly, makes the correct predictions for all the data except for three cases: (42c-d) and (47). I will first demonstrated the non-problematic cases for the right edge, mentioning, where appropriate, the cases that do not discriminate between (49) and (51); then I discuss the problematic cases.

In (38) above rule (51), just like (49), makes the right prediction since the same NP dominates /gorko/ and the following conjunction. (51) is able to account for the data in both (38b) and (39a-b) whereas these cases were problematic for (49). (40a) is unproblematic for both (49) and (51) because they both correctly put /ngu/ and /a/ in separate domains. However, (40b-c) are correctly predicted only by (51) which puts a bracket to the right of the leftmost NP, and VP, respectively. (41) is as unproblematic for (51) as it was for (49). As for (42a-b), recall that these were said to be the incorrect representation for these sentences. Consequently they fall outside the purview of (49) and (51). (43), (45), and (46a-b) are also correctly predicted by (51) whereas they were problematic for (49).

However, as pointed out earlier, (51) also gets in trouble. In particular, it is unable to account for (42c-d) and (47). In (42c) the parameter in (51) predicts a bracket to the left of the NP dominating /enen/, wrongly putting it and the preceding fous /ko/ in the same domain. In (42d) where both /mo/ and /am/ are dominated by the same maximal projection, (51) just like (49) wrongly puts the two words in the same domain. The other problematic case for (51) is in (47). In (47b) a bracket to the right of the NP that dominates /ndu rawaandu/ 'this dog' wrongly puts /rawaandu/ 'dog' and the following conjunction /e/ in separate domains whereas they should be within the same domain because VD applies in that context. These are the three cases that are problematic for (51) compared to many more problematic cases for (49). For this reason I am going to choose it to account for VD in Pulaar.

What I have been able to demonstrate so far is the fact that both the direct and indirect approach are unable to account for all the data presented in this paper in any unified way. For this reason, I propose a rule that has two components. The first component is going to ignore any c-command relation (the direct approach) or the Edge Parameter (the indirect approach). It targets heads of maximal projections and words that are not major lexical category items. It is going to account for (42c-d) as well as (47). The second component is going to account for all the remaining data. As I will demonstrate shortly, for the second component both the direct and the indirect approach are empirically equivalent. (52a-b) illustrate the two different components of the rule.

(52) Pulaar VD: a final formulation

a. Delete the final vowel of the final syllable of the head of a maximal projection and spread onto that syllable the initial vowel of a following word if the latter is not a major lexical category item;

b. XP/R

(52a) correctly predicts that in both (42c) and (42d) VD does not take place since neither first word in both cases is head of a maximal projection. In (47), however, (52a) is going to predict that since /rawaandu/ is a head (of the leftmost NP) and the word following it is not a major lexical category item, VD is going to apply. (52b) applies after (52a). An ideal situation would have been one in which reference is made just to heads of maximal projections and what follows them, as indicated in (52a). Most of the data, in fact, could be explained using that refer-

ence. However, headedness alone is insufficient. The data in (40c), for instance, are proof of that. In (40c) we have /calmin-mi/, head of the VP, which does not form a domain with the following determiner. (52a) cannot explain that. C-command or (52b) above can account for the data there. A further case that would be problematic for a solution based solely on the notion of headedness is in (45). In this example (52a) would predict /rokki/, head of the VP, to undergo VD. That is the wrong prediction. Again in this case, either (52b) or c-command correctly predicts VD to be blocked between the verb and the following object pronoun. (Since I discussed where c-command holds and where it does not I refer the reader to that discussion to better illustrate the fact that c-command, along with (52a) predicts all the data presented here in the same way (52a-b) do.)

Therefore, I conclude by saying that the rule that accounts for the data on Pulaar VD, necessarily requires the introduction of a statement like (52a), in addition to reference to either c-command or the Edge Parameter. I also come to the conclusion that the Pulaar data do not discriminate between the direct or the indirect approach to the syntax-phonology interface because either approach, along with the statement in (52a) correctly accounts for the data.

In this paper I have tried to do the following. After an introduction to the vowel system I presented a first set of data illustrating vowel deletion and vowel spreading in Pulaar. In that section the output of VD was shown to be: a long vowel, a short vowel, and a vowel-glide sequence. In (8f) I also showed that syllabification across morpheme boundaries is observed in Pulaar to the effect that the final consonant of a word and the following open-syllable word can form a syllable. Following this I introduced facts of the Pulaar syllable that are pertinent to the discussion and gave a phonological analysis of VD in the form of a two-step process: a (specific) rime merger rule that accounts for the vowel-glide sequences and a (general) nucleus merger rule for the cases where VD yields long or short vowels. These two rules, I suggested, were ordered following Kiparsky's Elsewhere Condition. After a series of derivations illustrating the rules at work I presented a new set of data (19a-b) that was apparently problematic for my general rule but I showed that these data are accounted for without changing or adding anything to the rule. The next discussion after that also shows how apparently problematic data (24) are easily accounted for without changing the rules. In (24) I demonstrated that failure of VD to apply in (24a-f) had to do with prosodic information, not the way the rule itself was formulated. In section 4 I showed that VD is sensitive to syntactic information. After presenting the data and the syntactic structures in which the two words that are supposed to be within the domain of VD appear, I introduced background information about the Pulaar noun, noun phrase, relative clause, and verbal complexes. In section 4.1 I presented the direct approach whereas the indirect approach was illustrated in 4.2. I demonstrated in these two sections that neither approach can handle all the Pulaar data in any unified way and that the correct generalization about the data required combining one of the approaches with the notion of headedness.

As I pointed out at the beginning of this paper there is a third approach to the syntax-phonology interfaced that I do not discuss here. It has to do with the issue

of what phonological rules can tell us about surface syntactic structure (Kenstowicz 1987). Another somewhat related issue that I am also not going to discuss has to do with the issue of whether phonological rules have access to deep (syntactic) structure or whether they can apply before or after the application of Move-α. The data in (45) above where VD applies after Move-α has moved the object NP raises such questions but I will leave them for further research.

NOTE

*This paper stems from chapter three of my doctoral dissertation (Diop 1993). I am indebted to Charles C. Kisseberth, James Yoon, Laura J. Downing, Alessandro Zucchi, and Elabbas Benmamoun for invaluable comments on earlier versions. All errors contained in this paper are entirely mine.

REFERENCES

CHO, Yu Young-Mee. 1990. Syntax and phrasing in Korean. The syntax-phonology connection, ed. by S. Inkelas & D. Zec, 47-62. Chicago: University of Chicago Press.

CHOMSKY, Noam. 1986. Barriers. Cambridge, MA: MIT Press.

——, & Morris Halle. 1968. The sound pattern of English. New York: Harper & Row.

CLEMENTS, G. N. 1978. Tone and syntax in Ewe. Elements of tone, stress, and intonation, ed. by D. J. Napoli, 21-99. Washington, D.C.: Georgetown University Press.

DIOP, Abdul Aziz. 1993. Aspects of Fula non-linear phonology: Towards more structure in moraic phonology. Urbana: University of Illinois Ph.D. dissertation in Linguistics.

GOLDSMITH, John. 1990. Autosegmental and metrical phonology. Oxford: Basil Blackwell.

HARRIS, J. 1983. Syllable structure and stress in Spanish. Cambridge, MA: MIT Press.

HAYES, Bruce. 1986. Inalterability in CV phonology. Language 62.321-351.

——. 1989. Compensatory lengthening in moraic phonology. Linguistic Inquiry 20:2.253-306.

HYMAN, Larry. 1985. A theory of phonological weight. Dordrecht: Foris.

——. 1990. Boundary tonology and the prosodic hierarchy. The syntax-phonology connection, ed. by S. Inkelas & D. Zec, 109-126. Chicago: Chicago University Press.

ITÔ, Junko. 1986. Syllable theory in prosodic phonology. Amherst: University of Massachusetts Ph.D. dissertation in Linguistics.

JACKENDOFF, Ray. 1977. X-bar syntax: A study of phrase structure. Cambridge, MA: MIT Press.

KAISSE, Ellen M. 1985. Connected speech: The interaction of syntax and phonology. Orlando, FL: Academic Press.

——. 1987. Rythm and the cycle. Chicago Linguistics Society 23:2.199-209.

——. 1990. Toward a typology of postlexical rules. The syntax-phonology connection, ed. by S. Inkelas & D. Zec, 127-144. Chicago: Chicago University Press.

KENSTOWICZ, Michael. 1987. The phonology and syntax of wh-expressions in Tangale. Phonology Yearbook 4.229-241.

——, & Charles Kisseberth. 1990. Chizigula tonology: the word and beyond. The syntax-phonology connection, ed. by S. Inkelas & D. Zec, 163-194. Chicago: Chicago University Press.

KIPARSKY, Paul. 1973. Elsewhere in phonology. A festschrift for Morris Halle, ed. by S. Anderson & P. Kiparsky, 93-106. New York: Holt, Rinehart and Winston.

——. 1982 Lexical phonology and morphology. Linguistics in the morning calm, ed. by I.-S. Yang, 3-91. Seoul: Hanshin.

MCCARTHY, John. 1986. OCP effects: gemination and antigemination. Linguistic Inquiry 17.207-263.

ODDEN, David. 1986. On the role of the Obligatory Contour Principle in phonological theory. Language 62.353-383.

——. 1987. Kimatuumbi phrasal phonology. Phonology Yearbook 4.13-36.

——. 1990. Syntax, lexical rules and postlexical rules in Kimatuumbi. The syntax-phonology connection, ed. by S. Inkelas & D. Zec, 259-278. Chicago: Chicago University Press.

PARADIS, Carole. 1986. Phonologie et morphologie lexicales: les classes nominales en peul (fula). Universite de Montreal Ph.D. dissertation in Linguistics.

——. 1992. Lexical phonology and morphology: the nominal classes in Fula. New York: Garland Publishing.

PRUNET, J. F., & C. Tellier. 1984. Interaction des niveaux en phonologie: l'abregement vocalique en Pulaar. McGill Working Papers in Linguistics 2:1.65-90.

SCHANE, Sanford A. 1987. The resolution of hiatus. Parassession on autosegmental and metrical phonology. CLS 23:2.279-290.

SELKIRK, Elizabeth. 1984. Phonology and syntax: The relation between sound and structure. Cambridge, MA: MIT Press.

——. 1986. On derived domains in sentence phonology. Phonology Yearbook 3.371-405.

——. 1987. Government and tonal phrasing in Papago. Phonology Yearbook 4.151-182

——, & Tong Shen. 1990. Prosodic domains in Shanghai Chinese. The syntax-phonology connection, ed. by S. Inkelas & D. Zec, 313-378. Chicago: Chicago University Press.

SELLS, Peter. 1985. Lectures in contemporary syntactic theories. Stanford: Center for the Study of Language and Information.

STERIADE, D. 1982. Greek prosodies and the nature of syllabification. MIT Ph.D. dissertation in Linguistics.

SYLLA, Yero. 1982. Grammaire moderne du Pulaar. Dakar-Abidjan-Lome: Nouvelles Editions Africaines.

Studies in the Linguistic Sciences
Volume 23, Number 2, Fall 1993 [publ. October 1996]

SUBVERSION OR CONVERGENCE? THE ISSUE OF PRE-VEDIC RETROFLEXION REEXAMINED*

Hans Henrich Hock

एतद्धस्तिदर्शन इव जात्यन्धाः (Śaṅkarācārya's bhāsya on Chāndogya-Upaniṣad 5.18.1)

As is well known, members of at least six distinct language families in South Asia have come to converge to a remarkable degree in their overall structure through millennia of bi- and multilingual contact. Most of the convergence has been in the syntax, but one phonological phenomenon, a contrast between dental and retroflex consonants (as in Skt. *dī-* 'shine' : *ḍī-* 'fly'), has the distinction of having been noticed earliest (Pott 1833, 1836). It is this phenomenon which I address in this paper.

Some scholars (most recently Thomason & Kaufman 1988, Kuiper 1991; see also and especially Emeneau 1980) argue that the source of retroflexion is Dravidian, for the dental : retroflex contrast can be reconstructed for Proto-Dravidian, while the ancestors of the other languages are said to have lacked it. Since the contrast is found in the earliest attested stage of Indo-Aryan, Vedic Sanskrit, convergence between Indo-Aryan and Dravidian must therefore have begun in the second millennium B.C., in terms of a SUBVERSION (my term) of Indo-Aryan by Dravidian.

In earlier publications (e.g. Hock 1975, 1984) I claimed that the arguments for early convergence are not cogent, since the Sanskrit dental : retroflex contrast can be explained by internal Indo-Aryan developments. Moreover, the contrast appears to be an innovation not only in Indo-Aryan, but also in Dravidian. This raises the possibility that the feature is a JOINT innovation of Dravidian and Indo-Aryan, reflecting direct or indirect bilingual contact.

At the same time, early Dravidian has a TRIPLE contrast, dental : alveolar : retroflex (or post-dental). This difference may be taken to cast doubt on the convergence hypothesis.

In this paper I present a somewhat speculative hypothesis that a triple dental : alveolar : retroflex contrast must be postulated for early stages of both Indo-Aryan (and Iranian) and Dravidian and further, that this contrast resulted from joint, convergent innovations. I support the hypothesis with comparative Indo-Iranian evidence and the dialectological evidence of early Middle Indo-Aryan (especially the Aśokan inscriptions). The latter evidence is especially interesting, since the development of the hypothetical Indo-Aryan alveolars to dentals in the more western regions and to retroflexes in the more eastern regions of Indo-Aryan is closely mirrored by corresponding

developments of the well-established alveolars to dentals in more western Dravidian and retroflexes in more eastern Dravidian. Alternative explanations of the observed data either are unnecessarily complex or are lacking in explanation.

The finding that the developments are the result of convergence, not of subversion, is significant, for it suggests that the social relationship between Indo-Aryan and Dravidian speakers in early India was not substantially different from what it is today—a relationship of (near-)equality, rather that the traditional picture of marauding Indo-Aryan invaders suppressing an indigenous Dravidian population and forcing it to learn their language.

1. Introduction

South Asia is a paradigm example of a multicultural, multiethnic, multilingual area. Members of at least six distinct language families coexist: Indo-Aryan, East Iranian, Munda, Austro-Asiatic, Dravidian, Tibeto-Burman, plus an "unaffiliated" language in the extreme north, Burushaski.[1] (For a simplified view of where these languages are spoken today see Map I, next page.) While only Indo-Aryan and Iranian are closely related, and Munda is remotely related to Austro-Asiatic, millennia of bi- and multilingual contact have led to a remarkable degree of structural convergence between these different language families and their members. As a consequence, South Asia has also come to be known as a paradigm case of a CONVERGENCE AREA. (For a good synchronic discussion see Masica 1976.)

Four features are commonly listed as characteristic of this area:

I. An unmarked major constituent order SOV, i.e. subject (S) before object (O) before verb (V), as in example (1);
II. A tendency to use non-finite absolutives, where modern European languages might use dependent clauses with finite verbs (2);
III. The marking of cited discourse by postposed quotative markers and a general absence of indirect discourse (3);
IV. "Retroflexion", i.e. a phonological contrast between dental and retroflex consonants (4).

(1) Hindi maiṁ (S) kitāb (O) paṛh rahā hūṁ (V) 'I am reading a book.'

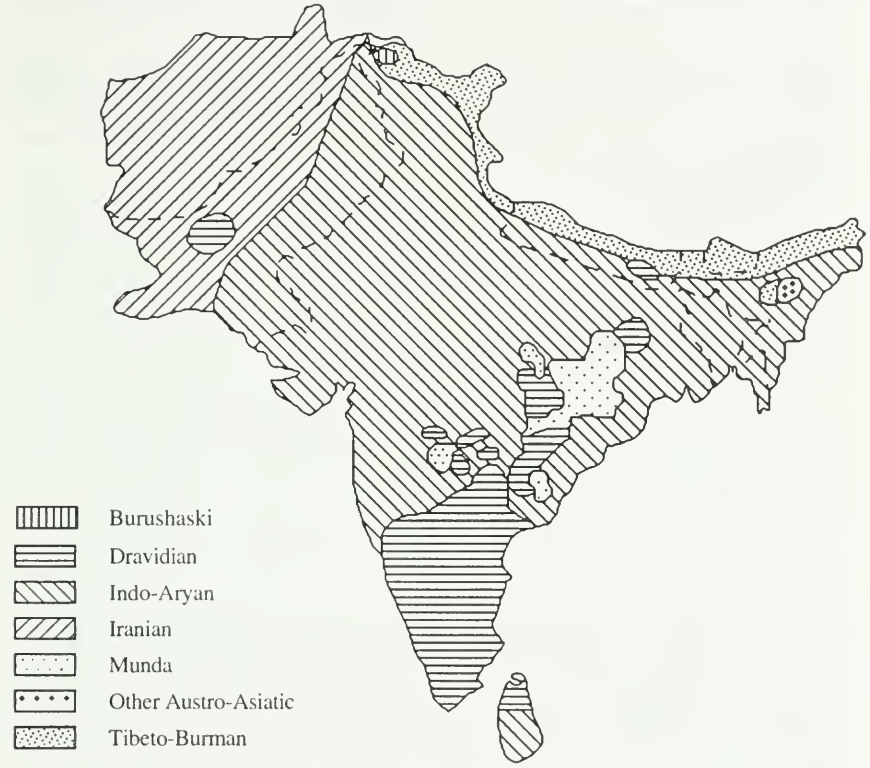
(2) Sanskrit tatra **gatvā** (abs.) na mucyase 'When you have gone there, you do not get free.'

(3) Sanskrit nakir vaktā 'na dād' **iti** (quot.) 'Nobody will say, "He shall not give."'

(4) Sanskrit pāta- 'flight' : pāṭa- 'portion'

As is common in convergence areas,[2] these features do not cover the entire region: Kashmiri places finite verbs in second position in main clauses and certain dependent clauses. SVO features are found in many Munda/Austro-Asiatic languages. Hindi-Urdu, Kashmiri, and many other Indo-Aryan languages, but

also Dravidian Brahui mark direct discourse by preposed *ki*, not by postposed quotative markers. Assamese and much of Tibeto-Burman lack retroflexion.



Map I: Distribution of modern South Asian languages (simplified)

The presence of all of the features in Sanskrit, the oldest attested stage of Indo-Aryan, suggests that the absence of some of these in some of modern Indo-Aryan results from secondary developments. On the origin of the Kashmiri verb-second position see Hock 1982a. The use of *ki* as a marker of direct discourse is no doubt due to Persian influence (see Hock 1982b, which needs to be updated; see also Marlow Forthcoming). The absence of the retroflexion in Assamese probably reflects contact with Tibeto-Burman.[3]

The question of when the linguistic convergence of the South Asian languages began and which group is responsible for it is highly interesting for anyone concerned with the early linguistic and ethnic history of South Asia. Attempts to answer the question, however, have led to some controversy.

Many scholars (most recently Thomason & Kaufman 1988, Kuiper 1991; see also and especially Emeneau 1980) argue that the source is Dravidian, since all of the four features can be reconstructed for Proto-Dravidian, while the ancestors of

the other languages are said to have lacked them. Since the features are found in the earliest attested stage of Indo-Aryan, Vedic Sanskrit, Indo-Aryan must according to this view have acquired them prior to that stage, in the second millenium B.C.

The sociolinguistic setting for subversion is usually considered one of inequality: The Indo-Aryan conquest forced the indigenous Dravidians to learn Sanskrit, the language of the Indo-Aryans; and as English has been influenced by the modern South Asian languages, the structure of Sanskrit was altered by transfer of Dravidian features.

The term used in traditional historical linguistics for such a development is "substratum influence". Using the terminology of Thomason & Kaufman 1988, the process can be characterized as "shift" of speakers from one language (usually of lower power or prestige) to another one (of higher power or prestige). For brevity's sake let me use the more compact term SUBVERSION to refer to "substratum influence" or the effects of "language shift".

Following the lead of others, I have claimed (e.g. Hock 1975 and 1984) that the arguments for prehistoric subversion are not cogent: The syntactic features (I - III) are either inherited from Proto-Indo-European or Proto-Indo-Iranian, or are typologically natural in early Indo-European; and the feature of retroflexion can be explained by internal Indo-Aryan developments.

Several recent reinvestigations of early Dravidian and Indo-Aryan/Indo-European syntax support the claim that the syntactic features, in their broad outlines (and including at least one other feature, the use of relative-correlative structures), were shared by the prehistoric ancestors of Dravidian and Indo-Aryan, going back to periods much earlier than the Indo-Aryan migration to South Asia. (See Steever 1988, Hock 1988b, 1992a, as well as Hock 1996).[4] For these reasons, and to keep the present paper within manageable limits, I concentrate on retroflexion, bringing in other evidence only where relevant to the argument.

In contrast to most earlier subversionist claims, but also breaking with my own earlier counterclaims, I present a speculative argument that retroflexion can be explained as resulting from CONVERGENCE, a process different from subversion, both in its effects and in its social setting. While subversion consists of the unidirectional transfer of features from one language to another, under conditions of strong inequality and sudden shift, convergence is a more complex, mutual or bidirectional development through which languages in long-standing bilingual contact come to be more similar in their overall structure. The required extended bilingualism is best maintained in a situation of approximate social equality; but it can also arise under other conditions, such as a "social imperative" of maintaining ethnic, religious, etc. identity by preserving linguistic distinctiveness.[5]

I argue that prehistoric convergence took place under social conditions that fostered extended bilingualism, similar to what we find in modern South Asia; that it involved Dravidian, Indo-Aryan and East Iranian, and possibly other languages as well; and that it led to a triple contrast (at least in Indo-Aryan and Dravidian) of dental : alveolar : retroflex, not just the simple dental : retroflex contrast ordinarily postulated for Indo-Aryan.

2. A survey of earlier views on Indo-Aryan retroflexion

Early scholars such as Pott 1833, 1836 and Caldwell 1855 could simply assert that Indo-Aryan (Sanskrit) retroflexion results from Dravidian subversion. But as time progressed it became necessary to go beyond such sweeping statements and to state more precisely HOW subversion exerted itself.

What was especially troubling is that in its general outlines early Indo-Aryan retroflexion could be explained by purely internal developments, with parallels in other languages (see e.g. Konow 1906, Bloch 1925). Compare the schematization in (5), and see also §5.1 below. Similar developments can be found in other Indo-European languages, most notably in Swedish and Norwegian dialects. Ever since Bühler 1864, anti-subversionists have taken these facts as evidence that we do not need to invoke Dravidian subversion to explain Indo-Aryan retroflexion. Something like a compromise position was offered by Konow 1906 and Bloch 1929 who claimed that Dravidian influence may have accelerated or aided in the propagation of these developments.

(5)

	I	II	III	IV	V
a.	(*liǵh-to- >)	*liẑdha- >	*liẓdha- >	*liẓdha- >	līḍha- 'licked'
cf.	(*wiǩ-to- >)	*wišta- >	*viṣṭa- >	viṣṭa- =	viṣṭa- 'entered'
b.	(*wiǩ-s >)	*wiš̌ >	*wiṣṣ >	*wiṭs >	viṭ 'people, clan' (N sg.)
cf.	(*wiǩ-su >)	*wiš̌šu >	*wiṣṣu >	*wiṭsu >	vikṣu (id.) (L pl.)
					(→ post-RV viṭ-su)

Emeneau 1956 and Kuiper 1967a introduced a much stronger and more specific claim: The presence of retroflexion in Dravidian led to the "redistribution" of pre-Indo-Aryan allophones as retroflex phonemes. Kuiper identified these allophones as Indo-Iranian *š̌ and *ž̌, elements generally recognized as the "triggers" for Indo-Aryan retroflexion, as in the above formulation.[6] This is now the standard subversionist position and has been accepted in Thomason & Kaufman 1988, the major general monograph on linguistic contact.

Subversionists moreover believe that it is highly unlikely that Indo-Aryan retroflexion arose independently from that of Dravidian, on the same South Asian subcontinent. And they consider irrelevant the fact that other Indo-European languages have developed retroflexion, since Indo-Aryan is the only EARLY branch of Indo-European with this feature (see e.g. Tikkanen 1987: 284).

Antisubversionists are not convinced of the logic of this argument: While it is true that Indo-Aryan retroflexion developed much earlier than retroflexion in other Indo-European languages, this does not mean that it must result from subversion. Different languages may exhibit similar phonological changes at different rates and at different stages. (Gothic, for instance, virtually leveled out the effects of Verner's Law many centuries before the other Germanic languages; but this does not mean that Gothic leveling resulted from subversion.) The fact that other Indo-European languages were able to acquire the feature demonstrates that retroflexion is not such an unusual phonological phenomenon that it must per-

force be attributed to outside subversion—pace Tikkanen 1987:284 or the extreme view of Bhat 1973.

At least some antisubversionists would however admit that the issue of whether it is LIKELY that Indo-Aryan retroflexion arose independently from Dravidian is a more serious one. Unfortunately, it is impossible to determine the likelihood of two similar phenomena arising independently in languages that come to share the same geographic area. As observed by Lyle Campbell (p.c. 1993), such an event is not impossible, as shown by the case of Brazil: Portuguese has come in contact with indigenous languages which, like Portuguese, have a contrast between oral and nasal vowels; but we know that the contrast existed before contact, in both groups of languages. Now, the Brazilian case merely establishes the possibility of chance similarity; it tells us nothing about its statistical likelihood. Nevertheless it further supports anti-subversionist reservations about the need to attribute Indo-Aryan retroflexion to Dravidian subversion.

More concrete arguments against Dravidian subversion are based on structural evidence.[7] Following Bloch 1925, Hock 1975 and 1984 observes a number of differences between early Dravidian and Indo-Aryan: Dravidian has a triple contrast (dental : alveolar : retroflex), while Indo-Aryan is considered to have a simple contrast between dental and retroflex (plus post-dental, alveolar *r*). Dravidian permits final retroflex and alveolar sonorants, early Indo-Aryan does not, except for the onomatopoetic nonce-form *bhāṇ* and coined terms of indigenous grammar. (Sanskrit word-final *r* is realized as *ḥ* utterance-finally.) From the earliest times, Indo-Aryan has at least one initial retroflex consonant (in Skt. *ṣaṭ* '6' and derivatives); Dravidian initial retroflex consonants are a late innovation. Indo-Aryan has retroflex sibilants which are absent in Dravidian, while the latter has a retroflex approximant *ɾ̤* which is absent in Indo-Aryan.[8] These extensive differences, summarized in Tables I and II, are considered difficult to explain if Indo-Aryan retroflexion resulted from Dravidian subversion.

The Brazilian parallel is interesting in this regard: The phonological effects of the oral : nasal vowel contrast, which we know to be of independent origin, differ considerably: In the indigenous languages, nasal consonants tend to become pre- or post-"oralized" next to oral vowels, a phenomenon without counterpart in Portuguese.

	SANSKRIT			DRAVIDIAN		
	DENT.	ALV.	RETR.	DENT.	ALV.	RETR.
STOP	t		ṭ	t	t̲	ṭ
	th		ṭh			
	d		ḍ			
	dh		ḍh			
SIB.	s		ṣ			
NAS.	n		ṇ	n	n̲	ṇ
LIQU.	l	r̲			l̲	l̤
					r̲	ɾ̤

Table I: Differences between the early Sanskrit and Dravidian systems

	SANSKRIT	DRAVIDIAN
Final retr./alv. sonor.	-	+
Initial retr./alv.	+	-
Idiosyncratic *ṣ*	+	-
Idiosyncratic *ṟ*	-	+

Table II: Other differences between early Sanskrit and Dravidian

Note further that many of the early phonological differences between Indo-Aryan and Dravidian disappear toward the modern period (except in the extreme south and northwest), as shown by Ramanujan & Masica's 1969 areal study of modern South Asian phonology. As argued in Hock 1984, in contrast to the prehistoric situation, this development does provide robust evidence for structural interaction. But it took place at a considerably later time and it involved convergence, not subversion.

In Hock 1975 I suggested that the dental : retroflex contrast may be an innovation, not only in Indo-Aryan, but also in Dravidian. My claim was based on speculative, and in one case clearly premature, attempts to genetically link Dravidian with outside languages which do not have the contrast (Uralic and Elamite). In a publication not accessible to me at the time, Zvelebil 1970 proposed a 'highly speculative' hypothesis that Dravidian consonant clusters and geminates result from large-scale assimilatory processes, some of which turned sequences of retroflex—or alveolar—sonorant plus dental stop into retroflex or alveolar stops.

Drawing on Zvelebil 1970 and on my 1975 suggestion, Tikkanen 1987 claims that retroflexion is innovated both in Dravidian and in Indo-Aryan. He attributes the impetus for the innovations to two separate substrata (295) and claims that convergent developments between Indo-Aryan and Dravidian took place later. On the Indo-Aryan side, he believes that the source for subversion was an unknown northwestern substratum, which in his view is also responsible for the large amount of early Indo-Aryan lexical items that can be traced neither to Proto-Indo-European nor to any of the known non-Indo-Aryan languages of South Asia. As for Dravidian, he entertains the idea that subversion is attributable to 'some lost sub- or adstratum in the pre-Indo-Aryan period' (323).

Given what we know—or do not know—about the distribution of languages in prehistoric South Asia, Tikkanen's proposal cannot be rejected out of hand. In fact, the modern presence of the language isolate Burushaski in the northwest may be taken as evidence for a prehistoric presence of a non-Indo-Aryan/non-Dravidian language in the area. (But see below on the difficulties in trying to draw prehistoric inferences from the modern situation.) Unfortunately, the hypothesis of an unknown substratum (or of several such substrata) is methodologically dubious, since by definition it is not open to verification or falsification.

Tikkanen is now doing intensive research on the northwestern languages of South Asia.[9] It is to be hoped that this research will eventually make it possible to identify a likely substratum; but the enormous time difference between the Indo-

Aryan arrival in South Asia and the first attestation of the northwestern languages places formidable obstacles in the way.

Moreover, the question must remain as to what the relation was, if any, between the two separate substrata that gave rise to Indo-Aryan and Dravidian retroflexion. Is it likely that the two substrata had developed retroflexion independently? Methodologically, invoking two separate substrata is problematic in that it merely projects the issue of Dravidian/Indo-Aryan prehistoric relationship to an even more remote—and uncertain—period in prehistory.

An alternative, and at this stage of our knowledge more feasible, hypothesis is that Indo-Aryan (as well as East Iranian) and Dravidian retroflexion and alveolarization are not just parallel innovations due to subversion by different unknown substrata, but that they result from CONVERGENT changes. It is this hypothesis which I want to support in the present paper, leaving open the question whether convergence took place under direct contact or whether it may have been mediated by other, intervening languages.

Before doing so let me briefly discuss some of the subsidiary arguments that have been raised in support of the view that Indo-Aryan retroflexion resulted from prehistoric Dravidian subversion. An examination of these arguments demonstrates the great difficulties facing anyone trying to make inferences about the prehistoric linguistic scene in the northwest (or any other part of South Asia) and the fact that any hypothesis about prehistoric contacts in this area—and their linguistic consequences—must by definition be speculative. At the same time, reexamination of one of the arguments establishes a possible building block for a new hypothesis.

3. Subsidiary arguments for prehistoric subversion

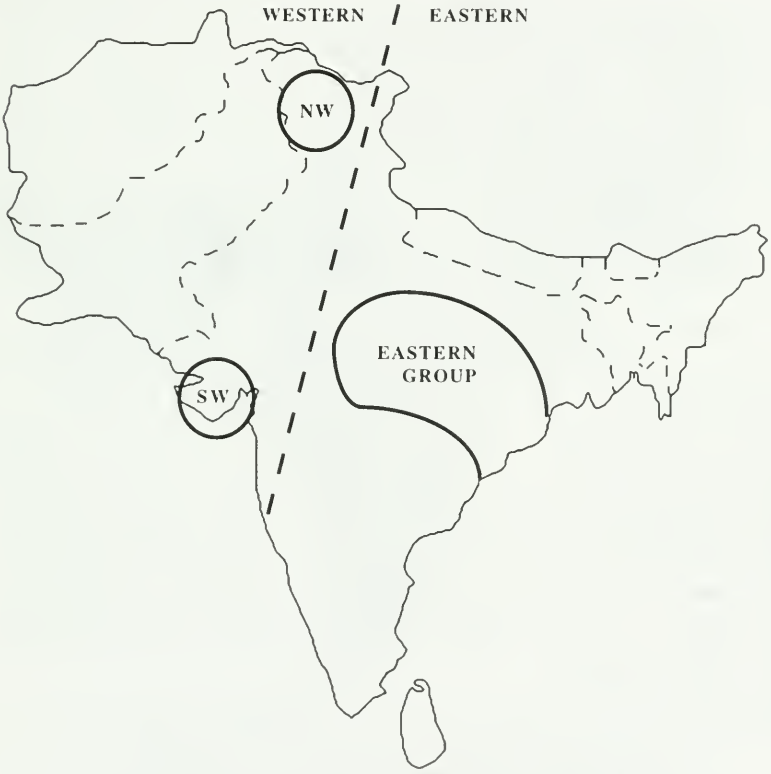
The greatest difficulty in dealing with the linguistic prehistory and early history of South Asia is the fact that we have more or less contemporary evidence from only one language family, Indo-Aryan; and even for this family the evidence is limited, because the early texts are composed in a language, (Vedic) Sanskrit, which is quite conservative and puristic. For other languages, we have to depend on evidence from much later periods. This is especially true for the non-Indo-Aryan languages spoken in present-day northwestern South Asia which are not attested before the nineteenth century.

The problems caused by this situation can be illustrated by examination of two arguments frequently raised in favor of prehistoric Dravidian subversion of Indo-Aryan.

One argument is based on the presence of a Dravidian language, Brahui, in Baluchistan. The geographical isolation of Brahui is taken to establish that it is a relic language, especially since migration is believed to normally take place only from north(west) to south. These facts are considered to legitimize the assumption of a Dravidian presence in the prehistoric northwest of first Indo-Aryan settlement. Further support is found in the fact that two other Dravidian languages, Kurukh and Malto, which with Brahui form the North Dravidian subfamily, are

spoken fairly to the north (in eastern Central India), suggesting that Brahui was part of a Dravidian subfamily which extended over a vast portion of northern South Asia.

A second argument, cited by Thomason & Kaufman 1988, rests on Southworth's 1974 attempt to establish a major east-west division of Indo-Aryan languages for the time of the Aśokan inscriptions (see Map II) and to link this division to more recent evidence that in his view suggests greater Dravidian influence in the west, i.e. in a region closer to the area of first Indo-Aryan settlement.

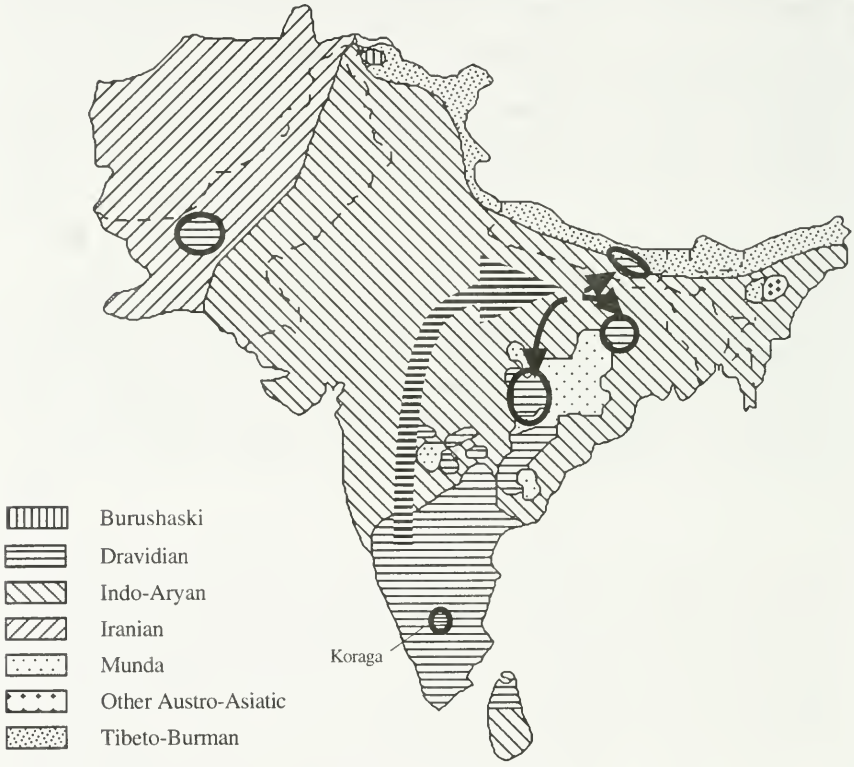


Map II: Dialect divisions of the Aśokan inscriptions according to Southworth 1974 (with reference to Bloch 1950)

Now, some 3,000 years separate the arrival of the Indo-Aryans from the time that North Dravidian languages begin to be attested. This fact in itself should give us pause. But there are more specific reasons for caution.

According to their own traditions, the Kurukh (and Malto) people migrated to their present locations from Karnataka, via the Narmada valley (Grierson 1903-1928, v. 4; Hahn 1911); see Map III. Several linguistic facts support this tradition: Bloch 1946 points out that the place names in present-day Kurukh and

Malto territory are Munda, not Dravidian, in origin. Kuiper 1966 demonstrates linguistic influence of Kurukh on Nahali and Kurku, a fact which supports the Kurukh tradition of an earlier settlement in the Narmada valley. Bhat 1971 produces linguistic evidence that Koraga in South Karnataka (see Map III) is more closely related to the North Dravidian languages than to the rest of Dravidian. In short, we have cumulative evidence that connects North Dravidian Kurukh and Malto to the south, not to the extreme northwest of Brahui.



Map III: Northern Dravidian languages and migrations

In fact, Bloch (1911, see also 1925, 1929) has suggested that Brahui, too, may have a southern origin, since according to their own traditions, the Brahuis have migrated to the area in which they live now.[10]

Such a northward migration would in fact not be unusual. As is well known, several Indo-Aryan groups likewise have followed this route, or migrated even farther. These include Gandhari or Niya Prakrit in early medieval Khotan and farther east; modern Dumaki in northwestern South Asia; the Parya who came to modern Uzbekistan via Afghanistan (Comrie 1981); and the 'Gypsies' or Ḍom who, via Central Asia, have spread all over Eurasia.

The present-day linguistic distribution, some 3000 years "after the fact" thus cannot be taken as cogent evidence for a prehistoric Dravidian presence in the northwest.

In all fairness, however, it must be admitted that the possible southern origin of modern North Dravidian does not preclude a Dravidian presence in the prehistoric northwest. We know that the just mentioned transplanted Indo-Aryan languages "remigrated" northward from areas well to the south. Thus, Comrie 1981 with references shows that the language of the Parya is closely affiliated to Hindi/Panjabi, and Kuiper 1966 adduces evidence for Kurukh influence on Ḍom in the Narmada valley. At the same time, we also know that the Indo-Aryan languages originally moved into South Asia from the northwest. Given the Indo-Aryan precedents for southward migration and subsequent remigration to the north, it is possible that there were Dravidians in the northwest when the Indo-Aryans came to South Asia, that these Dravidians moved southward under Indo-Aryan pressure (or that their languages died out), and that the present-day location of Brahui results from remigration to the north. But it is just as possible that the Dravidians, if they originally came from the north, had already departed from the northwest to the south by the time of Indo-Aryan arrival, and that only later did Brahui and the other Dravidian languages move north again.

The problem is, we simply do not have any reliable independent evidence that would permit a choice between these different possibilities.

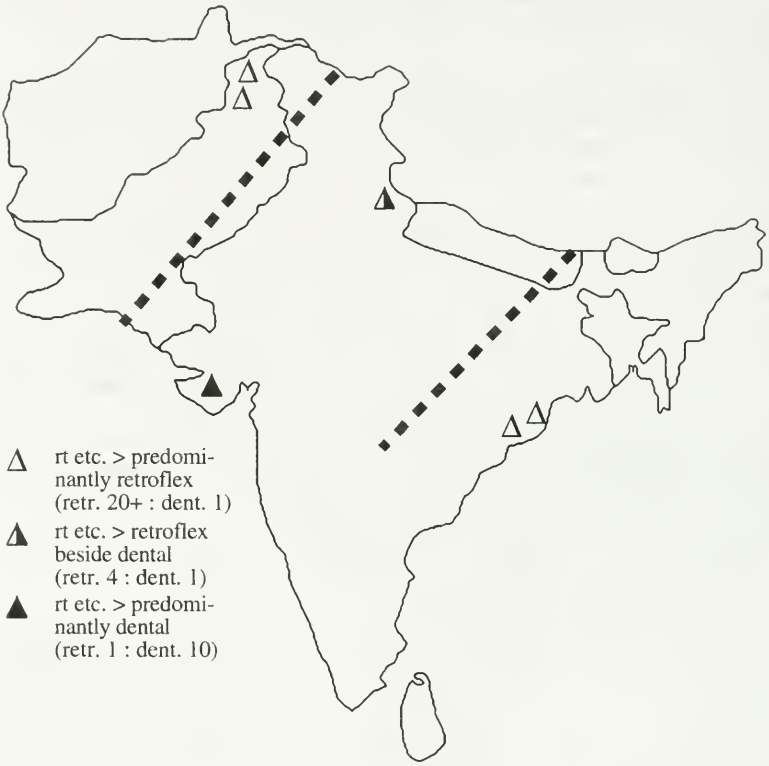
From the chronological perspective, Southworth's dialectological division of the Aśokan inscriptions rests on firmer grounds, since only a little more than a thousand years separate Aśoka from the time of Indo-Aryan arrival in South Asia. But his dual division is not supported by Bloch 1950, to whom he refers, who instead suggests a triple division: Center and East vs. [South]West vs. Northwest. A recent reexamination of the treatment of (syllabic and nonsyllabic) *r* + dental stop in the Aśokan inscriptions suggests a different division (superseding Hock 1991). If we exclude developments limited to specific lexical items which may be suspected of being borrowings, we can distinguish four different areas (see Map IV[11]):

a. A northwestern area with almost exclusively retroflex outcomes (beside cluster representations: Shahbazgarhi mainly *tr* etc., Mansehra *tr* etc.);
b. Southwestern Girnar with predominant dental;
c. North-central Kalsi with a retroflex : dental ratio of about 4 : 1;
d. An eastern area (Dhauli and Jaugada) with almost exclusive retroflex.

However, given the proximity of Kalsi to the east, its relatively high retroflex ratio may be attributed to eastern influence.[12] Under this assumption, it is possible to resolve Aśokan dialectology into three areas:

a. The northwest (predominant retroflex beside cluster representations);
b. A central area that originally includes Girnar and Kalsi (predominant dental);
c. The extreme east (predominant retroflex).

See Map IV (next page).



Map IV: Development of *r* + dental stop in the Aśokan inscriptions

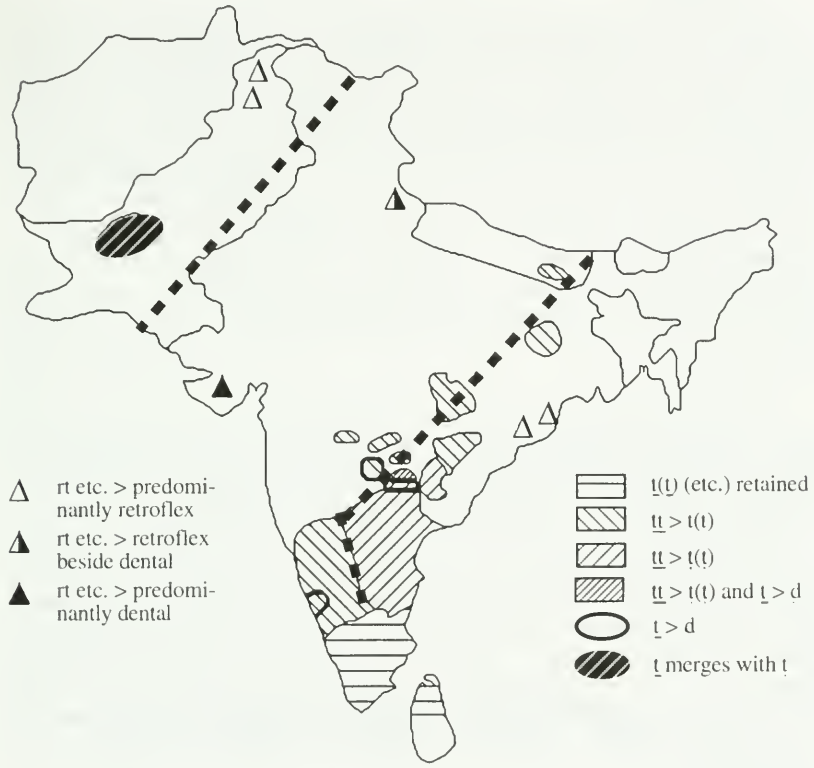
Significantly, this distribution agrees well with the Modern Indo-Aryan outcomes of *r* + dental stop discussed by Turner 1926 with 1921, 1924, a fact which suggests that the Aśokan inscriptions, at least on this count, offer a reliable window on developments in archaic Middle Indo-Aryan.[13]

By contrast, the evidence of the Aśokan inscriptions does not support Southworth's east-west division of Indo-Aryan and the concomitant claim that western Indo-Aryan exhibits stronger influence from Dravidian—whatever may be the merits of his findings for modern South Asia.[14]

Interestingly however—and surprisingly—the triple north-south division of Indo-Aryan as regards dental vs. retroflex outcomes of *r* + dental stop lines up amazingly well with a Dravidian areal division between languages in which alveolar stops have turned into dental vs. retroflex[15] (making allowances for some distributional irregularity in the transition area between dental and retroflex outcomes). See Map V (next page).[16]

To lay the foundation for an account that explains this parallelism as the result of convergent phonological developments, it is useful to briefly reexamine the

available evidence regarding the prehistoric and early historic social relationship between Indo-Aryans and non-Indo-Aryans.



Map V: Development of *r* + dental stop in the Aśokan inscriptions and Modern Indo-Aryan (according to Turner) compared with the Dravidian development of alveolar stops (mainly geminates).

4. The social relationship

Explicit, or at least implicit, in the subversionist view of early Indo-Aryan/Dravidian contact is the assumption of unilateral influence of Dravidian on Indo-Aryan. (No mention is ever made of prehistoric Indo-Aryan influence on Dravidian.) Such a unilateral development requires the assumption that the prehistoric social relationship between Indo-Aryans and Dravidians was one of considerable inequality. Dravidian speakers therefore would have had to speak Indo-Aryan and, just as happened to English in modern South Asia, in shifting to Indo-Aryan they transferred structural features of their own language(s). This interpretation of Indo-Aryan/Dravidian interaction is aptly summarized by Emeneau 1956: note 4:

> ... it was to their [the Indo-Aryans'] advantage, political, economic, religious, to have subjects and proselytes. Absorption, not displacement

is the chief mechanism in radical language changes of the kind we are considering.

This view is often considered supported by the belief that the Rig-Vedic Indo-Aryans made a strong ethnic distinction between themselves and the indigenous population, the *dāsas* or *dasyus*, frequently depicting the latter as 'infidels' (*adeva*), and characterizing them as 'black-skinned' in contrast to their own lighter hue. While not all subversionists accept this view (Emeneau, for instance, does not), it pervades much of the literature on the linguistic and general prehistory and early history of South Asia, as can be gauged from the following incomplete list of references: Zimmer 1879 (apparently the first propagator of the view); Macdonnell & Keith 1912: s.vv. dāsa and várṇa; Chatterji 1960:7 and 32; Elizarenkova 1995:36; Gonda 1975:129; Hale 1986:147 (see also 154); Kuiper 1991:17 (vs. ibid. 3-4); Kulke & Rothermund 1986:35; Mansion 1931:6; Rau 1957:16; Parpola 1988:104-106, 120-121, 125; see also Deshpande 1979:260, 1993:216-127.

Examination of the textual evidence of the Rig-Veda and general considerations regarding the early interaction of different ethnic groups suggest that, like several other aspects of early South Asian society,[17] this picture of radical inequality needs to be redrawn. (For a fuller discussion see Hock 1996.)

The Rig-Vedic passages in which adjectives meaning 'black' or 'dark' are used in reference to human enemies are of two kinds. One refers to the forts of the enemies, especially their 'womb', a term which may simply refer to their dark interior. The other passages seem to use the adjective in an ideological/metaphorical sense, contrasting the 'dark' world of the *dāsas* and *dasyus* with the 'light' world of the *āryas*. As far as I can tell, there is no unambiguous evidence for an awareness of color-related "racial" differences in the Rig-Veda. In fact, the notion "race" is a problematic invention of the colonial period, quite inappropriate, I believe, in the ancient world.

It is, I think, similarly inappropriate to project the supremacist ideology of modern colonial powers like the British into ancient and prehistoric times. True, those defeated in war often suffered a cruel fate, even extinction. At the same time, both "civilized" empires (such as the Roman one) and "barbarian" ones (such as the Huns) were multiethnic, multilingual, and multicultural. War-time alliances kept shifting and could pit members of the same ethnic group against each other (such as Germanic tribes allied with the Huns, and with the Romans). In fact, according to Classical Sanskrit political theory, alliances were to be made with people living on the other side of one's enemy, who would often be ethnically closer to the enemy's party than one's own.

Most important, the Rig-Vedic evidence suggests a fluid situation of this type, in which ethnicity played a relatively minor role. A dāsa, Balbūtha Tarukṣa, is mentioned as patron of a Vedic seer. Numerous passages refer in one breath to dāsa and ārya enemies and in one of these, both types of enemies are referred to as *ádeva* 'godless, infidel'. Especially instructive is the famous "Battle of the Ten Kings" (RV 7:18 with 7:33, 7:83; cf. also Kuiper 1991): On both sides of the bat-

tle we find people with "Aryan" names (such as Vasiṣṭha vs. Bharata) AND with names that sound "non-Aryan" (such as the Sṛñjayas vs. Śimyu).

The picture that emerges is rather different from the one commonly drawn: While there was hostility and warfare between Indo-Aryans and their dāsa/dasyu opponents, there was no social chasm comparable to that between the British colonialists and the South Asian people(s) they subjugated. Whatever the ethnic and linguistic differences, they did not prevent āryas and dāsas/dasyus from making shifting alliances with each other, requiring them to interact bi- or multilingually on a continuing basis.

If, then, there was indeed some kind of (direct or indirect) contact between Indo-Aryans and Dravidians, the linguistic consequence should be expected to have been bidirectional convergence, rather than the unidirectional subversion commonly assumed. This is the hypothesis which I will try to support in the remainder of this paper as regards the origin of Indo-Aryan and Dravidian retroflexion and alveolarization.

5. Retroflexion and alveolarization as convergent developments

As observed in §3, Indo-Aryan clusters of *r* + dental stop underwent a dual development in the Aśokan inscriptions and in later Indo-Aryan, either to dental or to retroflex stop. (Rig-)Vedic evidence shows that this development goes back to Old Indo-Aryan times: As is well known, although the Rig-Veda, and Sanskrit in general, ordinarily retain *r* + dental stop, we find occasional Rig-Vedic forms of the type (6), usually explained as borrowings from a contemporary "Vedic Prakrit".[18] Note especially (6b) vs. (6c) with retroflex vs. dental outcomes of the same element, *kṛta-*.

(6) a. RV kartá 'cavity, hole' > RV kāṭá 'cavity, depth'
b. RV vikṛta 'changed: misshapen' > RV vikaṭa 'hideous, terrible'
c. *kṛta-vat 'having the lucky throw in gambling' > RV kitava 'gambler'

In the remainder of this section I want to advance the hypothesis that these dual outcomes of *r* + dental stop, which we find from Old Indo-Aryan, through the time of Aśoka, to the present day, go back to an intermediate earlier stage with alveolar stops, and that the prehistoric developments giving rise to these alveolar stops, as well as to retroflex stops, were convergent with similar changes in Proto-Dravidian.

In support of this hypothesis, recall that the geographical distribution of dental vs. retroflex outcomes of *r* + dental in the Aśokan inscriptions is remarkably similar to the distribution of dental vs. retroflex outcomes of (geminate) alveolar stop in Dravidian (except in the extreme south which tends to retain the alveolars); see Map V above.

If we assume that *r* + dental changed directly into dental or retroflex in Indo-Aryan, this similarity in distribution would be accidental; but if we hypothesize that the development took place via an intermediate alveolar, then we can explain the similarity as resulting from convergent changes that eliminated alveolar stops in both Indo-Aryan and Dravidian (except in the deep south), in favor of either

dental or retroflex, depending on geographical region, along north-to-south lines.[19] (The fact that "rich" systems with a triple contrast dental : alveolar : retroflex may be relieved by merger of the alveolar with either dental or retroflex is demonstrated by most of Dravidian. Similar developments have taken place in dialectal Norwegian and Swedish; cf. Steblin-Kamenskij 1965.)

Although many of the specific arguments for the hypothesis are circumstantial and speculative, and although there are possible chronological problems, I believe that the Dravidian/Indo-Aryan parallelism, combined with the Vedic evidence on the prehistoric and early historic social relationship between Indo-Aryans and non-Indo-Aryans, makes it worth while to develop this convergence hypothesis so that it can be tested by other scholars.

§ 5.1 addresses the Indo-Aryan and East Iranian evidence in favor of the hypothesis. §5.2 deals with the more controversial issue of explaining Dravidian alveolar and retroflex stops as early Dravidian innovations. §5.3 draws on §§5.1 and 5.2 to set out the hypothesis that the Indo-Aryan/East Iranian and Dravidian developments are a common innovation. §5.4 discusses certain difficulties regarding the origin of the retroflex "triggers" for the Indo-Aryan and Dravidian developments. §5.5 is concerned with problems of chronology. §5.6 discusses alternatives and the consequences that arise from not accepting the convergence hypothesis.

5.1. Indo-Aryan

Let me begin with arguments and evidence that make it possible to support the "alveolarization hypothesis" that *r* + dental stop first changed into alveolar stop in some variety of Old Indo-Aryan and that this alveolar subsequently merged with either dental or retroflex, depending on the dialect. To do so it is necessary to remove several possible obstacles and, in the process, to examine relatively arcane aspects of early Vedic phonology, as well as parallels in the early East Iranian language, Avestan. The "fringe benefit" of this undertaking is that the alveolarization hypothesis raises interesting questions about both the early phonological history of Indo-Aryan and the dialectological or area-linguistic relationship between Indo-Aryan and ancient East Iranian.

As noted earlier, Indo-Aryan retroflexion is considered an innovation by all scholars, whether they attribute the change to subversion or to internal developments. The formulation of the changes in (5), repeated for convenience, is based on Hock 1975 and 1979 (with references).[20]

(5)	I	II	III	IV	V	
a.	(*liǵh-to- >)	*ližd͡ha- >	*liẓd͡ha- >	*liẓd͡ha-		> līḍha- 'licked'
cf.	(*wiǩ-to- >)	*wišta- >	*viṣta- >	viṣta-		= viṣṭa- 'entered'
b.	(*wiǩ-s >)	*wiš̌ >	*wiṣṣ >	*wiṭs		> viṭ 'people, clan' (N sg.)
cf.	(*wiǩ-su >)	*wiš̌u >	*wiṣṣu >	*wiṭṣu		> vikṣu (id.) (L pl.) (→ post-RV viṭ-su)

For present purposes, nothing depends on the specific formulation of the changes. Under any formulation, however, the changes between the last three

stages are the phonologically most important: From III to IV, two different processes introduce retroflex stops. Of these, the one in (5a) consists of a fairly ordinary assimilation of dental stop to preceding retroflex sibilant, a change with parallels in Swedish/Norwegian (except that here retroflexion is introduced after "dark" [l]; see Steblin-Kamenskij 1965); the change in (5b) is more "idiosyncratic", involving a pre-Indo-Aryan dissimilation of geminate sibilants (see Hock 1987[21]). The crucial next step is that from IV to V: Some of the "triggers" for the change are lost; as a consequence retroflex stops are no longer fully predictable and therefore become phonologically significant.

The alveolarization hypothesis can a priori be formulated as involving changes entirely parallel to those in (5a); and this parallelism may be considered an element in favor of the hypothesis. Dental stop assimilates to preceding alveolar *r*,[22] becoming alveolar; loss of the "trigger" *r* makes the alveolar stop (*t̲* etc.) unpredictable and hence phonologically significant. See the formulation in (7).

(7) I II III
*krta > *krt̲a > *ki/at̲a (hence RV kita- : -kat̲a-)

The hypothesis, however, runs into some empirical obstacles. First, Vedic, and following it, later Sanskrit, normally retains *r* + dental stop. If there was in fact assimilation of dental *t* to alveolar *r*, the change should only have progressed to stage II.

This particular obstacle can be taken care of by adopting the common assumption that forms of the type *kitava* and *vikat̲a* are borrowings from "Vedic Prakrits", more vernacular varieties of Old Indo-Aryan that coexisted with the puristic language of the Vedic poets and which we need to posit on independent grounds.[23] All we need to assume is that, in addition to other changes characteristic of the later Prakrits, the Vedic Prakrits carried out the change(s) in (7). The dental or retroflex of "puristic Vedic", then, can be accounted for in one of two ways: Either the Vedic Prakrits had already changed the alveolar into dental or retroflex, depending on dialect; or in the borrowing process the Vedic Prakrit alveolar was nativized variably as dental or retroflex. (The second alternative might receive support from the fact that river names which can plausibly be considered Tibeto-Burman in origin show suffix variation between *-ta* or *-t̲a* (Witzel 1995) which may result from different nativizations of a Tibeto-Burman element *-t̲a*, with alveolar; see note 20 above.)

Even so, we might expect puristic Vedic to at least show traces of stage II, with alveolarization still predictable because the trigger for the change is still present. The Vedic Prātiśākhyas, however, make no mention of an alveolar articulation of dentals after *r*, even though they observe all kinds of other fine phonetic details.

Fortunately, this difficulty, too, is amenable to explanation; but significantly, the explanation raises interesting issues for early Indo-Iranian dialectology, as well as for any contact-induced account of retroflexion and alveolarization, whether of the subversion or of the convergence variety.

As George Cardona has reminded me (p.c. 1991), according to Atharva-Prātiśākhya 1.101-102 and Rik-Prātiśākhya 6.13-14, a svarabhakti vowel is regularly inserted in Vedic recitation between *r* and consonant.[24] As a consequence, dental stops would not be directly preceded by *r* in this variety of Old Indo-Aryan and thus would not become alveolar.

Svarabhakti actually had a more general motivation than just to keep alveolar *r* apart from dental stops: As is well known, *r* is the weakest of the Old Indo-Aryan consonants. (For instance, it is the only consonant that is not permitted as a geminate.) If we assume that it was especially weak in syllable-coda position, we are able to account not only for the fact that the (Vedic) Prakrits lost it in this position (with compensatory length on the preceding vowel if *r* was non-syllabic) but also for the fact that *r* is the most pervasive trigger for the gemination of neighboring consonants in puristic Vedic (a compensation for the weakening of *r* and its loss of mora-bearing ability). Svarabhakti then can be seen as an alternative to *r*-weakening employed in careful recitation: Insertion of the vowel places *r* into the onset of a syllable and thus preserves it from weakening. (See Howell 1991 for Germanic parallels to this dual behavior of *r*.)

The Rig-Vedic cooccurrence of forms like *kartá-* and *kāṭá-* in example (6a) above, then, can be explained as reflecting two different traditions—one being the puristic tradition of mainstream Sanskrit which in Vedic times pronounced *kartá-* as [kar^{a}tá] and in so doing preserved the [r] as well as the dental articulation of *t*, the other a Prakritic tradition which did not have svarabhakti and which therefore permitted *rt* to develop to an alveolar stop (with loss of [r], except for compensatory lengthening).

As it turns out, the early East Iranian language of the Avesta exhibits a similar dual treatment of *r* + consonant. Here, too, we find a general tendency to insert a vowel, generally *ə*, as in (8a). But in combinations of *r* plus voiceless dental stop we find the alternative outcome *š̰*, as in (8b). With other voiceless stops we find a similar variation, as in (8c) vs. (8d).

(8) a. kərəta 'done' (cf. Skt. kṛta)
bərətar 'carrier' (cf. Skt. bhartṛ)
b. aš̰a 'truth' (cf. Skt. ṛta)
xᵛāš̰a 'food' (from xᵛar 'eat')
cf. c. vəhrka 'wolf' (cf. Skt. vṛka)
vs. d. Gath. marəka vs. YAv. mahrka 'destruction'

Of specific interest for present purposes is the dual development of *r* + *t* clusters, because it is highly reminiscent of the relationship between the 'puristic' Rig-Vedic type *kartá* [kar^{a}tá] and the 'Prakritic' type *kāṭá*, see (9).[25] That is, in both languages, combinations of *r* (or *ṛ*) + *t* either are broken up by svarabhakti or are fused, as it were, into a new sound.

(9)	Old Indo-Aryan	Avestan
	kāṭá (from kāṯá)	aš̰a
	kartá [kar^{a}tá]	bərətar

But there is more than simple parallelism: Hoffmann 1958/67, 1971 interprets Av. *š́* as a voiceless, perhaps retroflex, lateral, comparable to Pashto retroflex *ṛ* from *rt*.[26] A priori, of course, it is possible that *š́* designates an alveolar, rather than a retroflex. In that case the assimilation in (7) may have been shared by Indo-Aryan and East Iranian. In this regard it may be significant that the "retroflex" consonants of Pashto and other northwestern languages are commonly described as (post-)alveolar, not retroflex.[27]

Support for considering the two phenomena to be related comes from the fact that there is a remarkable shared idiosyncracy as regards svarabhakti. All four Prātiśākhyas (AP 1.101, RP 6.13-14, TP 21.15-16, VP 4.16) rule out svarabhakti in the context between *r* and a sibilant + stop cluster; cf. (10a). Gatha-Avestan normally does not have *ə*-insertion in the same context; cf. (10b); occasional forms with *ə*-insertion, such as *aibī.dərəštā* can be accounted for as analogical on the model of related forms with legitimate *ə*-insertion, such as the root √*dar**ə**s-* 'see'. Even more important, where **r* + voiceless sibilant is not followed by *t*, *ə*-insertion is absolutely regular. (All of the more than 27 Gatha-Avestan occurrences of such forms have *ə*-insertion.)

(10) a.	Vedic:	No svarabkakti between r/ṛ and Sib. + Stop	
b.	GAvest.:	Normally	darštōišca (Y 33.6) etc.
		Occasionally	aibī.dərəštā (Y 21.2, 50.5) motivated by √dar**ə**s-

The lack of vowel insertion in (10) is especially noteworthy since the clusters involved are more complex than those in which insertion does take place. It is therefore highly unlikely that this restriction on svarabhakti—and svarabhakti itself—are independent phenomena in Vedic and Avestan.

We can thus conclude that early Indo-Aryan and East Iranian share a dual treatment of *r* + consonant, one with svarabhakti, the other without. While the precise conditions for the choice between these two treatments is not entirely clear in Avestan, in Indo-Aryan it appears to be socially conditioned: Svarabhakti is a feature of puristic Vedic, its absence a feature of more vernacular Vedic Prakrits.

In the case of *r* + *t* this variation is responsible for a dual development. Svarabhakti permits retention of both *r* and *t*; its absence results in an interaction between the two consonants. The outcome of this interaction most likely was an alveolar, given the evidence for a retracted articulation of Avestan *š́* and the fact that the variation between dental and retroflex in puristic Vedic can be explained as reflecting an earlier alveolar in the Vedic Prakrits.

The present account raises interesting questions regarding prehistoric linguistic contacts in northwestern South Asia and neighboring areas. But its significance and fruitfulness extend farther: By pointing out parallel phenomena which ignore the boundary between Indo-Aryan and Iranian and by providing a UNIFIED explanation for these phenomena it raises important questions for early Indo-Iranian dialectology and/or areal linguistics.

5.2. Dravidian

Most Dravidianists would accept that morphophonemic alternations of the type (11) show that certain instances of Dravidian alveolar and retroflex stops result from secondary, assimilatory developments; see e.g. Zvelebil 1970 and Krishnamurti In Press (as well as Tikkanen 1987).

(11) Dravidian retroflexion and alveolarization (data from Tamil)

a.	ceṉ 'go'	+ -t- + -ēṉ	:	ceṉṯēṉ	'I went'
	uṇ 'eat'	+ -t- + -ēṉ	:	uṇṭēṉ	'I ate'
	kol 'kill'	+ -t- + -ēṉ	:	koṉṯēṉ	'I killed'
	āḷ 'rule'	+ -t- + -ēṉ	:	āṇṭēṉ	'I ruled'
b.	kal 'stone' + tūṇ 'pillar'		:	kaṯṯūṇ	'stone pillar'
	kaḷ 'booze' + tantāṉ 'gave'		:	kaṭṭantāṉ	'gave booze'

Following Krishnamurti 1961, Zvelebil 1970:178-180[28] attempts to extend this explanation to account for root variations of the type (12a), with final alveolar or retroflex sonorant alternating with alveolar or retroflex stop. As shown in (12b), the stops of these forms can be derived from the alternating sonorants by "fusions" parallel to those responsible for the alternations in (11b). Subsequent to fusion, the resulting forms evidently were reinterpreted as simple roots in their own right, ending in alveolar or retroflex stop. (See Krishnamurti In Press for a comprehensive discussion of this reinterpretation, its pervasive nature in Dravidian, and its consequences for Dravidian morphology.)

(12) a.

Tam. kāl 'air, wind'	:	kaṯṯu (id.)
Tam. uruḷ 'to roll (itr.)'	:	uruṭṭu (id., trans.)

b. kaṯṯu < kal-tu
uruṭṭu < uruḷ-tu

The discussion of Krishnamurti (In Press, see also 1995) almost exclusively deals with verbal roots; and perusal of DEDR yields ample evidence for verbal root alternations that can be explained along the lines of (12b). But Zvelebil's Tam. *kāl : kaṯṯu* shows that alternations also occur in nominal roots; and while such alternations do not appear to be as numerous, the examples in (13) illustrate that they are not limited to just one or two words. (The examples are drawn from Zvelebil 1970 and from DEDR; numbers in parentheses indicate the entry in DEDR.)

(13) Tam. il 'house' (494) : iṯai 'inside of a roof, eaves of a house ...' (528)
Tam. āḷ 'man ...' (399) : āṭṭi 'woman ...' (400)
Tam. kāl 'air, wind' : kaṯṯu (id.)
Tam. cil 'some, few, small' (1571) : ciṯu 'small, etc.' (1594)
Tam. col 'fine rice' (Zvelebil) : cōṯu 'boiled rice' (2897)
Tam. neru-nal 'yesterday' (3578) : neṯṯu 'recently' (ibid.)
Tam. pāṇ 'song, melody' (4068) : pāṭu 'sing, chant ...' (4065
Tam. paḷḷi 'hamlet' (4018) : pāṭi 'town, city, hamlet' (4064)
Tam. purai 'ulcer, fistula' (4297) : puṯṯu 'anything scrofulous or cancerous' (4336)
Tam. puṟai 'hole, tube ...' (4317) : puṭṭi 'flask, bottle' (4265a)
Tam. peru, pēr 'great' (4411) : peṯṯam 'greatness' (4425)

Tam. pēṟai 'box, chest' (4442) : peṭṭi 'box, chest ...' (4388)
Tam. pōl 'hollow object' (4604b) : pōṭai 'hole, hollow' (4604a)
Kan. mala 'other, next' (4732) : Ta. maṭu 'another, other, next' (4766)
Tam. vaḷ 'thong, lash' (5305) : vaṭam 'cable, cord, bowstring ...' (5220)
Tam. vīṟ/viṟutu 'aerial root' (5431) : viṭutu (id.; ibid.)

For Zvelebil, accounting for alternations of this type actually was of minor significance. His major claim is the 'highly speculative and hypothetic' proposal that many (though not all) consonant clusters and geminate consonants of Dravidian can be 'further analysed ... as results of assimilations' (178).

It remained for Tikkanen 1987:285 to interpret Zvelebil's account as supporting my earlier, rather poorly substantiated claim that all of Dravidian retroflexion and alveolarization is the result of secondary developments. As Tikkanen states it,

> Both alveolarization and retroflexion of dental stops in [P]roto-Dravidian are ... reflections of the same coarticulative process, i.e. the retraction of the point of articulation after retroflex and alveolar sonorants (with or without subsequent merger) ... (285)

The processes involved are summarized in (14). As in Indo-Aryan, the first step (stage I to stage II) consists in straightforward assimilations. The loss of some of the triggers for the change (II to III) then makes the alveolar and retroflex stops unpredictable and therefore phonologically significant.[29]

(14)

		I		II		III
a.	Retroflexion:	*ṇt	>	ṇṭ		
		*ḷt	>	ḷṭ	>	ṭ
		*ṟt	>	ṟṭ	>	ṭ
b.	Alveolarization:	*n̲t	>	n̲t̲		
		*l̲t	>	l̲t̲	>	t̲
		*rt	>	rt̲	>	t̲

Dravidianists like Zvelebil, Krishnamurti, and Emeneau acknowledge that many instances of Dravidian alveolar and retroflex stops can be explained along the lines of (14), but they are evidently not prepared to accept the view that, in principle, ALL Dravidian alveolar and retroflex stops are amenable to such an explanation. For instance, Krishnamurti 1995 distinguishes between two types of alveolars and retroflexes, the "alternating" type (12)-(13) and another type for which there is no evidence of alternation. Only the alternating type is considered to result from the changes in (14), within the linguistic history of Dravidian; the non-alternating type, by contrast, is believed to be directly inherited from Proto-Dravidian.

This is indeed a possible interpretation of the evidence. But a simpler interpretation would be that ALL alveolar and retroflex stops result from changes of the type (14).

The alternations in (12)-(13) do not seem to be confined to any particular subgroup of Dravidian and therefore must be considered a feature common to all of

Dravidian. No evidence requires the assumption that they arose in a post-Proto-Dravidian stage. It is therefore entirely possible that, just like the "non-alternating" type, they go back to Proto-Dravidian.

Moreover, lack of alternation does not guarantee different origin from the alternating type. It is certainly possible that "non-alternating" forms have the same origin as alternating ones and that the two types merely differ in terms of whether or not the original forms with root-final alveolar or retroflex sonorant happen to have been preserved (in meanings that are still relatable to those of the derived forms). In this regard note that a large number of early Indo-Aryan (Vedic) forms with voiced retroflex stop are synchronically "non-alternating"; it is only because we have access to earlier, pre-Proto-Indo-Aryan stages that one can propose for some of them the same historical derivations as for synchronically "alternating" ones (see Mayrhofer 1986- : 1: 69, 187, 204, 313, 385 (with Vine 1987), 413, 415; 2: 49, 136, 326, 357, 387). Even with this access to earlier stages, a number of "non-alternating" voiced retroflex stops (and other retroflex consonants) remain unexplained. Significantly, however, it is because we have access to these earlier stages that we know Indo-Aryan retroflexion to be an innovation.

A possible counterargument is that, in contrast to early Indo-Aryan, the number of Dravidian "non-alternating" forms is very large. But given the relatively late attestation of the Dravidian languages, the large number of "non-alternating" forms may simply result from the fact that over the centuries and millennia, many of the sonorant-final base forms have become obsolete, or that their meanings have diverged too much to permit linking them to roots in alveolar or retroflex stop. If our knowledge of Old Indo-Aryan had to be derived solely from Middle or even Modern Indo-Aryan sources, the number of "non-alternating" retroflexes would no doubt be much greater, too.

Moreover, once phonologically significant alveolar and retroflex stops have arisen, it is possible to extend these stops to new contexts. In Indo-Aryan, for instance, Hoffmann 1941 argues that retroflex *-ṇḍ-* is common in words belonging to two semantic categories, of "roundness" and of "breaking, crushing", and may have been analogically introduced in many of these words because of their meaning; see note 20 above. Indo-Aryan retroflex consonants are also commonly used in newly created onomatopoeia (Hoffmann 1956). And they are found in many suspected borrowings. Similar developments may have introduced some of the "non-alternating" alveolar and retroflex stops in Dravidian. Note in this regard that retroflex consonants are very common in Emeneau's 1969 collection of Dravidian onomatopoeia.

The hypothesis that all Dravidian alveolar and retroflex stops are an innovation along the lines of (14) moreover provides a motivation for the often-noted absence of these stops in initial position: The clusters that gave rise to them, i.e. sequences of sonorant followed by dental stop, are highly unlikely to have occurred in initial position. For similar reasons alveolar and retroflex stops are barred from initial position in Norwegian and Swedish, and the Rig-Veda has initial retroflexion only in one word (*ṣaṭ* 'six' and derivatives) where it results from distant assimilation.

5.3. The "convergence hypothesis"

Accepting the hypothesis that all of Dravidian alveolarization and retroflexion is an innovation and that varieties of early Indo-Aryan and East Iranian changed *r* + dental stop to alveolar stop has important consequences.

As shown by the comparison in (15) of the major changes[30] that gave rise to Dravidian and Indo-Aryan/East Iranian alveolar and retroflex stops, these changes are remarkably similar to each other.

(15)		Dravidian	Old Indo-Aryan
a.	RETR.	*nt > ṇṭ *lt > l̤t > ṭ	*st(h) > ṣṭ(h) *zd(h) > ẓḍ(h) > ḍ(h)
b.	ALV.	*n̲t > n̲t̲ *lt > l̤t̲ > t̲	*rt(h) > rt̲(h) > Ved. Pkt. t̲(h)) *rd(h) > rd̲(h) > Ved. Pkt. d̲(h)

The parallelism of these changes is striking enough to require a reassessment of the prehistoric linguistic interaction between Dravidian and Indo-Aryan/East Iranian.

As we have seen in §2, the early Dravidian and Indo-Aryan synchronic systems, as they are usually posited, are sufficiently different that antisubversionists may doubt the cogency of the Dravidian subversion hypothesis and would therefore attribute the presence of retroflex consonants in both groups to chance.

Such a chance explanation becomes extremely difficult to justify if the early Dravidian and Indo-Aryan/East Iranian systems result from the changes in (14); for the parallelism of the developments is simply too great. On both sides, dental stops assimilate to the same class of preceding consonants—alveolars and retroflexes. On both sides, the preceding consonants are higher in sonority than the dental stops—alveolar and retroflex sonorants in Dravidian, alveolar sonorant and retroflex continuant in Indo-Aryan/East Iranian. And on both sides, the same thing happens to make the results of assimilation unpredictable and therefore phonologically significant—loss of some of the triggers for the changes.

Even the differences between the two early systems becomes smaller, since under the present hypothesis, an alveolar series is found, not only in Dravidian, but also in Indo-Aryan/East Iranian. Any remaining differences, such as idiosyncratic Dravidian *ṟ* vs. idiosyncratic Indo-Aryan *ṣ/ẓ*, are of comparatively minor significance and can be attributed to preexisting differences between the two groups. (But see also §5.4 below.)

Now, the changes in (14) are innovations in both Dravidian and Indo-Aryan/East Iranian. Moreover, there is no evidence that would force us to locate the origin of the changes in one or the other linguistic group and to assume that they spread to the other group by subversion. Rather than making an arbitrary choice it is preferable to consider the changes to reflect CONVERGENCE between the two groups. This interpretation finds support in the fact that, as we have seen in §4, the evidence of the Rig-Veda is compatible with assuming a social situation

favoring the continuing bilingualism that would encourage convergence. Moreover, as shown in Hock 1988a, it is often difficult in convergence areas to pinpoint a particular language that may have been the source for a given shared innovation.

The convergence hypothesis, if correct, is significant, for it suggests that the social relationship between Indo-Aryan and non-Indo-Aryan speakers in early India was not substantially different from what it is today—a relationship of (near-) equality that encouraged continuing bilingualism, rather than the traditional picture of marauding Indo-Aryan invaders suppressing an indigenous Dravidian population and unilaterally forcing it to learn their language.

Note further that convergence does not require direct contact. As shown in Hock 1988a, languages in a convergence area behave very much like dialects in a monolingual situation; innovations can spread from language to language, eventually covering a vast territory. The convergent changes in (14) therefore could have resulted from mediated contact, possibly involving some ancestral form of Burushaski, Tikkanen's unknown northwestern substratum, and yet other languages. As a consequence, the convergence hypothesis does not depend on a resolution of the—highly controversial—question of whether there was a Dravidian presence in the prehistoric northwest (see §3)—a clear advantage of the convergence hypothesis over the subversion account.

A quasi-dialectological view of convergence further makes it possible to account for a fact that has so far been glossed over: While early Indo-Aryan has both retroflexion and alveolarization, Avestan offers no evidence for an assimilation of dental stop to preceding retroflex sibilant. In fact, most varieties of Middle and Modern East Iranian likewise do not exhibit such a change. Exceptions are Middle Iranian Saka (*hīṣta* 'sent', related to Skt. *iṣṭa* 'sent') and Modern Pashto (with *ləṛ* 'ache' < **duždah*) and Sanglīčī-Iškāsmī (with *ṭ*< *ṣṭ* < *šť*); see Emmerick 1989, Skjærvø 1989a,b, Payne 1989, and the discussion and references in Tikkanen 1987: 289.[31] But the Pashto development is limited to voiced sibilant + dental stop; voiceless **šť* results in dental *t*; and the Sanglīčī-Iškāsmī retroflex outcome *ṭ* alternates with dental *t*. Even Nuristani, somewhat intermediate between Iranian and Indo-Aryan, only has variable traces of this development; see Tikkanen 1987:287-288 (with references). In fact, the earlier change of *šť* to *ṣ*, which produced the trigger for retroflexion, likewise shows more limited or variable distribution in Nuristani and East Iranian.

Under the convergence hypothesis, the more limited and variable distribution of the change *šť* > *ṣ* and of the retroflexion of dental stop after *ṣ* in Nuristani and East Iranian finds a ready explanation as a peripheral, transition-area phenomenon. The core area of the change must have been in South Asia proper, from which the change spread only incompletely to the Nuristani and East Iranian languages on the northwestern periphery, before coming to a complete halt in geographically even more remote Iranian territory.

Alveolarization after *r* evidently was more "vigorous" in the northwest.[32] Plausible effects of the change are found in Saka (Emmerick 1989: 215), modern

Sanglīčī-Iškāsmī, Yidgha, Pashto, and Parachi (Skjærvø 1989), as well as of course in ancient Avestan. In Avestan, Saka, and Yidgha, however, the change is restricted to clusters of *r* + voiceless dental stop. This restriction would, again, be consonant with the view that East Iranian was on the periphery of South Asian convergence and therefore only partly affected by it.

The convergence account of alveolarization and retroflexion, thus, proves to be multiply fruitful. Not only does it explain the otherwise inexplicable parallelism in the changes that gave rise to alveolar and retroflex stops in Dravidian and Indo-Aryan/East Iranian, it also provides a principled account for the more variable effects of these changes in East Iranian and other languages on the northeastern periphery. Moreover, it does so without requiring the highly controversial assumption of direct prehistoric contact between Dravidian and Indo-Aryan—a clear advantage over the subversion hypothesis. Finally, it is consonant with the Rig-Vedic evidence for the social relation between Indo-Aryans and non-Indo-Aryans, a relationship that does not differ significantly from what we encounter in later, historic times.

5.4. The triggers for retroflexion and alveolarization

In spite of all its advantages, however, the convergence hypothesis encounters some problems of its own. The most important of these concerns the origin of the retroflex and alveolar continuants and sonorants that triggered the development of retroflex and alveolar stops.

On the Indo-Aryan side the picture is reasonably clear: Indo-Aryan *r* no doubt was alveolar to begin with (Hock 1992c). Indo-Aryan retroflex *ṣ/ẓ* continues Proto-Indo-Iranian *š́/ž́*, and as argued in Hock 1975, 1984, the change of *š́/ž́* to retroflex can be motivated by the principle of polarization (for which see the more general discussion in Hock 1986/1991), to maintain the contrast with distinctively palatal *ś* from Proto-Indo-Iranian *ć*. Tikkanen 1987: 289 argues against this explanation by observing that Proto-Indo-Iranian *ć* does not change to palatal *ś* in Nuristani, which nevertheless has retroflex *ṣ*. But as he himself notes, the Nuristani counterparts of Proto-Indo-Iranian *š́/ž́* show considerable fluctuation between retroflex, a sibilant marked *š́*, and even dental sibilant.[33] (In fact, Nuristani raises difficulties for Indo-Iranian comparative linguistics, including the usual reconstruction *š́/ž́*; see e.g. Morgenstierne 1975a.) This variability is compatible with the peripheral position of Nuristani noted in §5.3. It is therefore possible that the occasional Nuristani retroflex counterparts of *š́/ž́* result from the spread of the Indo-Aryan change *š́/ž́* > *ṣ/ẓ*. In that case, there would be no problems, since as noted, polarization is well motivated in Indo-Aryan.

But what are the sources for the Dravidian alveolar : retroflex contrasts in the sonorants—*r* vs. *ṛ*, *l* vs. *ḷ*, *n*/*n̰* vs. *ṇ*?

As far as I can tell, there is nothing in the comparative Dravidian literature to suggest that this contrast is secondary, comparable to the one between Old Indo-Aryan retroflex *ṣ*, palatal *ś*, and dental *s*.[34] However, examination of the data in DEDR yields a number of semantically relatable entries whose major difference

lies in the presence of a retroflex sonorant in one entry and a corresponding alveolar sonorant in the other; see the data in (16). If these, and perhaps other, similar doublets, should indeed turn out to be related, then the retroflex : alveolar contrast of the Dravidian sonorants must be secondary, resulting from some kind of phonological split (whose conditions may no longer be recoverable). In that case, the prehistoric parallelism between Indo-Aryan (and East Iranian) and Dravidian might extend even farther than envisaged in this paper.

(16) Tam. aṉai 'approach...' (120) : Tam. aṉuppu 'send (off)' (329)
Tam. eṉ 'thought' (793) : Tam. eṉ 'say so, utter ...' (868)
Tam. iṉai 'join together'(457) : Tam. iṉam 'class, group'(531)
Kan. gāḷ(i) 'air, wind' (1499) : Tam. kāl 'air, wind' (1481)
Tam. kāṟ 'blackness, blemish ...' (1494) : karu 'black' (1278a), kār 'blackness ...' (1278c)
Tam. koḷ 'strike, hurt' (2152) : Tam. kol 'kill, murder' (2132)
Tam. naḷ 'night' (3621) : Tam. nallam 'blackness' (3613)
Kan. taḷisu 'pound, beat' (3130) : Tam. tallu 'beat, crush' (3105)
Tam. paṟiccu 'praise' (4003) : Tam. paracu 'praise' (3951); cf. pārāṭṭu (4092) 'applaud'
Tam. puṟai 'hole, tube' (4317) : Tam. purai 'tube ...' (4197)
Tam. muṟu 'all, entire' (4992) : Tam. murañcu 'be full, abundant' (4970)
Tam. muṟaṅku 'roar ...' (4989) : Tam. mural 'make noise ...' (4973)
Tam. vaṟaṅku 'move ...' (5292) : Tam. var- 'come ...' (5270)
Tam. vāḷ 'lustre, splendor' (5377) : Tam. vāl 'whiteness, purity' (5364)

While to my knowledge the alternations in (16) have not been noted in earlier Dravidianist literature, an alternation that can be linked with this type has received attention, although it also has been subject to some disagreement. As Subrahmanyam 1983:350 reports, Krishnamurti 1961 'on the basis of a small number of examples, talks about alternation of *ṯ with *ṭ ...'. The alternations are given in fuller detail in Zvelebil 1970:98 and 102 (see also 178-179); see example (17) which also includes references to DEDR. As Zvelebil notes, the alternation is especially common in Telugu, 'where verb bases with *ṯ have transitives with *ṭ. In To[da] there are also traces of this alternation' (102).

(17) Tam. kaṭi 'chew, bite ...' (1390) : kaṭi 'bite, bite off' (1124)
Tam. vaṭa 'dry up, shrink' (5320) : vāṭu 'wither, fade' (5342)
Toda pïṛy 'dust' etc. (with *ṯ according to Zvelebil) (4481) : Tam. poṭi 'powder, dust ...' (ibid.)
Toda kwïḍy 'a family of children' (1655)[35] : Tam. kuṭi 'house, family' (ibid.)
Tam. ciṯu 'small, etc.' (1594) : ciṭṭu 'anything small' (2513)

While Krishnamurti and Zvelebil evidently are convinced that alternations of the type (17) are not just accidental, Subrahmanyam considers the connection doubtful 'since the two sounds are kept distinct in numerous etymologies' (1983: 350). And Burrow and Emeneau 1984 in their introduction to DEDR consider the explanation of the alternation 'still uncertain' (xv).

The last item in (17) makes it possible to argue that the connection is valid (even though the explanation may be different from Krishnamurti's). What is relevant here is that the two stop-final roots coexist with a third root which ends in alveolar sonorant: Tam. *cil* 'some, few, small' (1571); compare perhaps also Tam. *cil* 'small piece; potsherd ...' (1577).

The relationship between *cil* and *cil̲u* is, of course, of the type (13) above, where *cil̲u* can be derived from earlier **cil-tu*, with alveolar root-final sonorant followed by dental suffix. (In fact, *cil* : *cil̲u* is included in the examples under (13).) Given the evidence in (16) it is now possible to account for the form *cil̲ṭu* as derived from a parallel form **cil̤-ṭṭu*, with root-final retroflex sonorant followed by dental suffix. By extrapolating from this well-supported case it is possible to account for the other pairs of forms in (17) under the following assumptions:

- Proto-Dravidian had an alternation of root-final alveolar and retroflex sonorants (whose origin is at this point obscure).
- Just like other root-final alveolar and retroflex sonorants these alternating sonorants could be extended by dental stops and could thus yield alveolar and retroflex stops.
- As no doubt happened with many roots in "non-alternating" alveolar and retroflex stops, the original sonorant-final root may have become obsolete, thereby making it appear that in most cases the alveolar : retroflex stop alternation in (17) is primary, rather than secondarily built on an original alveolar : retroflex sonorant alternation.

The fact that the hypothesis of a Proto-Dravidian root-final alveolar : retroflex alternation thus helps explain the alternations in (17) shows that the hypothesis is a fruitful one.

What is most significant for present purposes, however, is that the hypothesis raises an interesting question regarding the prehistory of South Asian alveolarization and retroflexion. Given that alveolar *r* and *l* are less "marked" than their retroflex counterparts, it is reasonable to explain the alternation as the result of a change from alveolar to retroflex sonorant (under as yet unknown conditions). The triggers for stop retroflexion, then, are the result of an innovation. Now, the Indo-Aryan trigger for stop retroflexion, the retroflex sibilant (whether voiced or voiceless), likewise results from an innovation. What, then, is the likelihood that the Dravidian and Indo-Aryan innovations were independent from each other? Should we conclude that these changes, too, were convergent? And if so, what are the implications for the chronology of South Asian alveolarization and retroflexion?

I do not have any answers to these questions, and perhaps it will never be possible to give a satisfactory reply. But the fact that the hypothesis advanced in this paper encourages such questions may be taken as a further element in its favor.

5.5. Problems of chronology

Beyond the somewhat hypothetical chronology problems raised toward the end of the preceding section, there is a much more concrete chronological prob-

lem. Recall that the starting point for this paper's hypothesis consisted in the similar geographic distribution of dental and retroflex outcomes of alveolar stops in Dravidian and *r* + dental stop clusters in Indo-Aryan. This similarity was explained under the assumption that Indo-Aryan *r* + dental stop did not directly change to dental or retroflex, but that it did so via an intermediate stage with alveolar stop. The alternative would have been to consider the similarity to be accidental.

From the geographical perspective, this line of argumentation is quite reasonable. But the chronology creates greater problems: Even if we assume that Vedic forms like the *kitava* and *vikaṭa* of example (6) are to be explained as different nativizations of Vedic Prakrit forms with alveolar stops (see §5.1), the evidence of the Aśokan inscriptions shows that by the third century BC, the dialectally differentiated merger of alveolar stop with dental or retroflex had been completed. In the oldest stages of the literary Dravidian languages, however, distinct alveolar geminates are maintained not only in Old Tamil, but also in the earliest records of Kannada, i.e. as late as the ninth to tenth century AD. Only Old Telugu (7th century AD) no longer distinguishes alveolar from retroflex geminates. (See Zvelebil 1970:100.) The evidence of Kannada suggests that the Dravidian merger took place some thousand years after the Indo-Aryan one.

Does this mean that we have to consider the geographical similarities in the distribution of dental and retroflex outcomes to be accidental? If so, do we have to abandon the idea that Indo-Aryan *r* + dental stop changed to dental or alveolar stop via an intermediate alveolar stop? And what are the effects for the convergence hypothesis of this paper?

The chronological difference is indeed troublesome. But there are possible ways of getting around this difficulty. One possibility is that the Old Kannada texts reflect a conservative form of the language which retained the dental : alveolar distinction, while the popular language had long abandoned it—an early stage of diglossia.

A second possibility lies in taking a closer look at the geography: Within the literary Dravidian south, the merger of geminate alveolars with dental or retroflex is even later in Tamil than in Kannada; and Malayalam still preserves geminate alveolars. That is, the merger appears to have been spreading from north to south—and at a fairly slow pace. If we back-project the direction of the spread, we will eventually reach Indo-Aryan territory; and we may hypothesize that the change originated there. Now, except for "transplanted" texts like those of Yerragudi, the Aśokan inscriptions come from locations considerably to the north of the literary Dravidian languages. If the spread of the merger was as slow-paced in its early Indo-Aryan stages as it was later in southern Dravidian, it is possible to speculate that the time difference between the Kannada merger and the Aśokan merger results from the interaction between geographical distance and the slow pace of spread.

5.6. Alternatives and their consequences

If the explanations in the preceding section for the time difference between Dravidian (Kannada) and Indo-Aryan (Aśokan) merger of alveolar with dental or retroflex are not accepted, it becomes necessary to examine alternative accounts and their consequences.

An obvious consequence of rejecting the explanations would consist in rejection of the significance of the geographical alignment between the dental and retroflex outcomes of Dravidian alveolars and Indo-Aryan *r* + dental stop clusters. Such a rejection would entail an alternative interpretation of the geographical evidence as being the result of chance.

Now, the geographical alignment was an important building block for the Indo-Aryan/East Iranian "alveolarization" hypothesis. If the alignment is considered to be due to chance, one might be tempted to reject the "alveolarization" hypothesis, too, and claim instead that *r* + dental stop changed directly to Indo-Aryan dental or retroflex, without an intervening alveolar stage. The Avestan development of *rt* to *š̌*, then, might either be an unrelated phenomenon or, if related, simply another instance of *r* + dental stop directly going to retroflex.

If the "alveolarization" hypothesis is rejected, then of course the prehistoric parallelism between Dravidian and Indo-Aryan/East Iranian is diminished. As a consequence, the "convergence" hypothesis might be rejected, too. Additional reasons for such a rejection might come from the Dravidianist side, by insisting that "non-alternating" alveolar and retroflex stops are inherited from Proto-Dravidian, rather than the result of assimilations between alveolar or retroflex sonorants plus dental stops.

In that case, we might have to return to earlier subversionist accounts of Indo-Aryan and East Iranian retroflexion. Moreover, we would have to choose between the simple Dravidian subversion hypothesis favored in traditional accounts and Tikkanen's hypothesis of an unknown northwestern substratum.

What would be the consequences of these various alternatives to the hypotheses presented in this paper?

Most obviously, return to subversionist accounts would mean a return to all the difficulties that have been observed for such accounts. To my mind the most important among these is the fact that unilateral subversion is not what we would expect, given the Rig-Vedic evidence on the social relationship between Indo-Aryans and non-Indo-Aryans, as well as the general uncertainty as to the identity of these non-Indo-Aryans (were they Dravidians, Mundas, or speakers of yet other languages?).

Beyond that, several steps in the arguments against the hypotheses of this paper call into the question the very foundations of subversionist claims:

If we attribute the geographical alignment of dental vs. retroflex outcomes of Dravidian alveolars and Indo-Aryan/East Iranian *r* + dental to chance, by what right, then, do we decide that the similarities between early Dravidian and Indo-

Aryan retroflexion can NOT be due to chance? We would have to develop a much better theory of chance similarities before we can make such a decision without appearing to be arbitrary or self-serving. (True, there is a chronological problem with the geographical alignment; but for all we know, there may have been similar chronological problems as regards Indo-Aryan retroflexion. The fact that we do not have access to relevant information on the chronology of retroflexion does not necessarily give us license to assume that there were no problems.)

Rejecting the "alveolarization" hypothesis likewise raises questions about chance: Given the alveolar articulation of Indo-Aryan *r*, is it likely that combinations of *r* + dental stop directly went to dental or retroflex, rather than to alveolar—especially in light of the fact that Dravidian and Norwegian/Swedish furnish precedents for dialectally differentiated merger of alveolar with either dental or retroflex? Moreover, rejection of the hypothesis makes the early Indo-Aryan phonological system more different from that of Dravidian, and thereby reduces the plausibility of hypotheses that want to link Indo-Aryan retroflexion to Dravidian, whether by subversion or by convergence.

Rejection of the hypothesis that all Dravidian alveolar and retroflex stops result from assimilations of dental stops to preceding alveolar or retroflex sonorants has consequences, too, since it rejects a simple, general account in favor of a more complicated one, deriving some alveolar and retroflex stops by assimilation, but others by inheritance from Proto-Dravidian. Opting for the more complex account calls into question a fundamental assumption of subversionists like Emeneau 1971b and Thomason & Kaufman 1988 that the Dravidian subversion hypothesis should be accepted because it is the SIMPLEST account. True, this is not necessarily an argument against subversion, since as noted in Hock 1996, there are independent reasons for doubting the cogency of claims based entirely on simplicity; but it does constitute a problem for subversionist argumentation. (Ultimately, the issue is not merely one of simplicity, but of "Occam's Razor" which states that elements in an argument should not be multiplied WITHOUT NECESSITY. While the "necessity rider" clearly is relevant, it also opens the way for disagreement over when a more complex argument accounts better for the data than a simpler one.)

At numerous points in this paper I noted that the hypotheses advanced in the paper are fruitful in that they explain interesting linguistic issues that go beyond the question of convergence or subversion. Rejection of these hypotheses would require offering alternative accounts for these issues.

In some cases, this should not be too difficult. For instance, my claims concerning the alternations in (16) and (17) can be maintained, even if the hypothesis is rejected that all Dravidian alveolar and retroflex stops are the result of innovation. But note that examples such as the last one in (17) suggest that the number of lexical items with secondary, rather than "non-alternating, inherited", alveolar or retroflex stop may be much larger than is commonly assumed. And this fact may raise questions about the claim that there were such "non-alternating, inherited" stops.

6. Summary and conclusions

This paper proposes a set of related hypotheses concerning South Asian retroflexion that differ significantly from earlier views. In contrast to earlier Dravidian subversion explanations and in contrast to simple rejections of such explanations, I argue for CONVERGENT changes which introduced not only retroflexion, but also alveolarization, both in Dravidian and in Indo-Aryan/East Iranian. As common in convergence areas, the changes lost momentum on the periphery, in East Iranian and Nuristani. Later convergent developments led to the merger of alveolar stops with dental or retroflex stops in Indo-Aryan and most of Dravidian. The hypothesis of convergence, rather than subversion, finds support in the Rig-Vedic testimony regarding the social relationship between Indo-Aryans and non-Indo-Aryans.

My claims and findings are significant on several counts. First, they suggest that the prehistoric relationship between Indo-Aryans and non-Indo-Aryans was not substantially different from what we find in observable history—a relationship that encouraged extended bilingual interaction with bidirectional linguistic consequences, rather than the usually assumed forced shift from non-Indo-Aryan to Indo-Aryan with unilateral linguistic consequences. By drawing on the evidence of Old Iranian Avestan, I expand the horizon for convergence (or subversion) hypotheses to East Iranian and, in so doing, raise interesting questions about early dialectal or bilingual interactions in Indo-Iranian. The assumption of convergence, rather than subversion, makes it possible to provide an explanation of this relationship (in terms of the peripheral location of East Iranian). Moreover, because convergence does not require direct contact, the hypothesis avoids the difficulty encountered by subversion hypotheses that independent evidence for prehistoric Dravidian/Indo-Aryan contact is highly controversial and that vocabulary evidence favors Indo-Aryan contact with neither Dravidian nor Munda, but possibly with some unknown northwestern language (Tikkanen 1987, 1988). Finally, in the process of developing the convergence hypothesis I advance a number of subsidiary arguments which shed an interesting light on Indo-Aryan and Dravidian historical phonology.

While reactions to earlier versions of this paper by advocates of Dravidian subversion suggest that they will not be convinced by my claims, I hope that they will consider the arguments presented in this paper to be worthy of serious discussion. Whatever the outcome of the discussion, if it is supported by alternative explanations and new data, it is bound to advance our understanding of the prehistory and early history of South Asia and of historical linguistics in general.

Notes

* This paper grows out of continuing research on the issue of prehistoric and early historic South Asian convergence. The present paper is a thorough revision of Hock 1995, which itself is a revised version of a paper read at the 1992 Annual Meeting of the American Oriental Society. A related paper has been presented on

numerous occasions, including the 1993 Linguistic Institute at Ohio State University, lectures at the Universities of Hamburg and Freiburg, and most recently at the November 1994 International Seminar on 'Ideology and Status of Sanskrit in India and Asia', International Institute for Asian Studies, Leiden (NL). (See Hock 1996.) Part of the research has been supported by grants from the University of Illinois Research Board and a spring 1995 sabbatical leave. I am indebted to Rahul Peter Das for kindly making a copy of Hoffmann 1941 available to me. I am also grateful for comments I received on earlier versions of this paper and related papers, especially from Lyle Campbell, George Cardona, Jan Houben, Murray B. Emeneau, Bh. Krishnamurti, and Sarah Thomason. I know that the three last-mentioned scholars do not agree with many of the claims in this paper; but I sincerely hope that our disagreement will stimulate further fruitful discussion. As usual, the responsibility for any errors and omissions rests with me.

[1] Nahali might constitute the remnant of yet another language family (Kuiper 1966 with references). Witzel 1995 further adds Kusunda in central Nepal, as well as possibly other languages, including that of the Veddas.

[2] See e.g. Masica 1976 and Hock 1988a.

[3] Interestingly, in other areas of close contact, Tibeto-Burman has converged with Indo-Aryan, by acquiring the contrast.

[4] For Emeneau's 1974 lexical-syntactic arguments regarding Skt. *api* : Drav. *-um*, see Hock 1975 and Gil 1994 (apparently independent of Hock 1975). For Abbi's 1992 monograph on 'reduplicated' structures, see the review in Hock 1993b.

[5] See Hock 1986/1991: Chapter 16 for general discussion. The notion "convergence" and "convergence area", Germ. Sprachbund, was introduced by Jakobson 1931 and Trubetzkoy 1931. Emeneau has introduced and popularized an alternative term, "linguistic area". For treatments of more recent South Asian convergence, both "global" and more localized, see e.g. Emeneau 1989, Gumperz & Wilson 1971, Krishnamurti 1991, Pandharipande 1982.

[6] The precise manner in which Indo-Iranian **š̌* and **ž̌* gave rise to Indo-Aryan (Vedic Sanskrit) retroflexion, the conditions under which the development took place, and the extension of retroflexion beyond its original domain are still a matter of controversy. For earlier views and literature see Wackernagel 1896, espcially pp. 164-177, and Debrunner's supplement of 1957. More recent literature is found, and referred to, in Kuiper 1967b, Hock 1975 (with 1974), Hock 1984 (with 1979, 1987), Kuiper 1991, see also Hock 1991, as well as note 20 below.

[7] Deshpande 1979 claims that the Sanskrit dental : retroflex contrast developed in post-Rig-Vedic. If correct, this would be another argument against prehistoric subversion. However, as noted in Hock 1979, the Rig-Veda offers evidence for a highly patterned, rule-governed DEGENERALIZATION of retroflex

sandhi across word boundary, an early phase of a change that gets virtually completed in the Classical period. We must therefore assume that retroflexion was introduced prehistorically, before the attested Rig-Vedic texts.

[8] Dravidian *ṟ* is occasionally written *ẓ*, but Krishnamurti 1969:318, n. 18 notes that there is no strong empirical evidence for this phonetic interpretation. Typologically, a system with a voiced obstruent not matched by a corresponding voiceless one is rare enough to require more than cursory justification. Note further that in the traditional Tamil alphabetical arrangement retroflex *ṟ* holds the same position relative to retroflex *ḷ* as alveolar *r* to alveolar *l*.

[9] See for instance Tikkanen 1988.

[10] The value of the tradition is weakened by the claim that the Brahuis came from Aleppo, in present-day Syria [!], but this element may reflect a later "Islamization" of an earlier tradition according to which the Brahuis are immigrants to the area.

[11] The extreme southeastern inscriptions from Yerragudi are most similar to those of north-central Kalsi. Since their language is clearly transplanted (the inscriptions are found deep in Dravidian territory), they cannot be relied on for dialectological judgments. The evidence of Gandhari Prakrit (with predominant dental) must likewise be ignored, since the language is transplanted, too. (The dental outcome might suggest an original affiliation closer with southwestern Girnar than with northwestern Gandhara, since the northwestern Aśokan inscriptions have predominant retroflex.)

[12] Note in this regard that unlike the south- and northwestern inscriptions, Kalsi does not substitute *r* for the *l* of the eastern inscriptions in words like *lājā/rājā* 'king'.

[13] As I became aware only after having examined (and reexamined) the Aśokan inscriptions, Turner did in fact connect the modern distribution to the Aśokan northwest : central : eastern distribution advocated here.

[14] Southworth's modern retroflex distribution is based on text frequencies. The highest retroflex : dental ratios are found in Sindhi, Gujarati, Marathi and the Dravidian south, the lowest ratios in Panjabi, Hindi, the Bihari languages, and Bangla, with the remaining areas having an intermediate ratio. Unfortunately Southworth does not indicate the texts on which his statistics are based.

Examining versions of the "Prodigal Son", representative of the different languages (and major subdialects) in Grierson 1903-1928, I arrive at rather different text frequencies and distributions: The highest retroflex : dental ratios (1 : 1-2.5) are found in a discontinuous Indo-Aryan area comprising Sindhi, Rajasthani, and Pahari dialects, and in Malayalam. Among the major languages, Kashmiri, Nepali, and Bangla have the lowest ratios (1 : 30 for Kashmiri, 1 : 9-15 for Nepali and Bangla). Intermediate ratios of 1 : 3-8 are found in most of South Asia, including the northwest and most of the Dravidian south. The northwest has

a number of pockets with significantly lower retroflex ratios (beside Kashmiri, note e.g. Burushaski with 1 : 33, Ormuri with 1 : 13.5, and Khowar with 1 : 58.5); there are similar pockets in the central area around Nahali (Gondi of Mandla with 1 : 13.5, Kurku with 1 : 34; see also Kuiper 1962: 255).

Whatever these geographical distributions may indicate about the history of South Asian languages, they do not support Southworth's grouping of Gujarati and Marathi with Sindhi and the entire Dravidian south, a grouping which is crucial for his claim that there was a strong prehistoric Dravidian presence in present-day western Indo-Aryan.

[15] See Subrahmanyam 1983, Zvelebil 1970. Northwestern Brahui does not offer any conclusive evidence for geminate alveolar *tt*; but as Emeneau 1971b observes, single alveolar *t* merges with retroflex *ṭ*, not with the dental; that is, its outcome is parallel to the dominant retroflex outcome of *r* + dental stop in northwestern Indo-Aryan.

[16] For most of Dravidian, only the geminate alveolar stop is considered; single alveolar stop generally changes to a liquid, commonly an [r]-sound. For Tamil, the conservative, literary retention of *tt* is assumed. (The dialectology of non-conservative, colloquial, and vernacular Tamil is quite complex and also, to my knowledge, not yet fully investigated. It appears that different varieties prefer dental or retroflex outcomes.) The geographically easternmost Dravidian languages have assibilated outcomes of geminate alveolar stops; these are not included in Map V.

[17] See for instance Hock 1991 and 1993a on early Indo-Aryan.

[18] Lengthening of the vowel preceding the cluster appears to depend in these early attestations on whether the *r*-sound was non-syllabic or syllabic. In the former case, *r* added a mora in the coda of the syllable so that its loss resulted in compensatory lengthening; in the latter case, there was no compensatory lengthening. (In later attestations, the loss of non-syllabic *r* in coda more commonly resulted in compensatory lengthening of the following consonant.)

[19] As I hope to show elsewhere, a similar north-to-south alignment between Indo-Aryan and Dravidian can be observed in the modern distribution of retroflex vs. dental (or rather, alveolar) nasals and laterals.

[20] See also note 6. More problematic is the question of "sponteneous retroflexion", as in RV *atati* 'wanders' vs. later *aṭati*. Developments of this type, too, have been attributed to subversion (e.g. Kuiper 1967a, Emeneau 1974); but alternative solutions have been proposed.

Some retroflexes have been explained by sporadic internal developments (e.g. dissimilation), others as borrowings from "Vedic Prakrits" or as anticipations of changes that become regular in Middle Indo-Aryan; cf. the discussion and references in Wackernagel 1889, Hoffmann 1941, Hock 1975, 1984, and 1991, Vine 1987.

If some of the developments should have been the result of contact, languages other than Dravidian might furnish alternative, or even more plausible sources: As argued in Hock 1984, since Dravidian has a contrast dental : retroflex, it is 'difficult to see how [the substitution of retroflex consonants for dentals] could be attributed to the mistakes made by Dravidians trying to speak a Sanskrit with undifferentiated dentals (cf. Emeneau 1974). Rather, just as in the case of modern-day contacts between Westerners and South Asians, I would expect speakers LACKING the contrast to make mistakes in trying to speak a language which has it. (Perhaps speakers of early forms of Munda, or of Tibeto-Burman, might be involved?)' In fact, if Indo-Aryan "dentals" had really been post-dental/alveolar, one would have expected the substitution of Dravidian alveolars; this is precisely what we find in Malayalam in nativizations of English words with alveolars.

Support for Tibeto-Burman provenience of some lexical items with "spontenous retroflexion" may be found in the fact that, except where it has undergone South Asian influence, Tibeto-Burman has undifferentiated alveolars which could be nativized either as dentals or as retroflexes in languages like Indo-Aryan which already had a contrast dental : retroflex (whether that contrast was due to subversion or not). That this is not just a thought experiment is suggested by the evidence in Witzel 1995 for river names ending in *-ta* or *-ṭa* (with apparent dental : retroflex variation) at the Himalayan border of Vedic Sanskrit, i.e., an area where a Tibeto-Burman presence is most likely. Interestingly, *kirāta*, the name of a non-Aryan people mentioned in the Rig-Veda and tentatively identified as Tibeto-Burman by Witzel, has a Pali variant with retroflex, *kirāṭa*.

Hock 1991 adds the further possibility that some "spontaneous" changes of dental to retroflex may have resulted from inner-Indo-Aryan differences suggested by the Prātiśākhyas (such that a 'tooth-root' *t* of one variety of Vedic could be reinterpreted as postdental and therefore retroflex in another variety whose *t* was interdental).

We should also consider the possibility that retroflexes replaced earlier dentals through sporadic analogical developments. In this regard note Hoffmann's 1941 observation that most Sanskrit words with retroflex *-ṇḍ-* belong to one of two semantic categories, that of "roundness" and that of "breaking, crushing". As Hoffmann correctly notes, this fact makes it possible that *-ṇḍ-* was secondarily extended to words belonging to one or the other of these categories.

[21] As I realized when rereading Kuiper 1967b for this paper, my 1987 account is similar, even though by no means identical, to that of Kuiper. I take this opportunity to add the reference to Kuiper's article to my 1987 paper.

[22] On the Vedic articulation of *r* see Hock 1992c.

[23] Kuiper's 1991 rejection of this explanation ignores the well-known independent evidence for the existence of Vedic Prakrits (for which see the discussion and references in Hock 1991). See also the evalution by Oberlies 1994.

[24] The Vājasaneyi- and Taittirīya-Prātiśākhyas specify svarabhakti only between *r* and sibilant (4.16 and 21.15-16, respectively). This is the context for which the Atharva- and Rik-Prātiśākhyas teach a fuller (half- or quarter-mora) version of svarabhakti, while before stop they recognize a shorter (1/8-mora) version. One suspects that the Vājasaneyi and Taittirīya-Prātiśākhyas do not describe a different 'dialect', but merely overlook the shorter, less noticeable variety of svarabhakti. (See further below for a restriction on svarabhakti even before sibilant, if followed by stop.)

[25] Earlier discussions of the Avestan situation that I am aware of do not consider this parallel with Old Indo-Aryan. Miller 1968 argues that structures of the type (8a), with *r* + *ə* + voiceless dental stop, are morphological renewals which, by undoing the conditioned development in (8b), reassert the morphological transparency of synchronically analysable forms. Kellens 1989 and Beekes 1988 ignore Miller's position and reassert the old claim that the difference between (8a) and (8b) is accentually conditioned. After a careful survey of all relevant Gatha Avestan forms, Beekes is able to maintain this view only by claiming that Iranian accentuation differed in a number of forms from the one of Vedic, even though the conditions for accent shift are no longer discernible. (He further claims that the change to *ṣ̌* was post-Gathic, but I do not find his arguments convincing.) None of these approaches offer a satisfactory account for the fluctuation in (8d).

[26] The view that the sound designated by the symbol «š» is some kind of retracted sound is supported by the fact that it reflects an earlier cluster involving POST-DENTAL, alveolar *r* and that in early Avestan it is distinguished from two other *š*-sounds of different, non-alveolar origins: a plain «š» (reflecting ordinary PIIr. *š*) and a palatal «ṧ» (resulting from palatal *c* + *y*). (See Hoffmann 1971.)

[27] Could the alveolar articulation of retroflex stops in many varieties of modern Hindi-Urdu be due to the influence of the Muslim conquerors, at least some of whom came from the northwest? Other, apparently less urban (or urbane) varieties have strong retroflex articulation, a fact which suggests that alveolar pronunciation is an urban overlay.

[28] See also p. 175, note 5.

[29] The fact that in (11a) the triggers are not lost may be cause for concern. But their presence can be accounted for as the result of analogical restoration of the root-final sonorants; in synchronically opaque structures such as (12) and (13), non-nasal triggers regularly are lost. The difference between geminate and non-geminate retroflex and alveolar stops in (12) and (13) can be explained with Krishnamurti (In Press) as reflecting the difference between geminate and simple dental stop in the input. In example (12b), taken from Zvelebil, the inputs therefore should be rewritten as **kal-ttu* and **uruḷ-ttu*.

[30] The (retroflex) sibilant dissimilation in (5b) is an idiosyncracy of Indo-Aryan and is therefore not included in the comparison.

[31] Tikkanen adds Yidgha, for which Skjærvø does not list the variant with retroflex stop.

[32] There is also a wide-spread and robust distribution of *ṣ* resulting from earlier *s* + *r* or *r* + *s* in the northwest, found in Nuristani, Middle and Modern East Iranian, and even northwestern Middle and Modern Indo-Aryan; see Tikkanen 1987:287-289, Morgenstierne 1947:234-235, von Hinüber 1986:28-29. The antiquity of the phenomenon is not clear; in Indo-Aryan, the input *s* may reflect earlier palatal **ś*. Since **ś* > *s* is a Middle Indo-Aryan innovation, it appears that the change of *s* + *r* or *r* + *s* may be a relatively recent phenomenon.

[33] For another source of *ṣ* see the preceding note.

[34] Zvelebil 1970:177, however, notes in passing the correspondence Tam. *aṇmai* 'nearness, being near' : *aṉpu* 'love', which is obviously related to the first item in (16) below, but he does not discuss its significance. A further exception might be Levitt 1989; but I find the arguments of the paper difficult to penetrate.

[35] The DEDR puts a question mark next to the Toda word.

References

DEDR = Burrow & Emeneau 1984.

ABBI, Anvita. 1992. Reduplication in South Asian languages: An areal, typological, and historical study. New Delhi: Allied Publishers.

BEEKES, Robert S. P. 1988. A grammar of Gatha-Avestan. Leiden: Brill.

BHAT, D. N. S. 1971. The Koraga language. Poona: Deccan College, Postgraduate and Research Institute.

——. 1973. Retroflexion as an areal feature. Stanford University Working Papers on Language Universals 13: 27-67.

BLOCH, Jules. 1911. Review of Denys de S. Bray, The Brahui language, 1. Journal asiatique, 10e série, 17: 1: 162-167. (Repr. in Bloch 1985.)

——. 1925. Sanskrit et dravidien. Bulletin de la Société de Linguistique de Paris 25: 1-21. (Repr. in Bloch 1985.)

——. 1929. Some problems of Indo-Aryan philology. Bulletin of the School of Oriental Studies (London) 5: 719-756. (Repr. in Bloch 1985.)

——. 1946. Brahui et tsigane. Journal of the Royal Asiatic Society 1946: 199-201. (Repr. in Bloch 1985.)

——. 1950. Les inscriptions d'Asoka. Paris: Les Belles Lettres.

——. 1985. Recueil d'articles de Jules Bloch (1906-1955), ed. by C. Caillat. Paris: Collège de France, Institut de Civilisation Indienne.

BÜHLER, George. 1864. On the origin of the Sanskrit linguals. Madras Journal of Literature and Science, No. I, Third Series, 116-186. [I have not been able to see this article; reference in Wackernagel 1896: 165 and elsewhere.]

BURROW, Thomas, & Murray B. Emeneau. 1984. A Dravidian etymological dictionary. 2nd ed. Oxford: Oxford University Press.

CALDWELL, Robert. 1856. A comparative grammar of the Dravidian or South Indian family of languages. First edition. London. (Third edition repr. 1974, New Delhi: Oriental Books.)

CHATTERJI, Suniti Kumar. 1926. The origin and development of the Bengali language. Calcutta: University Press. (Repr. 1970, London: Allen & Unwin. Also distributed by Motilal Banarsidass, Delhi.)

COMRIE, Bernard. 1981. The languages of the Soviet Union. New York: Cambridge University Press.

DESHPANDE, Madhav. 1979. Genesis of Ṛgvedic retroflexion: A historical and sociolinguistic investigation. Aryan and non-Aryan in India, ed. by M. Deshpande and P. E. Hook, 235-315. Ann Arbor: Center for South and Southeast Asian Studies, University of Michigan.

——. 1993. Aryans, non-Aryans, and brāhmaṇas: Processes of indigenization. Journal of Indo-European Studies 21: 215-235.

ELIZARENKOVA, Tatyana J. 1995. Language and style of the Vedic ṛsis. Albany: State University of New York Press.

EMENEAU, Murray B. 1954. Linguistic prehistory of India. Proceedings of the American Philosophical Society 98: 282-292. (Repr. in Emeneau 1980.)

——. 1956. India as a linguistic area. Language 32: 3-16. (Repr. in Emeneau 1980.)

——. 1969. Onomatopoetics in the Indian linguistic area. Language 45: 274-299. (Repr. in Emeneau 1980.)

——. 1971a. Dravidian and Indo-Aryan: The Indian linguistic area. Symposium on Dravidian civilization, ed. by Andrée F. Sjoberg, 33-68. Austin & New York: Jenkins Publishing & Pemberton Press. Repr. in Emeneau 1980.

——. 1971b. Kodagu and Brahui developments of Proto-Dravidian *ṟ. Indo-Iranian Journal 13: 176-198. (Repr. in Emeneau 1994.)

——. 1974. The Indian linguistic area revisited. International Journal of Dravidian Linguistics 3: 92-134. (Repr. in Emeneau 1980.)

——. 1980. Language and linguistic area: essays selected by A. S. Dil. Stanford, CA: University Press.

——. 1989. The language of the Nilgiris. Blue Mountains: The ethnography and biogeography of a South Indian region, 133-143. (Repr. in Emeneau 1994.)

——. 1994. Dravidian studies: Selected papers, ed. by Bh. Krishnamurti. Delhi: Motilal Banarsidass.

EMMERICK, Ronald E. 1989. Khotanese and Tumshuqese. Schmitt 1989: 204-229.

GELDNER, Karl Friedrich. 1951. Der Rig-Veda ..., 3 vols, posthumously edited by Charles R. Lanman. Cambridge, MA: Harvard University Press.

GIL, David. 1994. Conjunctive operators in South-Asian languages. SALA XV: Papers from the Fifteenth South Asian Language Analysis Roundtable Conference 1993, ed. by A. Davison and F. M. Smith. Iowa City: South Asian Studies Program, University of Iowa.

GONDA, Jan. 1975. Vedic literature. (A history of Indian literature, 1:1.) Wiesbaden: Harrassowitz.

GRIERSON, George. 1903-1928. Linguistic survey of India, 11 vols. Calcutta. Repr. 1968, Delhi: Motilal Banarsidass.

GUMPERZ, John J., & Robert Wilson. 1971. Convergence and creolization: a case from the Indo-Aryan/Dravidian border. Pidginization and creolization of language, ed. by D. Hymes, 151-168. Cambridge: University Press.

HAHN, Ferd. 1911. Kurukh grammar. Calcutta: Bengal Secretariat. Repr. (no date), Delhi: Mittal.

HALE, Wash Edward. 1986. Ásura in early Vedic religion. Delhi: Motilal Banarsidass.

HINÜBER, Oskar von. 1986. Das ältere Mittelindisch im Überblick. (Österreichische Akademie der Wissenschaften, Phil.-hist. Klasse, Sitzungsberichte 467.)

HOCK, Hans Henrich. 1974. Historical change and synchronic structure: The case of the Sanskrit root nouns. Toward tomorrow's linguistics, ed. by R. W. Shuy & C.-J. N. Bailey, 329-42. Georgetown University Press. (Repr. in International Journal of Dravidian Linguistics 4: 215-28 (1975).)

——. 1975. Substratum influence on (Rig-Vedic) Sanskrit? Studies in the Linguistic Sciences 5: 2: 76-125.

——. 1979. Retroflexion rules in Sanskrit. South Asian Languages Analysis 1: 47-62.

——. 1982a. AUX-cliticization as a motivation for word order change. Studies in the Linguistic Sciences 12:1: 91-101.

——. 1982b. The Sanskrit quotative: A historical and comparative study. Studies in the Linguistic Sciences 12: 2: 39-85.

——. 1984. (Pre-)Rig-Vedic convergence of Indo-Aryan with Dravidian? Another look at the evidence. Studies in the Linguistic Sciences 14: 1: 89-107.

——. 1986/1991. Principles of historical linguistics. Berlin: Mouton de Gruyter. (Second revised edition, 1991, ibid.)

——. 1987. Regular contact dissimilation. A Festschrift for Henry Hoenigswald, ed. by G. Cardona & N. Zide, 143-53. Tübingen: Narr.

——. 1988a. Historical implications of a dialectological approach to convergence. Historical dialectology, ed. by J. Fisiak, 283-328. Berlin: Mouton de Gruyter.

——. 1988b. Review article: Finiteness in Dravidian. (Sanford B. Steever (1988): The serial verb formation in the Dravidian languages.) Studies in the Linguistic Sciences 18: 2: 211-31

——. 1991. Dialects, diglossia, and diachronic phonology in early Indo-Aryan. Studies in the historical phonology of Asian languages, ed. by W. G. Boltz and M. C. Shapiro. Amsterdam & Philadelphia: Benjamins.

——. 1992a. Reconstruction and syntactic typology: A plea for a different approach. Explanation in historical linguistics, ed. by G. W. Davis & G. K. Iverson, 105-121. Amsterdam & Philadelphia: Benjamins.

——. 1992b. Review of Tikkanen 1987. Kratylos 37: 62-68.

——. 1992c. Were *r̥* and *l̥* velar in early Sanskrit? Vidyā-Vratin: Professor A. M. Ghatage felicitation volume, ed. by V. N. Jha, 69-94. (Sri Garib Dass Oriental Series, 160.) Delhi: Sri Satguru Publications.

——. 1993a. A critical examination of some early Sanskrit passages alleged to indicate dialectal diversity. Comparative-historical linguistics: Indo-European and Finno-Ugric: Papers in honor of Oswald Szemerényi III, ed. by B. Brogyanyi & R. Lipp, 217-232. Amsterdam & Philadelphia: Benjamins.

——. 1993b. Review of Abbi 1992. Studies in the Linguistic Sciences 23: 1: 169-192.

——. 1995. Subversion or Convergence? The issue of pre-Vedic retroflexion re-examined. Symposium on Language and Prehistory in South Asia, University of Hawaii, March 1995.

——. 1996. Pre-Ṛgvedic convergence between Indo-Aryan (Sanskrit) and Dravidian? A survey of the issues and controversies. Ideology and status of Sanskrit: Contributions to the history of the Sanskrit language, ed. by J. E. M. Houben, 17-58. Leiden: Brill.

HOFFMANN, Karl. [1941]. Die alt-indoarischen Wörter mit *-ṇḍ-*, besonders im Ṛgveda. Unpublished doctoral dissertation, München.

——. 1956. 'Wiederholende' Onomatopoetika im Altindischen. Indogermanische Forschungen 60: 254-264. (Repr. in Hoffmann 1975.)

——. 1958/67. Altiranisch. Handbuch der Orientalistik 1: 4: 1: 1-19. Leiden: Brill. (Repr. in Hoffmann 1975:58-76.)

——. 1971. Zum Zeicheninventar der Avesta-Schrift. Festgabe deutscher Iranisten zur 2500Jahrfeier Irans, 64-73. (Repr. in Hoffmann 1975:316-326.)

——. 1975. Aufsätze zur Indoiranistik, vol. 1, ed. by J. Narten. Wiesbaden: Reichert.

HOWELL, Robert B. 1991. Modern evidence for ancient sound changes: Old English breaking and Old High German vowel epenthesis revisited. Stæfcræft: Studies in Germanic linguistics, ed. by E. H. Antonsen & H. H. Hock, 65-73. Amsterdam: Benjamins.

JAKOBSON, Roman. 1931. Über die phonologischen Sprachbünde. Travaux du Cercle Linguistique de Prague 4: 234-240.

KELLENS, Jean. 1989. Avestique. Schmitt 1989: 32-55.

KONOW, Sten. 1906. (Discussion in vol. 4: 279 of Grierson 1903-1928.)

KRISHNAMURTI, Bh. 1961. Telugu verbal bases: A comparative and descriptive study. Berkeley & Los Angeles: University of California Press. [In spite of repeated attempts I have not been able to secure a copy of this book. My citations are based on references in other works.]

——. 1969. Comparative Dravidian studies. Current trends in linguistics, ed. by T. A. Sebeok, 5: 309-33. The Hague: Mouton.

——. 1991. The emergence of the syllable types of stems (C)VCC(V) and (C)V̄C(V) in Indo-Aryan and Dravidian: Conspiracy or convergence?. Studies in the historical phonology of Asian languages, ed. by W. G. Boltz and M. C. Shapiro, 160-175. Amsterdam & Philadelphia: Benjamins.

——. 1995. Patterns of sound change in Dravidian. Symposium on Language and Prehistory in South Asia, University of Hawaii, March 1995.

——. In Press. The origin and evolution of primary derivative suffixes in Dravidian. Historical, Indo-European, and lexicographical studies for Ladislav Zgusta in honor of his 70th birthday, ed. by H. H. Hock. Berlin: Mouton de Gruyter.

KUIPER, F. B. J. 1962. Nahali: A comparative study. (Mededelingen der Koninklijke Nederlandse Akademie van Wetenschappen, Afd. Letterkunde, N.R., 25: 5.) Amsterdam: Noord-Hollandsche Uitgevers Maatschappij.

——. 1966. The sources of Nahali vocabulary. Studies in comparative Austroasiatic linguistics, ed. by N. H. Zide. The Hague: Mouton.

——. 1967a. The genesis of a linguistic area. Indo-Iranian Journal 10: 81-102. (Repr. in International Journal of Dravidian Linguistics 3: 135-153 (1974).)

——. 1967b. The Sanskrit nom. sing. *viṭ*. Indo-Iranian Journal 10: 103-125.

——. 1991. Aryans in the Rigveda. (Leiden Studies in Indo-European, 1.) Amsterdam and Atlanta, GA: Rodopi.

——. 1992. Rigvedic loanwords. International Journal of Dravidian Linguistics 21: 1-49.

KULKE, Hermann, and Dietmar Rothermund. 1986. A history of India. Repr. 1990, New York: Dorsett.

LEVITT, Stephan Hillyer. 1989. The alternation of *r* and *l* in Dravidian. Indian Linguistics 50: 130-147. (Published in 1991.)

MACDONNELL, Arthur Anthony, & Arthur Berriedale Keith. 1912. Vedic index of names and subjects, 2 vols. Repr. 1958, Delhi: Motilal Banarsidass.

MANSION, Joseph. 1931. Esquisse d'une histoire de la langue sanscrite. Paris: Geuthner.

MARLOW, Patrick. Forthcoming. Origin and development of the Modern Indo-Aryan quotative and complementizer: an areal approach. University of Illinois Ph.D. dissertation in Linguistics.

MAYRHOFER, Manfred. 1986-. Etymologisches Wörterbuch des Altindoarischen. Heidelberg: Winter.

MILLER, D. Gary. 1968. *rt*-clusters in Avestan. Language 44: 274-283.

MORGENSTIERNE, Georg. 1947. Metathesis of liquids in Dardic. (Det norske videnskapsakademi i Oslo, Skrifter, 1947: 2.) (Repr. in Georg Morgenstierne 1975b.)

——. 1975a. Die Stellung der Kafirsprachen. Morgenstierne 1975b: 327-343.

——. 1975b. Irano-Dardica. Wiesbaden: Reichert.

OBERLIES, Thomas. 1994. Review of Kuiper 1991. Indo-Iranian Journal 37: 333-349.

PANDHARIPANDE, Rajeshwari. 1982. Counteracting forces in language change: convergence vs. maintenance. Studies in the Linguistic Sciences 12: 2: 97-116.

PARPOLA, Asko. 1988. The coming of the Aryans to Iran and India and the cultural and ethnic identity of the dāsas. International Journal of Dravidian Linguistics 17: 2: 85-229.

PAYNE, John. 1989. Pāmir languages. Schmitt 1989: 417-444.

POTT, August Friedrich. 1933, 1836. Etymologische Forschungen auf dem Gebiete der indogermanischen Sprachen, 1 and 2. Lemgo: Meyer.

RAMANUJAN, A. K., & Colin Masica. 1969. Toward a phonological typology of the Indian linguistic area. Current trends in linguistics, ed. by T. A. Sebeok, 5: 543-577. The Hague: Mouton.

RAU, Wilhelm. 1957. Staat und Gesellschaft im alten Indien nach den Brāhmaṇa-Texten dargestellt. Wiesbaden: Harrassowitz.

SCHMITT, Rüdiger (ed.). 1989. Compendium linguarum iranicarum. Wiesbaden: Reichert.

SKJÆRVØ, Prods O. 1989 (a). Modern East Iranian languages. Schmitt 1989: 370-383.

——. 1989 (b). Pashto. Schmitt 1989: 384-416.

SOUTHWORTH, Franklin C. 1974. Linguistic stratigraphy of North India. International Journal of Dravidian Linguistics 3: 201-223.

STEBLIN-KAMENSKIJ, M. I. 1965. Om alveolarer og kakuminaler i norsk og svensk. Norsk Tidskrift for Sprogvidenskap 20: 18-27.

STEEVER, Sanford B. 1988. The serial verb formation in the Dravidian languages. (MLBD Series in Linguistics, IV.) Delhi: Motilal Banarsidass.

SUBRAHMANYAM, P. S. 1983. Dravidian comparative phonology. Annamalainagar: Annamalai University.

THOMASON, Sarah Grey, & Terrence Kaufman. 1988. Language contact, creolization, and genetic linguistics. Berkeley & Los Angeles: University of California Press.

TIKKANEN, Bertil. 1987. The Sanskrit gerund: A synchronic, diachronic, and typological analysis. (Studia Orientalia, 62.) Helsinki: Finnish Oriental Society.

——. 1988. On Burushaski and other ancient substrata in northwestern South Asia. Studia Orientalia 64: 303-325.

TURNER, R. L. 1921. Gujarati phonology. Journal of the Royal Asiatic Society. Repr. in Turner 1973 = 1985: 88-145.

——. 1924. Cerebralization in Sindhi. Journal of the Royal Asiatic Society. Repr. in Turner 1973 = 1985: 206-227.

——. 1926. Middle Indian *-d-* and *-dd-*. Festgabe Hermann Jacobi. Repr. in Turner 1973 = 1985: 239-250.

——. 1973 = 1985. Indo-Aryan linguistics: Collected papers, ed. by J. Brough. Repr. 1985, Delhi: Disha Publications.

VINE, Brent. 1987. Vedic *kūláyati* and "spontaneous cerebralization" in Sanskrit. 14th International Congress of Linguists, Berlin, 1987. (Published in the Proceedings, ed. by W. Bahner, J. Schildt, and D. Viehweger, Berlin: Akademie-Verlag.)

WACKERNAGEL, Jakob. 1896. Altindische Grammatik, 1. Göttingen: Vandenhoeck & Ruprecht. Second edition, 1957, with a new introduction by L. Renou and Nachträge by A. Debrunner, ibid.

WITZEL, Michael. 1995. The linguistic situation in 1000 B.C. Symposium on Language and Prehistory in South Asia, University of Hawaii, March 1995.

ZIMMER, Heinrich. 1879. Altindisches Leben. Berlin: Weidmann.

ZVELEBIL, K. 1970. Comparative Dravidian phonology. The Hague: Mouton.

Studies in the Linguistic Sciences
Volume 23, Number 2, Fall 1993 [publ. October 1996]

REQUEST FORMATION IN ECUADORIAN QUICHUA

Joni Kay Hurley
Clemson University

This study analyzes request formation in Ecuadorian Quichua, focusing specifically on verb forms, morphological softeners, and lexical softeners. Perceptions of the degree of politeness conveyed by the various grammatical forms are presented, followed by a discussion of the influence of Spanish on Quichua. The findings will then be reviewed within the framework of language contact.

1. Introduction

The study of Quichua[1] has traditionally focused on the writing of grammar books and dictionaries (such as those of Catta Quelen 1987, Cordero 1989, Grimm 1896, Guzmán 1920, Leonardi 1966, París 1961, Stark and Carpenter 1973, and Vásquez 1990). Pragmatics, which studies language use to accomplish conversational goals such as requests, invitations, and offers, has apparently not been studied in Quichua. This paper is part of a larger research project to determine request formation in the Spanish and Quichua spoken in the Otavalo area of Ecuador, and how request formation in each of these languages may have been altered as a result of the language contact situation. The present study analyzes request formation in Quichua, the differing degrees of politeness conveyed by different grammatical structures, and possible Spanish influence.

2. Review of the literature: The study of requests

The study of requests originated with Austin 1962 and Searle 1976, 1979 and their Theory of Speech Acts. They were the first to relate grammatical form to the purpose of the utterance in the conversation. Originally they classified utterances into categories based upon the type of verb used. For example, *I request that you come* would be classified as a Directive in Searle's classification due to the presence of the verb *request*. The inherent defects in this classification system became apparent when Searle 1979 attempted to classify both direct and indirect speech acts. In direct speech acts there is a one-to-one correspondence between the syntactic form of a sentence and its illocutionary meaning. For example, if an utterance contains an imperative then it conveys a request. Indirect speech acts involve more than one possible interpretation. The syntactic form may convey one type of speech act, but the utterance is being used to execute a different speech act. Proper interpretation of the utterance is dependent upon background knowledge, the context in which the utterance is said, and the roles of the speaker and the hearer. The large number of utterances that Searle listed as forms of indirect speech acts caused him to label them as idiomatic since each conveys more than one meaning.

Language specific studies of request formation have been conducted on English (Searle 1976, Ervin-Tripp 1976, Wardhaugh 1985, Blum-Kulka, House, and Kasper 1989, Fraser and Nolen 1981), Portuguese (Koike 1986, 1989, Wherritt 1983), Athapaskan (Rushforth 1985), Tzeltal (Brown 1979), and Spanish (Blum-Kulka, House, and Kasper 1989, Fraser and Nolen 1981, Haverkate 1979, Wilson 1965). The results of these studies show that different grammatical formations result in requests that are perceived by speakers of these languages to be more, or less, polite. (The concept of politeness has been elaborated at length by Brown and Levinson 1978). Politeness in requests is determined by a combination of grammatical/lexical form and patterns of use in a given language/culture.

Blum-Kulka, House, and Kasper 1989 have conducted an ambitious study of requests and apologies across five languages (Spanish, German, Hebrew, English, and French). They found that all the languages in question have direct and indirect request strategies, but that each language may form these strategies with different grammatical forms. In addition, perceptions of politeness[2] and patterns of use for the different request strategies can vary from language to language.

3. Study site

The research for this study was conducted in Otavalo and neighboring small towns in Ecuador over a period of seven months from 1989 to 1992. Otavalo is located about two hours north of Quito in the Andes mountains in the province of Imbabura. Residents of Otavalo speak either Spanish or Quichua as their maternal language. Many native Spanish speakers have some knowledge of Quichua since it is heard daily in the marketplace, but they do not normally acquire the ability to speak it. All Indians speak Quichua as their native language and most are also bilingual in Spanish to some degree.[3] The only Indians who remain monolingual in Quichua are those who live in isolated areas.

4. The Quichua language

Quichua is a member of the Andean-Equatorial language family and may be closely related[4] to the Aymara language which is spoken in areas of Peru and Bolivia (Escobar 1986). Quichua (Quechua) is spoken by some seven million people from Ecuador to northern Argentina (Fromkin and Rodman 1988). Quichua does not have a standardized orthography and has been declared an official language only in Peru. This language is largely connected with Indian culture and is therefore highly stigmatized in Andean countries.

Quichua is an agglutinating language in which the accumulation of suffixes conveys grammatical relations that are expressed in Indo-European languages by syntactic means. The following example illustrates the use of suffixes in Quichua:

(1) Raimicunapica
Raimi + cuna + pi + ca
holiday / plural / in, on / topic marker
'On holidays'

It is sometimes difficult to determine the meaning of a Quichua suffix, especially when it serves a function unknown in Indo-European languages, such as the topic marker *-ca* and the validation suffix *-mi*. These suffixes require further study.

5. Methodology and sample

Both elicited data and naturally-occurring requests were tape recorded in the Otavalo area. A questionnaire to elicit role-play (based upon the model of Blum-Kulka, House, and Kasper 1989) was developed in Quichua with the help of a bilingual Indian. Various situations were devised in which requests would commonly occur, such as:

(2) Shuc mamaca, imashinata chunga ishcai huatata charic churita mañan, cuyman sarata jihuata carachun?

How does a mother ask her 12 year old son to give corn and grass to the guinea pigs?

(3) Imashna shuc tiuca, quilcana cashpita, paipac mashita mañan?

How does a man ask his friend for a pen?

Seventeen such situations were developed, and were mixed with elicitations of thirteen other speech acts (greetings, offers, expressions of gratitude, complaints, etc.) so that the informants would not realize that requests were being solicited.

Another bilingual Indian tape-recorded the interviews and answers to informant profile questions (age, occupation, education, language ability, etc.) with 75 Quichua-speaking males between the ages of 20 and 50.[5] Each interview lasted from 20 to 45 minutes.

Six and a half hours of natural conversations were recorded in a store in downtown Otavalo that sells ponchos, blouses, and blankets to the Indians. At least 56 different speakers were represented.

6. Corpus

A total of 1,873 requests were recorded in Quichua, of which 1,803 were produced in interviews and 69 occurred in natural conversations. The small number of requests in natural conversations is due to the extensive use of Spanish in commerce transactions (Hill and Hill 1986).

7. Data analysis

Native speakers of Quichua listened to the tapes and transcribed what was said. These data were then entered into LOTUS 123 spreadsheet computer programs which contained columns dedicated to the interview number, the number of the question, and the utterance. All the data were coded for grammatical and lexical forms that would have a bearing on the politeness of the requests.[6] This allowed the data to be sorted in a variety of ways to determine patterns. In a follow-up study, Quichua speakers ranked a series of requests (selected from the interviews) from least to most polite.[7]

8. Findings

This analysis of Quichua requests will be divided into verb forms (modal verbs, imperatives, and other verb forms) and morphological and lexical softeners (diminutives, politeness suffixes, courtesy expressions, interjections, and vocatives). This will be followed by a discussion of the perception of politeness in Quichua, and possible Spanish influence on the Quichua language.

8.1 Verb forms

8.1.1 Modal verbs— *carana* 'to give'

The only modal verb in this study is *carana* 'to give', which is used with the gerund to convey softened requests.[8] This structure translates as *do me the favor of* or *please* and can only be used with transitive verbs (i.e. verbs that permit a direct object) that do not clearly indicate benefit to the speaker.[9] As a result, *carana* as a softener is of low frequency, occurring in 11.2 percent (n=209) of the elicited data, and did not occur in the natural conversations due to the scarcity of transitive verbs. An example of this structure is:

(4) Papaguta randishpa carahuay.
Papa + gu + ta randi + shpa cara + hua + y
potato / dim.[10] / acc. / buy / -ing / give / me / imp.
'Do me the favor of buying (me) some potatoes.'

In a few cases (4 out of 209 occurrences) another verb meaning 'give', *cuna*, is used in the same way:

(5) Ashtahuan, ribajashpa cuhuay.
Ashtahuan ribaja + shpa cu + hua + y
too much / reduce / -ing / give / me / imp.
(That's) too much, do me the favor of reducing (the cost).

Both these verbs are used in Ecuadorian Quichua with the meaning of 'to give' and are used as softeners with the gerund *-shpa* (-ing) (Albor 1973, Catta Quelen 1987, Toscano Mateus 1953). Studies of the Cuzco dialect of Quechua describe the verbal suffix *-cu* as conveying cordiality and personal interest when used with imperatives (Gutiérrez 1990, Solá and Yupanqui 1970). However, the meaning added by *-cu* is not translated in their examples,[11] such as:

(6) Kapuliyta ranti**cu**huay.
Kapuli + y + ta ranti + **cu** + hua + y
cherry / my / acc. / buy / me / imp.
Buy from me my cherries.

As this example illustrates, *cuhuay* could easily have been separated from the verb *ranticuhuay* 'buy from me' and used as a separate verb, equivalent in form to 'give me', *cuhuay*. This usage has apparently been transferred to the other verb for 'to give', *carana*, in Ecuadorian Quichua.

8.1.2 Imperatives (Commands)

8.1.2.1 Present imperatives

Present imperatives, formed in Quichua by suffixing *-y* to the verb root as in *shamuy* 'come', were the most common means of conveying requests in Quichua. They occurred in 73.4 percent (n=1,325) of the elicited requests and in 52.2 percent (n=36) of the naturally-occurring requests. These commands are used when the addressee is expected to carryout the request immediately (Leonardi 1966). An example of such a command is:

(7) Mañachiy lapizguta escribingapac.
mañanchi + y lapiz + gu + ta escribi + ngapac
lend / imp. / pencil / dim. / acc. / write / in order to
Lend (me) a pencil in order to write.

8.1.2.2 Future imperatives

The future imperative (formed by suffixing *-ngui* to the verb root) was used in 10.6 percent (n=192) of the elicited data and in 29 percent (n=20) of the naturally-occurring data. The future imperative is used for commands that are to be executed at a time subsequent to *right now* (Leonardi 1966, Mugica no year).[12] This can be clearly seen in the data in which two commands occur together, as in:

(8) Shamuy, randipangui, caipi yapachishpa cusha.
Shamu + y randi + pa + nguicai + pi yapachi + shpa
come / imp. / buy / please / fut.imp. / this / in / to give one extra / -ing /
cu + sha
give / I will
Come (pres. imp.), please buy (fut. imp.) here, I will give (them to you) giving (you) one extra.

As this example illustrates, a present imperative command to come (*shamuy*) is followed by a command in the future imperative, *randipangui* 'please buy' conveying what is to be done after the addressee comes.

The future imperative is also used to convey politeness (Carpenter 1982). This was supported by research conducted by the author of this paper in 1990, in which the average ranking of various request structures by forty-eight Quichua Indians indicated that requests in the future imperative are considered to be more polite than requests in the present imperative.

8.1.3 Other verb forms

The remaining 15.4 percent of the elicited requests and 18.2 percent of the naturally-occurring requests were formulated with six other grammatical strategies. The frequencies of these strategies are presented in Table 1.

TABLE 1. FREQUENCIES OF OTHER VERBAL STRATEGIES

Grammatical Structure	Frequency of Occurrence in Elicited Data	Frequency of Occurrence in Natural Data
1. Questions[13]	6.8% (n=122)	8.7% (n=6)
2. Statements (not 'need')	2.3% (n=42)	5.8% (n=4)
3. *-shun* 'let's'	4.2% (n=76)	2.3% (n=2)
4. 'Need' statements	1.8% (n=33)	0%
5. *-chun* (subjunctive)[14]	.3% (n=5)	0%
6. Softener only	0%	1.4% (n=1)

An example of each of these request structures is provided below.

a. Questions

(9) Nachu chai puchata charingui?
Na + chu chai pucha + ta chari + ngui
neg. / quest. / that / yarn / acc. / have / 2nd pers.
Don't you have that yarn (that I am wanting to borrow)?

b. Statements other than *need*

(10) Pero chai preciupacca na ushani.
Pero chai preciu + pac + ca na usha + ni
but / that / price / for / topic / neg. can / 1st pers.
But for that price I can't (buy it, so reduce the price).

c. 'Let's' ...

(11) Jacu futbulta pucllashun.
Jacu futbul + ta puclla + shun
let's go / soccer / acc. / play / let's
Let's go, let's play soccer.

d. 'Need' statements

(12) Por Diosmanda, sacota ahuangapac, trabajangapac munani.
Por Dios + manda saco + ta ahua + ngapac trabaja + ngapac muna + ni
Please / by / sweater / acc. / weave / to / work / to / need/want / 1st pers.
Please, I need to work, (I need) to weave sweaters.

e. Subjunctive

(13) Ricungui, cunan charini shuc carruguta, cunan como can yachangui manejanaca, munani que can trabajachun ñuca carrupi.[15]
Ricu + ngui cunan chari + ni shuc carru + gu + ta cunan como can
look / fut.imp. / now / have / 1st pers. / a / car / dim. / acc. / now / since / you/
yacha + ngui maneja + na + ca muna + ni que can trabaja + chun
know / 2nd pers. / drive / inf. / topic / want / 1st pers. / that / you / work / subj./
ñuca carru + pi
my / car / in
Look, I have a car now, and since you know how to drive, I want you to work (for me) in my car.

f. Softener only

(14) Por Dios, cumari.
Por Dios, cumari
please / godmother
Please, godmother (sell it to me cheaper).

8.2 Morphological softeners

8.2.1 Diminutives

There is much variation in the diminutives used in the various dialects of Ecuadorian Quichua, and they include *-hua*, *-cu*, and *-lla* (Jara [no year]). Diminutives were used in 61 percent (n=1,101) of the elicited requests, and in 24.6 percent (n=17) of the naturally-occurring requests, and were found primarily on direct objects (55.6 percent [n=612] of the diminutives) and on vocatives (14.2 percent [n=156] of the diminutives). Three diminutives were used in this sample of Otavalo Quichua: *-cu* (and its voiced variant *-gu*), *-lla*, and the Spanish diminutive *-ito*.

The most common diminutive was *-cu*/*-gu*, representing 84.7 percent (n=932) of all diminutives in the elicited data, and 41.2 percent (n=7) of the diminutives in the naturally-occurring data. Examples include:

(15) esferu**cu**ta
esferu + cu + ta
pen / dim. / acc.
a pen (**dim**.)

(16) tanda**gu**
tanda + gu
bread / dim.
a **little** bread

There appears to be no difference in meaning or usage between *-cu* and *-gu*, and the voiced variant is not due to the phonological environment. Catta Quelen 1987 reports the distribution to vary according to geographic region. In the present study, usage could be closely linked to town of residence. For example, all informants from Peguche used only *-gu*, except on the word *taita* 'father, sir', which always contained the diminutive *-cu*. Informants residing in Otavalo used only *-cu*. The only informants who would use both diminutives had been born in one town and were now living in another. It is not clear why the variant *-cu* is the only one used on *taita* 'father, sir' regardless of the diminutive used in a particular informant's town of residence.

The diminutive *-lla* was much less common, representing 10 percent (n=110) of the diminutives used in the elicited data, and 41.2 percent (n=7) of the diminutives used in the naturally-occurring data. The use of this diminutive is largely restricted to specific lexical items, especially the Spanish loanwords *barato* 'cheap', *tío* 'uncle, sir' and *amo* 'master'. An example is:

(17) baratu**lla**
baratu + lla
cheap / dim.
a **little** cheap

The two diminutives *-cu/-gu* and *-lla* can be combined on the same word to strengthen the minimizing effect:

(18) ratu**gulla**
ratu + gu + lla
while / dim. / dim.
a **very little** while

Forty Quichua-speaking Indians ranked nouns with diminutives as more polite than nouns without diminutives, but there was little difference in perceived politeness between *-cu/-gu* and *-lla*.

The last diminutive found in this sample was the Spanish *-ito*, which represented 5.4 percent (n=59) of the diminutives used in elicited requests, and 17.6 percent (n=3) of the diminutives used in naturally-occurring requests. A careful analysis of the use of this diminutive reveals that it occurred only on Spanish loanwords, principally vocatives. For example:

(19) amigu**ito**
amigo + ito
dear friend

8.2.2 The suffixes *-pa, -lla,* and *-ya(ri)*

The suffix *-pa* is attached to verb forms to convey courtesy and respect on the part of the speaker toward the addressee (Quintero and Cotacachi 1986, Carpenter 1982) and is often loosely translated as 'please' (Jara [no year], Catta Quelen 1987). This suffix occurred in 25.4 percent (n=458) of the elicited requests, and in 11.6 percent (n=8) of the naturally-occurring requests. An example is:

(20) Ñucapac carrupi trabajangapac shamu**pay**.
Ñuca + pac carru + pi trabaja + ngapac shamu + pa + y
I / of / car / in / work / to / come / please / imp.
Please come to work in my car.

The suffixes *-lla* and *ya(ri)* are placed on imperatives to achieve opposite effects. The suffix *-lla* is used to soften verbs and is usually translated as 'just' (Mugica [no year], Quintero and Cotacachi 1986, Carpenter 1982, Stark and Carpenter 1973). It was used in only .1 percent (n=2) of the elicited requests, and in 10.1 percent (n=7) of the naturally-occurring requests. An example of this suffix is:

(21) Shamuy**llá**.[16]
Shamu + y + lla
come / imp. / just
Just come (on over).

It is possible that this is simply the diminutive suffix *-lla* placed on verb forms. Further support for this idea is found in the use of *nomás* 'just' in Andean Spanish, which serves as a softener for both nouns and imperatives (Naula Gaucho 1975, Stratford 1989). It is claimed that this use of *nomás* in Spanish is due to the influence of the Quichua suffix *-lla* (Quintero and Cotacachi 1986, Catta Quelen 1987). In many instances there were strong parallels in the data gathered in this study between the use of *-lla* in Quichua and *nomás* in Spanish, as is illustrated in the following examples:

(22) Compadrito, caiman shamuy**llá**.
Compadre + ito cai + man shamu + y + lla
godfather / dim. / this / to / come / imp. / just
Godfather, **just** come (on) over here.

(23) Ya, ya, bueno, bueno, vendrá **nomás**.
Okay, okay, fine, fine, **just** come (on) (fut. imp.).

When *-lla* is used in commands there is a distinctive intonational contour that is not found when it is placed on nouns. The pitch rises suddenly to a higher level on the suffix *-lla*, and then drops off rapidly.

The suffix *-ya(ri)*, also translated as 'just' or 'come on' (Gutiérrez 1990), is attached to commands to make them more emphatic (*Centro de Investigaciones para la Educación Indígena* 1983). The longer form, *-yari*, is considered to be more emphatic than the shorter form *-ya* (Catta Quelen 1987, Quintero and Cotacachi 1986). Gutiérrez 1990 reports that *-yari* can also serve to emphasize pleading, as in *come on* in English or *ya pues* in Spanish. The suffix *-yari* was used in 1.5 percent (n=26) of the elicited requests, and in 10.1 percent (n=7) of the naturally-occurring requests. For example:

(24) Cuatrupac cuhuay**yari**.
Cuatru + pac cu + hua + y + yari
four / for / give / me / imp. / just (emphatic)
Come on, give (it) to me for four (thousand sucres).

This suffix serves as the pattern for the use of *pues* 'just, come on' as a suffix in Andean Spanish, which emphasizes utterances (Quintero and Cotacachi 1986). As is true for the suffix *-lla* in Quichua, *pues* is placed after the word or words that the speaker wishes to emphasize:

(25) Venderáme[17] a 50 **pues**.
Sell (future imperative) (it) to me for **just** 50 (sucres).

In this example, *pues* is placed after *50* conveying that what is being emphasized is the reduced price of 50 sucres, compared to the asking price.

8.3 Lexical softeners

8.3.1 Courtesy expressions

The sample gathered in this study contains six expressions that convey 'please' and all are either borrowed directly from Spanish, translated from

Spanish, or a combination of both. These expressions were used in 24.8 percent (n=448) of the elicited requests, and in only 1.4 percent (n=1) of the naturally occurring requests. The lexical items and their frequencies are presented in Table 2.

TABLE 2. LEXICAL SOFTENERS AND THEIR FREQUENCIES

Courtesy Expression	Frequency of Occurrence in Elicited Data	Frequency of Occurrence in Natural Data
Por dios(manda) 'By God'	20.2% (n=364)	1.4% (n=1)
Ama shinagu cashpa 'Don't be that way'	3.0% (n=55)	0%
Favor 'Please'	.2% (n=4)	0%
Por favor 'Please'	.4% (n=7)	0%
Favorta rashpa 'Doing the favor'	.7% (n=12)	0%
Favorta shinashpa 'Doing the favor'	.3% (n=6)	0%

The first expression, *por Dios(manda)* 'by God', was used in Old Spanish and is presently the most frequent means of conveying 'please' in the Quichua spoken in Otavalo. In many cases the Quichua suffix for 'by' is added, forming *por Diosmanda* 'by God by'. *Ama shina cashpa* 'don't be that way',[18] is the translation of the Spanish politeness expression *no sea(s) malito* 'don't be bad' which has been documented in Ecuador, Bolivia, and Mexico. Both *favor* and *por favor* 'please' have been borrowed directly into Quichua, and the last two expressions listed in Table 2 are loanblends, combining the Spanish word *favor* with the Quichua verbs *(ru)rana* and *shinana* meaning 'to do'. The resulting expressions are equivalent to the Spanish *hága(me) el favor de...* 'do (me) the favor of...'.

8.3.2 Interjections and vocatives

Interjections were used in 6.4 percent (n=116) of the elicited requests, and in 4.3 percent (n=3) of the naturally occurring requests. Only two interjections, *jala* 'look' and *jaica* 'take', are purely Quichua. The others were either borrowed or translated from Spanish, such as: *bueno* 'well', *oye* 'listen', *uyay* (from the Spanish *oye*, 'listen'), and *ricuy* 'look'. The most common interjection was *jala* 'look', representing 59.5 percent (n=69) of the interjections in the elicited data, and 66.6 percent (n=2) of the interjections in the naturally occurring data.

The vocatives that occurred in the corpus were primarily terms of family relationship or friendship, and 75 percent were of Spanish origin. They were used in 26.2 percent (n=472) of the elicited requests, and in 20.3 percent (n=14) of the naturally-occurring requests. Vocatives in this sample included: *compadre* 'godfather', *comadre* 'godmother', *pana* 'sister', *guambra* 'guy', *taita* 'father/sir', *amigo* 'friend', and *señor* 'sir'.

8.4 Request strategies and perceptions of politeness

Forty-eight Quichua-speaking Indians were asked to rank a series of requests to *come* in order of politeness (without considering factors of the situation). Their ranking is presented in Table 3. The resulting scale demonstrates that: the future

imperative is more polite than the present imperative, the future imperative is more polite than the suffix *-pa* 'please', *por Diosmanda* 'by God' is the strongest courtesy expression, and the most polite request formation consists of the future imperative, *-pa*, and *por Diosmanda* 'by God'. Gutiérrez 1990 reports that the first example, *shamuy* 'come', is a command, and that the most polite request, *Por Diosmanda, shamupangui* 'By God, please come [future imperative]', conveys almost pleading.

TABLE 3. RANKING OF QUICHUA REQUEST FORMS IN ORDER OF POLITENESS (least polite to most polite)

Shamuy.	Come (pres. imp).
Shamuyllá.	Just come (pres. imp.).
Shamupay.	Please come (pres.imp.).
Shamungui.	Come (fut. imp).
Shamupangui.	Please, come (fut. imp.).
Por Diosmanda, shamuy.	Please, come (pres. imp.).
Por Diosmanda, shamupay.	Please, please come (pres. imp.).
Por Diosmanda, shamungui.	Please, come (fut. imp.).
Por Diosmanda, shamupangui.	Please, please come (fut. imp.).

A similar ranking of requests to buy potatoes demonstrates that *carana* 'to give' makes requests more polite and that *por favor* is the least polite expression for 'please', with *por Diosmanda* being considered the most polite. This ranking is presented in Table 4.

TABLE 4. POLITENESS RANKING INVOLVING THE VERB *CARANA* 'TO GIVE' (least polite to most polite)

Papaguta randihuay.
Buy me potatoes (dim.).

Papaguta randishpa caray.
Do (me) the favor of buying (me) potatoes (dim.).

Papaguta randishpa carahuay.
Do me the favor of buying (me) potatoes (dim.).

Por favor, carahuay randishpa papaguta.
Please, do me the favor of buying (me) potatoes (dim.).

Ama shinagu cashpa, papaguta randishpa carahuay.
Don't be that way, do me the favor of buying (me) potatoes (dim.).

Por Diosmanda, papaguta randishpa carahuay.
By God, do me the favor of buying (me) potatoes (dim.).

Finally, forty Quichua Indians ranked vocative terms with diminutives as more polite than vocatives without diminutives, but there was very little difference in politeness noted between *-lla* and *-cu*.

8.5 The influence of Spanish on Quichua

In the elicited data the most obvious influence of Spanish on Quichua is the large percentage of loanwords, which represent from 7 to 49 percent of the vocabulary used, depending upon the informant. However, a careful analysis of the words themselves reveals that the 2,828 occurrences of Spanish words represent only 300 different words. The majority refer to items and concepts brought by the Spanish such as *carro* 'car', *trabajar* 'to work for money', *llamada* 'telephone call', and *bautizar* 'to baptize'. Also frequently borrowed are connecting words and phrases such as *pero* 'but', *y* 'and', *o sea que* 'or rather', and *entonces* 'then' which often replace the Quichua suffixes with similar meanings. In addition, the only lexical items (as opposed to suffixes) that are used to soften requests in these data are apparently all of Spanish origin.

Spanish influence is also seen in the borrowing of one, and possibly two, suffixes into Quichua. The Spanish suffix *-dor* 'the person who' is occasionally used with this meaning on Quichua words, replacing the Quichua equivalent *-c*. Examples include *pucllador* 'ball player' and *ahuador* 'weaver'. It is also possible that the Quichua diminutive suffix *-cu/-gu* is from the Spanish diminutive *-ico*, although this has not been documented in any of the Quichua grammars. Support for this idea is found in Bolivian varieties of Quechua, which use the Spanish diminutive *-ito* as the primary diminutive, as in *jamp'atitu* 'little toad' (Urioste 1955:21). In the Otavalo area, *-ito* is used only on Spanish loanwords.

In naturally-occurring conversations the most obvious influence of Spanish on Quichua is language mixing: either code-switching between Quichua and Spanish, or *media lengua* 'middle language' - Quichua syntax with approximately 90 percent Spanish vocabulary (Muysken 1981). *Media lengua* has been reported by Muysken 1981 in the southern dialects of Ecuadorian Quichua. This mixed language is described as a combination of the Quichua grammatical system with the majority of the lexicon of Spanish origin. *Media lengua* occurs to a limited extent in the Otavalo area. In the example provided below, the vocabulary of Spanish origin is written in capital letters, and Quichua words and suffixes are written in lower case letters:

(28) CUCINA URA ISQUINAcupi; VINTANAcuna, SILLAcuna, tianmi.
CUCINA URA ISQUINA + cu + pi; VINTANA + cuna
kitchen / now / corner / dim. / in / window / pl.
SILLA + cuna tia + n + mi
chair / pl. / exist / 3rd pers. / validator
The kitchen now (is) in the corner (diminutive); there are windows and chairs.

In this example, all the vocabulary except *tian* 'there are' is of Spanish origin. Word order and grammatical relations are completely Quichua.

9 Conclusions

The grammatical strategies used in the formulation of requests in the Quichua data recorded for this study are summarized in Table 5. The elicited data is characterized by more extensive use of softeners than is the case in the naturally-

occurring conversations. This is most likely due to the fact that the interview situation is more formal and there is no true relationship between the people in the hypothetical situations. In addition, the natural conversations dealt with commerce, in which the banter between the customer and the vendor is relatively direct and to the point. In both types of data the primary verb form is the imperative, both present (*-y*) and future (*-ngui*). Elicited speech contained many Spanish loanwords (ranging from 7 to 50 percent of the vocabulary used by each individual), whereas the naturally-occurring conversations were characterized not only by many loanwords but by code-switches to Spanish as well.

TABLE 5. SUMMARY OF QUICHUA REQUEST STRATEGIES

Strategy	Frequency of Occurrence in Elicited Data	Frequency of Occurrence in Natural Conversations
Modal verbs:		
Carana 'to give'	11.2% (n=209)	0%
Verb forms:		
Present imperative	73.4% (n=1,325)	52.2% (n=36)
Future imperative	10.6% (n=192)	29.0% (n=20)
Questions	6.8% (n=122)	8.7% (n=6)
-shun 'let's'	4.2% (n=76)	2.3% (n=2)
Statements (not 'need')	2.3% (n=42)	5.8% (n=4)
'Need' statements	1.8% (n=33)	0%
Softener only	0%	1.4% (n=1)
-chun (subjunctive)	.3% (n=5)	0%
Morphological softeners:		
-cu/-gu	51.7% (n=932)	10.1% (n=7)
-lla (added to nouns)	6.1% (n=110)	10.1% (n=7)
-ito (from Spanish)	3.3% (n=59)	4.3% (n=3)
-pa (polite)	25.4% (n=458)	11.6% (n=8)
-lla (added to verbs)	.1% (n=2)	10.1% (n=7)
Lexical softeners:		
Courtesy Expressions:		
Por Diosmanda 'By God'	20.2% (n=364)	1.4% (n=1)
Ama shinagu cashpa 'Don't be that way'	3.0% (n=55)	
Favor 'please'	.2% (n=4)	
Por favor 'please'	.4% (n=7)	
Favorta rashpa 'Doing the favor'	.7% (n=12)	
Favorta shinashpa 'Doing the favor'	.3% (n=6)	
Interjections	6.4% (n=116)	4.3% (n=3)
Vocatives	26.2% (n=472)	20.3% (n=14)

When the data collected in Quichua is compared to that collected for Spanish in the Otavalo area (Hurley 1992), important observations can be made concerning language contact and request formation. Both data sets support the idea that indirect request strategies in Quichua and Spanish have been greatly reduced (as compared to other varieties of these languages) and replaced with a higher frequency of direct strategies that translate easily from one language to another, specifically present and future imperatives softened with lexical expressions and diminutives. This shared pragmatic system can be clearly seen in Table 6.

Table 6. REQUEST STRATEGIES IN THE SPANISH AND QUICHUA OF OTAVALO, ECUADOR

Grammatical Form	Frequency in Spanish	Frequency in Quichua
¿Puede..? 'Can you..?'	2.2% (n=63)	0%
Present imperative	61.9% (n=1,807)	72.7% (n=1,361)
Future Imperative	7.7% (n=226)	11.3% (n=212)
'Give' as a softener	10.2% (n=298)	11.2% (n=209
Diminutives	36.8% (n=1,133)	59.7% (n=1,118)
Lexical softeners	23.2% (n=714)	24.0% (n=449)
Interjections	20.1% (n=618)	6.4% (n=119)
Vocatives	28.4% (n=875)	25.9% (n=486)

To reach this point of shared pragmatics, the following changes have apparently occurred in Otavalo Quichua: a decreased reliance on the use of politeness suffixes and the borrowing of Spanish lexical courtesy expressions, word order changes so that softening suffixes in Quichua (such as *-lla*) occupy the same syntactic slot as their Spanish equivalents, the adoption of at least one (*-ito*) and possibly all three Spanish diminutives (*-lla* [from the Spanish *-illo*?] and *-cu/-gu* [from the Spanish *-ico*?]), and a preference for direct verbal request strategies (i.e. imperatives). Changes in Otavalo Spanish (as compared to the findings of Blum-Kulka 1989) include: the virtual abandonment of *poder* 'to be able' as a request softener, an increased use of imperatives, the adoption of both a present and a future imperative, the use of the future imperative to signal compliance at a future time or increased politeness, the loan translation of the Quichua modal verb *carana* 'to give', and the use of *nomás* 'just' and *pues* (emphatic) to express shades of meaning that are conveyed in Quichua through suffixes.

The solution to cross-cultural communication problems in a language contact situation lies in the development of a shared set of pragmatic strategies. In order to reach this point, strategies that are used by both linguistic groups are used with greater frequency (such as imperatives in Quichua and Spanish), and those which are not shared are either borrowed (such as the borrowing of *give* as a modal verb into Spanish and of lexical courtesy expressions into Quichua) or discarded (such as the modal verb *poder* 'to be able' in Spanish). Since the shared pragmatic system used in the Quichua and Spanish spoken in Otavalo is based upon direct request strategies (present and future imperatives plus softeners) which can be used in all conversational situations, there is very little possibility of being misunderstood.

NOTES

[1] Quichua is referred to as Quechua outside of Ecuador.

[2] There have been many studies on the concept of politeness, beginning with Goffman 1967 and Brown and Levinson 1978. They determined the basic principles of politeness, which are considered universal, such as 'saving face'. However, each language and culture possesses a variety of linguistic forms that are considered more, or less, polite by members of that culture. There may be some overlap between cultures in the grammatical structures and the relative degree of politeness they are perceived to convey. Many language-specific studies have been conducted on request formation and perceived politeness (such as Searle 1976, Fraser and Nolen 1981, Koike 1986, 1989 and Brown 1979). In order to determine which grammatical structures are considered more polite by speakers of a given language, they are commonly asked to rank a series of requests that vary in grammatical/lexical choice from least to most polite (such as Fraser and Nolen 1981 and Koike 1986, 1989). While this method does not associate grammatical structure with actual patterns of use, it does shed light upon what linguistic features convey increased politeness.

[3] Bilingualism is regarded as a continuum, ranging from knowing a few words and phrases in the second language to being a fluent speaker of two languages.

[4] There is disagreement among linguists as to whether or not Quichua/Quechua and Aymara are genetically related.

[5] Only males were used due to the limited amount of time spent in Ecuador and the desire to have as homogeneous a group as possible.

[6] The determination of what grammatical/lexical categories are important in request formation was based upon the coding manual developed by Blum-Kulka (1989:273). As proposed by Blum-Kulka, a request can contain the following components: the head act (the minimal requesting unit), alerters (vocatives, interjections), the directness of the request (grammatical moods such as the imperative, 'want' statements, hints, etc.), syntactic downgraders (interrogative form, tense, and aspect), and lexical and phrasal downgraders (politeness expressions such as 'please', hedges, cajolers, etc.).

[7] A set of eight index cards were presented to each informant. Each card contained the same request, but worded differently. They were asked to order them from least to most polite. A similar ranking process was used by Koike (1989:195) for Portuguese and by Fraser and Nolen (1981:106) for Spanish.

[8] This structure is not found in Southern Peruvian varieties of Quechua (Gutiérrez 1990).

[9] The use of *give* as a modal verb has been translated into the Spanish spoken in the Ecuadorian Andes, and is used like its Quichua counterpart. An example would be: *Dáme abriendo la ventana* 'Do me the favor of opening the window'.

[10] The abbreviations in this paper are the following:

dim.	diminutive
acc.	accusative case, marking the direct object
imp.	imperative, or command form
pres.	present
fut.	future
neg.	negative
quest.	question
pers.	person, as in the form of the verb
inf.	infinitive ending
subj.	subjunctive mood
pl.	plural

[11] Gutiérrez 1990 reports that the suffix *-cu* in Peruvian Quechua is used instead of the words *allichu* and *ichu*, which are commonly used to soften requests. She provides the following example: *Allichu tantata ruwapuway?* 'Would you please make bread for him/her?'. In this case, *allichu* is translated by 'would you please...?'.

[12] The same tense usage was observed in the Spanish sample from the Otavalo area. An example is: *Deja por ahora - buscarás algún rato* 'Leave (it) (present imperative) for now, look for (it) (future imperative) some other time)'.

[13] Questions are commonly used to convey requests, especially questions of ability and availability. In this study, questions are considered requests when they are generated in response to a stimulus using the requesting verb *mañana* 'to ask' in the elicited data. In naturally-occurring conversations, questions are considered requests when they expect the hearer to comply by performing an action.

[14] This strategy was only used by one informant whose speech was highly influenced by Spanish.

[15] This example also shows the heavy syntactic influence of Spanish in this speaker's Quichua. The word order is completely Spanish (subject, verb, object) and this sentence could be translated into Spanish merely by substituting Spanish vocabulary (with the exception of the final preposition *-pi* 'in').

[16] Gutiérrez 1990 reports that in Southern Peruvian Quechua *-lla* is affixed to words before the tense/aspect marker *-y*, as in *shamullây* '**Just** come on'. This was also reported for the Quichua spoken in the Ecuadorian jungle (Catta Quelen 1987). The positioning of *-lla* in word final position is apparently typical of the Quichua spoken in Imbabura (Stark and Carpenter 1973). It could be theorized that *-lla* was moved to word final position in this area to "match" the equivalent structure in Spanish: *Venga **nomás*** '**just** come', *Shamuyllá* '**Just** come'.

[17] The future tense is used in the Spanish of Otavalo to formulate requests that are to be executed at a time subsequent to the present. This is the same tense usage as in Quichua.

[18] In Southern Peruvian Quechua the equivalent expression is *ama hina kaychu* (Gutiérrez 1990).

REFERENCES

ALBOR, Hugo. 1973. *Da* + gerundio, ¿un quechuismo? Hispania 56.316-318.

AUSTIN, John. 1962. How to do things with words. London: Oxford University Press.

BLUM-KULKA, Shoshona, Juliane House, & Gabriele Kasper. 1989. Cross-cultural pragmatics: Requests and apologies. Norwood, N.J.: Ablex Publishing Corp.

BROWN, Penelope. 1979. Language, interaction, and sex roles in a Mayan community: A study of politeness and the position of women. University of California Ph.D. dissertation.

——, & Stephen Levinson. 1978. Universals in language usage: Politeness phenomena. Questions and politeness, ed. by E. Goody, 56-288. Cambridge, Mass.: Cambridge University Press.

CARPENTER, Lawrence. 1982. Ecuadorian Quichua: Descriptive sketch and variation. University of Florida Ph.D. dissertation.

CATTA QUELEN, Javier. 1987. Gramática del quichua ecuatoriano. Quito, Ecuador: Ediciones Abya-Yala. Centro de investigaciones para la educación indígena.

CENTRO DE INVESTIGACIONES PARA LA EDUCACIÓN INDÍGENA. 1983. Ñucanchic llactapac shimi. Vol. 1. Quito, Ecuador.

CORDERO, Luis. 1989. Diccionario quichua. Quito, Ecuador: Corporación Editora Nacional.

ERVIN-TRIPP, Susan. 1976. Is Sybil there? The structure of some American English directives. Language in Society 5:1. 25-66.

ESCOBAR, Alberto. 1986. Types and stages of bilingual behavior: A sociopragmatic analysis of Peruvian bilingual Spanish. State University of New York at Buffalo Ph.D. dissertation.

FRASER, Bruce, and Nolen, William. 1981. The association of deference with linguistic form. International Journal of the Sociology of Language 27.93-103.

FROMKIN, Victoria, & Robert Rodman. 1988. An introduction to language. Fourth edition. Chicago: Holt, Rinehart and Winston.

GOFFMAN, Irving. 1967. Interaction ritual: Essays on face-to-face behavior. Garden City, N.Y.: Anchor Books.

GRIMM, Juan. 1896. La lengua quichua: dialecto de la República del Ecuador. Quito, Ecuador: Proyecto de educación intercultural.

GUTIÉRREZ, Salomé. 1990. Personal communication.

GUZMÁN, Manuel. 1920. Gramática de la lengua quichua. Quito, Ecuador: Prensa católica.

HAVERKATE, Henk. 1979. Impositive sentences in Spanish. New York: North-Holland Publishing Co.

HURLEY, Joni K. 1992. A cross-cultural pragmatic study of Spanish and Quichua request strategies as influenced by language contact in Otavalo, Ecuador. University of Pittsburgh Ph.D. dissertation.

HILL, Jane, & Kenneth Hill. 1986. Speaking Mexicano: Dynamics of syncretic

language in central Mexico. Tucson: The University of Arizona Press.
JAKE, Janice. 1979. Validation suffixes in Imbabura Quechua. Papers from the fifteenth regional meeting of the Chicago Linguistic Society, ed. by P. Clyne, W. Hanks, & C. Hofbauer, 172-184.
JARA, Fausto. No year. Morfología quichua. Quito, Ecuador: Ediciones Mundo Andino.
KOIKE, Dale. 1986. Differences and similarities in men's and women's directives in Carioca Brazilian Portuguese. Hispania 69:2.387-394.
——. 1989. Requests and the role of deixis in politeness. Journal of Pragmatics 13.187-202.
LEONARDI, P. José. 1966. Lengua quichua (dialecto del Napo) gramática y diccionario. Quito, Ecuador: Editora `Fénix'.
MUGICA, P. No year. Aprenda el quichua: Gramática y vocabularios. Quito, Ecuador: CICAME.
MUYSKEN, Pieter. 1981. Halfway between Quechua and Spanish: the case for relexification. Historicity and variation in creole studies, ed. by A. Highfield & A. Valdman, 52-78. Ann Arbor: Karoma Publishers Inc.
NAULA GAUCHO, Juan. 1975. Bosquejo gramatical del quichua de Chimborazo. Quito, Ecuador: Universidad Central.
Ñucanchic llactapac shimi. Vol. 1.1983 Quito, Ecuador.
ORR, Carolyn. 1981. Algunos rasgos característicos del discurso en el quechua del Napo. Revista latinoamericana de estudios lingüísticos 1.135-175.
PARÍS, J. 1961. Gramática de la lengua quichua. Quito, Ecuador: Editorial Santo Domingo.
PARKER, Gary. 1964. Gramática del quechua ayacuchano. Lima, Perú: Universidad Nacional Mayor de San Marcos.
QUINTERO, María, & María Cotacachi. 1986. Quiquinllatac quichua shimita yachacupai. Vol. 4. Quito, Ecuador: Ediciones Abya-Yala.
RUSHFORTH, Scott. 1985. Some directive illocutionary acts among the Bear Lake Athapaskans. Anthropological Linguistics 27:4.387-411.
SALOMON, Frank. 1983. El quichua de los Andes ecuatoriales: algunos aportes recientes. Revista Andina 1:2.393-405.
SEARLE, John. 1976. A classification of illocutionary acts. Language in Society 5.1-23.
——. 1979. Expression and meaning: Studies in the theory of speech acts. Cambridge, Mass.: Cambridge University Press.
SOLÁ, Donald, & D. Yupanqui. 1970. Quechua hablado. New York: Cornell University.
STARK, Louisa, & Lawrence Carpenter. 1973. El quichua de Imbabura: Una gramática pedagógica. Otavalo, Ecuador: Instituto Interandino de Desarrollo.
STRATFORD, Billie. 1989. Structure and use of Altiplano Spanish. University of Florida Ph.D. dissertation.
TOSCANO MATEUS, Humberto. 1953. El español en el Ecuador. Madrid, Spain.
URIOSTE, Jorge. 1955. Gramática de la lengua quechua y vocabulario. La Paz, Bolivia: Editorial Canata.

VÁSQUEZ, P. Victor. 1990. Runa shimita yachacushun: Aprendamos quichua. Quito, Ecuador: Ediciones Abya-Yala.

WARDHAUGH, Ronald. 1985. How communication works. New York: Basil Blackwell.

WHERRITT, Irene. 1983. Directives in Brazilian Portuguese: Mother-child interaction. Spanish and Portuguese in social context, ed. by J. Bergen & G. Bills, 105-118. Washington, D.C.: Georgetown University Press.

WILSON, Robert. 1965. Polite ways to give orders. Hispania 48:1.

Studies in the Linguistic Sciences
Volume 23, Number 2, Fall 1993 [publ. October 1996]

ADVANCEMENT IN SOME ASIAN AND AFRICAN LANGUAGES*

Nkonko M. Kamwangamalu
University of Swaziland

'Advancement', a Relational Grammar rule which promotes a nominal bearing a given grammatical relation in a clause to a higher relation in the same clause (Perlmutter 1983), has been one of the central themes in Relational Grammar (RG) for the past twenty years or so. In RG, examples of advancement include such traditional rules as dative movement, raising, and passive. This paper discusses advancement of accusative, dative, and locative nominals in passive constructions in some South Asian and African languages, with a focus on Hindi and Ciluba. The paper is especially concerned with the claim in RG that 'the relational network of every passive clause in any human language has a nominal bearing the 2-relation and the 1-relation in successive strata (Perlmutter & Postal 1983:17). The data presented not only challenges this claim, but also has far-reaching implications for the relational laws resulting therefrom, viz. the Agreement Law, the Chomeur Law, and the Stratal Uniqueness Law. The implications of the data for relational concepts such as 'Terms' will also be discussed. It will be suggested that RG modify its claim, laws, and concepts to accommodate the data presented here and elsewhere in the literature on South Asian (e.g. Y. Kachru et al. 1976, Pandharipande 1981, Hock 1982, Mohanan 1990) and African (e.g. Dalgish 1976) languages.

1. Introduction

The subject of this paper is 'advancement', a Relational Grammar (RG) rule that promotes a nominal bearing a given grammatical relation in a clause to a higher relation in the same clause (Perlmutter 1983). The paper aims to discuss advancement of accusative, dative and locative nominals in passive constructions in some Asian and African languages, and in Hindi and Ciluba in particular. More specifically, the paper addresses the claim in RG that the 'relational network of every passive clause in any human language has a nominal bearing the 2-relation and the 1-relation in successive strata (Perlmutter & Postal 1983:17). Before I discuss this claim and the conditions or laws resulting therefrom, I shall, by way of background, first present a brief introduction to RG theory. Subsequently, I shall discuss accusative advancement in Hindi, and accusative, dative, and locative advancement in Ciluba, with a focus on how these nominals achieve subjecthood in passive constructions. This will be followed by a discussion of the implications of the Hindi and Ciluba data for the RG's claim under consideration. It is worth noting here that the discussion of advancement in Hindi will draw heavily from previous works in which this topic has received extensive coverage (e.g. Y. Kachru

1980, Pandharipande 1981, Hock 1982, Mohanan 1990). This discussion will be limited to accusative nominals only because these are the only ones that can advance to subject in passive constructions in Hindi and related languages (e.g. Y. Kachru et al. 1976, Hock 1982).

2.0 Relational Grammar

2.1 Background

Central to RG theory is the notion of grammatical relations in a clause. RG views a clause as consisting of a network of grammatical relations such as SUBJECT (SU), DIRECT OBJECT (DO), INDIRECT OBJECT (IO), LOCATIVE (Loc), INSTRUMENTAL (Ins), BENEFACTIVE (Ben), etc. These are referred to as primitives of syntactic theory. The primitives are divided into two main categories: central relations and oblique relations. Central relations include SU, DO, and IO, known as TERMS or as the 1-relation, the 2-relation, and the 3-relation, respectively. Oblique relations include the remaining relations, viz. Loc, Ins, Ben, etc. These are known as NON-TERMS.

Also central to RG is the notion of linguistic levels. RG argues that multiple syntactic levels must be recognized in the analysis of clause structure. This is because in a clause a nominal may bear a range of relations to the predicate at different syntactic levels and also because certain syntactic phenomena are sensitive to some grammatical relations but not to others. For example, in (1a) below the term *banana* bears the 2-relation to the predicate, while in (1b) it bears the 1-relation. Similarly, in (2a) the term *Paul*, for instance, bears the 3-relation to the predicate, whereas in (2b) and (2c), it bears the 2-relation and the 1-relation, respectively. Related to the question of linguistic levels is the distinction in RG between initial and final grammatical relations. For instance, in (1a) the term *child* is an initial 1, while the term *banana* is an initial 2. In (1b), however, the term *child* bears the chomeur relation to the predicate, while the term *banana* bears the final 1-relation.

(1) a. The child ate **the banana.**
b. **The banana** was eaten by the child.

(2) a. John gave food to **Paul**
b. John gave **Paul** food
c. **Paul** was given food by John

Similar examples can be drawn from Asian languages, e.g. Malayalam and Hindi, or from African languages, e.g. Ciluba and Lingala, as shown in (3)-(4). In (3a) the highlighted terms each bear the 2-relation to the predicate, while in (3b) they bear the 1-relation. The data in (4a) shows that the term *Paul* bears the 3-relation in the initial stratum, while where applicable in the final stratum in (4b) it bears the 1-relation.

(3) a. 'The child ate the banana.'
(M=Malayalam, H= Hindi, C=Ciluba, L= Lingala)

M: **kutti** param tunnu
child-N / banana-N / eat-PT

H: **bacce**-ne kelaa k^{h}aayaa
child-Erg / banana-N / eat-Perf

C: **mu-ana** u-aku-di-a ci-bota
pf-child / Ag-PT-eat-FV / pf-banana

L: **mu-ana** a-li-aki e-tabi
pf-child / Ag-eat-PT / pf-banana

b. 'The banana was eaten by the child.'

M: **kutti**yaal param tinnappetu.
child-Ins / banana-N / eat-PSV-PT

H: **bacce**-dvaaraa kelaa k^{h}aayaa gayaa
child-through / banana-N / eat-Perf go-Perf

C: **ci-bota** ci-aku-di-ibwa kudi mu-ana
pf-banana / Ag-PT-eat-PSV / by / pf-child

L: **e-tabi** e-li-am-aki na mu-ana
pf-banana / Ag-eat-PSV-PT / by / pf-child

(4) a. 'John gave food to Paul (gave Paul food).'

M: **John** Paul b^{h}aksanam kotu^{tt}u.
John-N / Paul-D / food / give-PT

H: **John** Paul k^{h}aanaa diyaa
John-E / Paul-D / food-N / give-Perf

C: **Jean** u-aku-pa *ci-akudia Paul (Paul ci-akudia)
John / Ag-PT-give / pf -food / Paul (Paul / pf-food)

L: **Jean** a-pes-aki *bi-lei Paul (Paul bi-lei)
John / Ag-give-PT / pf-food / Paul (Paul / pf-food)

b. 'Paul was given food by John'

M: X (no equivalent)

H: **Paul**-ko John-dvaaraa k^{h}aanaa diyaa gayaa
Paul-D / John-through / food-N / give-Perf / go-Perf

C: **Paul** u-aku-p-ibw-a ci-akudia kudi Jean
Paul / Ag-PT-give-PS-FV / pf-food / by / John

L: **Paul** a-pes-am-aki bi-lei na Jean
Paul / Ag-give-PSV-PT / pf-food / by / John

Considering data such as (1)-(4), the question is how does RG explain the fact that the term *Paul*, for instance, which is an initial 3 in (2a), turns out to be a final 1 in (2c). This is where the notion of 'advancement' comes into the picture, a point to which I turn below.

2.2 Advancement and passive in RG

In view of data such as (1)-(4), and in particular the English data in (1) and (2), Perlmutter (1983:17) makes the claim given in (5) about advancement in passive clauses not just in English but in human languages in general:

(5) the RN of every passive clause in any human language has a nominal bearing the 2-relation and the 1-relation in successive strata

In line with this claim Perlmutter 1983 defines passive as:

(6) a rule which sanctions 1-hood in an immediately successive stratum for a nominal which is a 2 of a clause at a stratum in which some nominal is a 1.

In other words, what both (5) and (6) mean is that in passive constructions nothing can be a final 1 (i.e. subject) which was not a 2 (i.e. direct object) in a preceding stratum. Advancement in passive constructions, as defined above, is governed by certain restrictions or laws in RG terminology, including the following (Perlmutter & Postal 1983:88-101, Frantz 1981:71):

(7) a. The AGREEMENT LAW:
'only Terms can trigger verb agreement. That is, only a nominal bearing Term relation in some stratum may trigger verb agreement.'

b. The CHOMEUR LAW:
If some nominal N_a, bears a given Term relation 'n' in a given stratum S_i, and some other nominal N_b bears the same Term relation in the following stratum, S_{i+1}, then N_a bears the chomeur relation (n) in S_{i+1}.

c. The STRATAL UNIQUENESS LAW:
Each Term bears one and only one grammatical relation to the predicate.

It is the above claims and laws that I shall be concerned with in this paper. It should be pointed out that these claims and laws have been challenged in recent literature on the syntax of Hindi and related languages (e.g. Pandharipande 1981). My concern here is to determine to what extent the claims and laws are applicable in passive constructions involving accusative, dative and locative nominals in Ciluba. First, I shall argue that contrary to the RG view of passive, in Ciluba dative and locative nominals may passivize directly from their initial grammatical relation as 3 or loc to the 1-relation, and that attempting to advance these nominals to 1 via 3>2 or loc>3>2 would yield ungrammatical sentences. Second, I shall show that in Ciluba the distinction between terms and non-terms does not hold since, contrary to the Agreement Law, non-terms do also trigger verb agreement in this language. Third, I shall show that the facts of Ciluba receive support from previous works on languages as distant as Asian languages, such as Hindi. For instance, there is evidence from the literature on Hindi syntax that shows that contrary to one of the RG laws referred to earlier, viz. the Stratal Uniqueness Law, in Hindi a term may simultaneously bear two grammatical relations to the predicate, the subject relation on the one hand, and the direct object relation on the other (e.g. Y. Kachru 1980, Pandharipande 1981).

3.0 Hindi and Ciluba

3.1 Background

Hindi is an Indo-Aryan language spoken on the Indian subcontinent. Ciluba is a Bantu language spoken in the Republic of Zaire. Both languages differ in many important respects. Here I shall highlight some of the features that are relevant to this paper. In terms of word order, Hindi is an SOV language, while Ciluba is an SVO language. In Hindi direct daughters of S can scramble freely, but this is not allowed in Ciluba, the latter being a strict word order language. Hindi has a case-marking system whereby in a clause the syntactic function of a given nominal is signaled. In the clause *Ninaa-ne bacce-ko kitaab dii* 'Nina gave the child a book', the nominal *Ninaa* carries the ergative case while the nominal *bacce* carries the dative case, as signaled by the clitics *-ne* and *-ko*, respectively. In Hindi syntax, a nominal that does not bear a clitic, such as *kitaab* 'book' is conventionally assumed to bear a nominative case (Y. Kachru et al. 1976, 1977; Y. Kachru 1980; Pandharipande 1981, 1990; Mohanan 1990). In terms of agreement, in Hindi a verb agrees in number, gender, person with its subject if it is nominative. And if the subject is not nominative, the verb agrees with the object if that is nominative (Mohanan 1990:14).

In Ciluba, as in most Bantu languages (e.g. Bresnan & Kanerva 1989), a finite verb must agree with its subject noun in person, number and noun class by means of an agreement prefix. To ensure subject-verb agreement, each Bantu language, and Ciluba is not an exception, has a noun class system whereby each noun consists of two basic morphemes, a noun prefix and a noun stem. In the noun *ba-ana* 'children', for instance, *ba-* is the noun prefix, and *-ana* the noun stem. The noun prefix provides a clue to determining the type of agreement that must obtain between a subject noun and a verb (Kamwangamalu 1985:110). For instance, in the clause *ba-ana ba-di ba-dila* 'the children are crying' the prefix *ba-* in *ba-ana* ensures that whatever verb comes after the noun *ba-ana* 'children' must bear this same prefix for agreement, as evidenced by the presence of the prefix *ba-* in *ba-di* 'are' and *ba-dila* 'crying'.

Hindi and Ciluba may be different from each other in many other important respects, but describing such differences is beyond the scope of this paper: advancement of accusative nominals in Hindi, and accusative, dative and locative nominals in Ciluba.

3.2 Accusative/dative/locative nominals and subjecthood in Hindi and Ciluba

It is generally agreed that a nominal that bears an accusative case ranks higher in the subject accessibility hierarchy. The questions I would like to raise in this section concern mainly accessibility to subject of dative and locative nominals. First, can dative and locative nominals advance to 1-relation (i.e. subject) in passive constructions in Hindi and Ciluba and, if they can, how is this advancement process done? Is it the case that a dative/locative nominal that advances to 1 does so in one step, that is from its initial grammatical relation as 3/loc to the sub-

ject relation; or does it achieve subjecthood through intermediate stages, such as exemplified in (2a-c) above?

Let us first address the question of subjecthood of accusative/ dative/locative nominals in Hindi and Ciluba, digressing briefly on the concept of subject. I shall start with Hindi, drawing heavily on the works of Y. Kachru 1980, 1981, 1990, Pandharipande 1981, and Mohanan 1990. According to the works just cited, in Hindi there are two types of nominals that are considered canonical or unmarked subjects, viz. the ergative subjects and the nominative subjects. However, such nominals are not the only ones that can function as grammatical subject in a Hindi clause. Other nominals that behave like subject include those I am concerned with in this paper, viz. the accusative, dative and locative nominals. Determining the subjecthood of these nominals is not a straightforward affair in Hindi. To determine the subjecthood of these or any other nominals most Hindi grammarians usually appeal to syntactic phenomena such as case-marking, agreement, word order, pronominal coreference, passivization, gap control, reflexive binding, conjunction reduction, etc. Here I shall refer to few of these phenomena, as discussed in recent works on Hindi syntax (Y. Kachru 1990, Mohanan 1990). In their works, Y. Kachru and Mohanan are of the opinion that in Hindi a nominal that is claimed to be a subject must behave like one that is, it must have the properties associated with subject in the language, including the following, among others:

i) it must be able to control reflexivization
ii) it must be able to control conjunction reduction
iv) it must be able to control equi-NP deletion

No universality is claimed for these conditions on subjecthood. That is, a nominal that meets these conditions and therefore qualifies for subjecthood in Hindi, for instance, may not necessarily qualify as subject in other South Asian languages and vice versa. It is not surprising, then, that in languages such as Maithili, for instance, dative nominals are treated as objects rather than subjects (e.g. Mishra 1990).

Unlike Maithili, there seems to be enough evidence from recent works on Hindi syntax that in Hindi accusative/dative/locative nominals also behave like subjects (e.g. Pandharipande 1981, Mohanan 1990). While accusative nominals may function as subject with any class of predicate, there are in Hindi certain classes of predicates which govern dative/locative subjects. For instance, Y. Kachru 1990 notes that predicates that denote a set of 'inherent properties' such as *utsaah* 'enthusiasm', *dhairy* 'patience', *himmat* 'courage', etc. require a locative subject, while those that denote perception (e.g. *dikhaaii denaa* 'to be visible'), liking (e.g. *pasand aanaa* 'to like'), knowledge (e.g. *maaluum honaa* 'to come to know'), etc. require a dative subject. In what follows I present data which show that accusative, dative and locative nominals do indeed have properties associated with subject, for they meet the above and other diagnostics for subjecthood in Hindi.

3.2.1 Accusative/dative/locative subjects in Hindi

3.2.1.1 The reflexive *apnaa* binding

According to Kachru & Bhatia 1977 and Pandharipande 1981, in Hindi the reflexive *apnaa* can take as its antecedent a subject, grammatical or logical, but no other argument. In the literature this phenomenon is also known as reflexive binding. The data in (8)-(9) is illustrative. In both (8a) and (9a) the dative nominals *Rita* and *Vijay* are the logical subjects in their respective structures and, therefore, they qualify as antecedent of the reflexive *apnaa*. (8b) and (9b) show that in contrast to the reflexive, a pronoun cannot be coreferent with the subject of its minimal clause (Mohanan 1990). The dative subjects in the (b) sentences in (8)-(9) therefore cannot be coreferential with the pronoun *uske*.

(8) a. ritaa-ko apnaa ghar bahut yaad aa rahaa thaa
Rita-Dat / self's / home / much / memory / coming / was
'Rita$_i$ was missing self's$_i$ home very much.'
(Y. Kachru 1990:70)

b. ritaa-ko ghar uska bahut yaad aa rahaa thaa
Rita-Dat / home / pron / much / memory / coming was
'Rita$_i$ was missing her$_{j/*i}$ home very much.'

(9) a. vijay-ko kitaab apnee g^{h}ar-me milii
Vijay-Dat / book-N / self-Gen house-L / find-Perf
'Vijay$_i$ found the book in self$_{i/*j}$ house.'
(Mohanan 1990:197)

b. vijay-ko kitaab uske g^{h}ar-me milli
Vijay-Dat / book-N / pron / home-L / find-Perf
'Vijay$_i$ found the book at his$_{j/*i}$ home.'

Reflexive binding, as described above, holds not only for dative nominals but also for locative nominals, as can be seen in (10). (10a) shows that the locative nominal, *niina-me*, is the only eligible antecedent of the reflexive *apnaa*. In the Hindi grammarians' view, this suggests that either the logical subject, namely the locative *niina-me* is the subject, or that there is no subject at all in (10a). It is noted that the facts of pronominal coreference support the former alternative. Pronouns cannot be coreferent with the subject in their minimal finite clause. This is borne out in (10b), where it is shown that the pronoun *uskii* is not coreferent with the locative nominal *Ninaa-me*. This suggests that the latter is indeed the subject in both (10a) and (10b) (e.g. Mohanan 1990:235-36). (The list of abbreviations used in the data below is given in the footnotes section[1]).

(10) a. niinaa-me apnii mausii-ke liye badii mamtaa h
Nina-Loc / self-Gen / aunt-Gen / for / much / affection-N / be-pres
'Nina$_i$ has a lot of affection for self's$_i$ aunt.'

b. niinaa-me uskii mausii-ke liye badii mamtaa hai
Nina-Loc / pron-Gen / aunt-Gen / for / much / affection / be-pres
'Nina$_i$ has a lot of affection for her$_{*i}$ aunt.'

The facts of reflexive binding presented in (8)-(10) obtain also in construc-

tions with accusative nominals, as illustrated in (11). Note that (11a) is the active counterpart of the passive construction in (11b). In (11a) the ergative nominal *John*, the unmarked subject, is obviously the eligible antecedent of the reflexive *apne*, as required in Hindi. In (11b), however, the ergative nominal under consideration has been demoted from its initial grammatical relation of subject to the chomeur relation as a result of passive, thus leaving the initially accusative nominal, *Paul*, as the binder of the reflexive *apne*. Since the latter can only have a subject, logical or grammatical, as its antecedent, it is correct to assume that the nominal *Paul* is the grammatical subject, and it is, in the passive construction in (11b). It is worth pointing out here that in addition to (11b), there is an alternative passive to the construction in (11a). This alternative, which I shall discuss later, is given in (11c). This construction differs from (11b) in terms of case-marking: in (11c) the nominal *Paul* is case-marked, while in (11b) it is not case-marked.

(11) a. John-ne Paul-ko apne kamre me dekhaa
John-Erg / Paul-Acc / self / room / in / saw-Perf
'John$_i$ saw Paul$_j$ food in self$_i$/*$_j$ home.'

b. Paul apne kamre me dekhaa gayaa
Paul / self / room / in / seen / was-Perf
'Paul$_j$ was seen (by John$_i$) in self$_j$/(*$_i$) home.'

c. Paul-ko apne kamre me dekhaa gayaa
Paul-Acc / self's / room / in / seen / was-Perf
'Paul$_i$ was seen in self's$_i$ room.'

3.2.1.2 Conjunction reduction[2]

The data in (12)-(13) shows that the dative subject behaves like a subject because it controls conjunction reduction, as in (12), though it does not undergo this process, as can be seen in (13) (e.g. Y. Kachru 1990:63).

(12) tasviir dekh kar use gussa aayaa
picture / see / CP / him / Dat / anger came
'He$_i$ became angry_$_i$/having seen the picture.'

(13) *gussa aa kar us-ne sab ko bahut DaaTaa
anger / come / CP / he-Erg / all / DO / much / scolded
'He$_i$ scolded everyone_*$_i$ having become angry.'

3.2.1.3 Equi-NP deletion

In Hindi, like subject the dative nominal both controls and undergoes equi, as shown in (14) and (15) (Y. Kachru 1990: 64).

(14) larke-ko film dekhnaa pasand hai
boy-Dat / film / viewing / liking / is
'The boy likes to view films.'

(15) larke-ne film pasand aane kii carcaa nahii kii
boy-Erg / film / liking / coming of / mention / not / did
'The boy did not mention (his) liking the film.'

As can be seen from the data presented above (e.g. (8)-(17)), dative, locative and accusative nominals prove to function as subject in Hindi, a point that is demonstrated at length by Y. Kachru and Mohanan. Rather than pursue this point any further, I shall assume the correctness of the conclusions reached by Y. Kachru and Mohanan and others regarding the subjecthood of the above-mentioned nominals in Hindi and will, instead, focus on how these nominals achieve their status as subject in this language. But first, a word on the subjecthood of accusative, dative and locative nominals in Ciluba.

3.2.2 Accusative/dative/locative subjects in Ciluba

We have seen that in Hindi one needs a number of diagnostics to show that accusative/dative/locative nominals can behave like subjects. In Ciluba, however, the situation is much simpler. Compared to Hindi, in Ciluba it simply takes one test to determine the subjecthood of not just accusative/dative/locative nominals, but of any nominal that claims subjecthood in a Ciluba clause. The most common test is agreement: In Ciluba, as in related Bantu languages (e.g. Swahili, Lingala, Kikongo), the verb must agree in person, number and noun class with nothing else but the subject, as can be seen from (16)-(17). In (16) and (17a) the verb agrees with the nominative nominals *mwana* 'child' and *bibota* 'bananas', respectively, while in (17b) the verb agrees with the inverted locative *pa-mesa* 'on the table.'

(16) mu-ana u-aku-di-a bi-bota
pf/sg-child / Ag-PTs-eat-FV / pf.pl-bananas
'The child ate the bananas.'

(17) a. bi-bota bi-di **pa-mesa**
pf.pl-banana / Ag-are. / Loc.on-table
'The bananas are on the table.'
b. **pa-mesa** pa-di bi-bota
Loc.on-table / Ag-are / pf.pl-banana
'Lit: On the table is (are) bananas.'

Agreement, as shown in (16)-(17), obtains also in passive constructions with accusative, dative and locative nominals, as can be seen in (18)-(19). Note that in the active clause in (18a), the verb *-pa* 'give' agrees with the subject *John* by means of the (singular) agreement prefix *u-*. Note also that in (18) both the initially accusative nominal, *ci-akudia* 'food', and the initially dative nominal, *ba-ana* 'children' each can be passivized, as shown in (18b) and (18c), respectively. In the passive construction in (18b), the verb agrees with the initially dative nominal *ba-ana* 'children', which in this case is the grammatical subject of the clause under consideration. Here agreement is done by means of the (plural) agreement prefix *ba-*. In (18c), the verb agrees with the initially accusative nominal *ci-akudia* 'food' by means of the agreement prefix *ci-*. In (19a) agreement is the same as in (18a). In (19b), which is the passive counterpart of (19a), the verb agrees with the locative nominal *mu-cikuku* 'in the kitchen' by means of the locative prefix *mu-*.

(18) a. Jean u-aku-p-a ba-ana ci-akudia
John / Ag-PTs-give-FV / pf.pl-child / pf-food
'John gave the children food/food to the children.'

b. **Ba**-ana **ba**-aku-p-ibw-a ci-akudia kudi Jean
pf.pl-child / Ag-PTs-give-PSV-FV / pf-food / by / John
'The children were given food by John.'

c. **Ci**-akudia **ci**-aku-p-ibw-a ba-na kudi Jean
pf.sg-food /Ag-PTs-give-PSV-FV / children / by / John
'Food was given to the children by John.'

(19) a. Jean u-aku-p-a ba-ana ci-akudia mu-cikuku
John / Ag-PTs-give-FV / pf.pl-child / pf-food / Loc.in-kitchen
'John gave the children food in the kitchen.'

b. **mu**-cikuku **mu**-aku-p-ibw-a ba-ana ci-akudia kudi Jean
Loc.in-kitchen / Ag-PTs-give-PSV-FV / pf-child / pf-food / by / John
Lit: 'In the house was given the children food by John.'

In addition to the facts presented in (18)-(19), elsewhere I have shown that in Ciluba, accusative, dative, and locative nominals behave like subject-Terms not only in terms of their ability to govern agreement on the verb, but also in terms of other properties associated with Terms, such as the ability to passivize, to relativize, to incorporate onto the verb, to cleft, and to topicalize (e.g. Kamwangamalu 1985)

Having shown that in both Hindi and Ciluba accusative/dative/locative nominals may also function as subjects, I shall now move on to the other concern of this paper, viz. how these nominals achieve their status as subject in passive constructions in the languages under consideration.

4. Accusative/dative/locative advancement to subject in Hindi/Ciluba

It was observed earlier that in Hindi, dative and locative subjects are base-generated rather than derived through processes such as advancement. Therefore, they will not be included in the discussion of advancement that follows. As background for this discussion, let us recall the claim in (5) regarding RG's conception of the relational network of a passive clause. Again, RG claims that the relational network of every passive clause in any human language has a nominal bearing the 2-relation and the 1-relation in successive strata. Applying this claim about passive to Hindi and Ciluba, the following analyses can be envisaged for accusative (in addition to dative/ locatives for Ciluba) advancement in passive constructions in these languages.

One analysis, which follows directly from and is consistent with the above-stated claim of RG, is that in Hindi and Ciluba, an accusative nominal behaves like a subject that has undergone 2 to 1 advancement. Following this analysis, Ciluba nominals such as locative/dative, for instance, cannot be promoted to subject unless they have first undergone loc>3>2 / 3>2 advancement, respectively.

The other analysis, one that I shall suggest in this paper, is that for Ciluba, locative/dative nominals do not have to undergo loc>3>2>1 / 3>2>1 advance-

ment, and that they undergo loc/3>1 advancement instead. For Hindi, the literature (e.g. Pandharipande 1981, Mohanan 1990) suggests that an accusative nominal may behave like subject in a given construction without necessarily having undergone 2>1 advancement. This analysis conflicts with the claim of Perlmutter and others (e.g. Johnson 1981), but it is consistent with the data of Hindi and Ciluba presented thus far in this paper. For the sake of illustration, let us look again at the passive constructions given earlier in (11) for Hindi and in (4) for Ciluba, repeated here below as (20) and (21), respectively.

Regarding Hindi, it is clear that in (20b) the accusative nominal *Paul* has advanced to 1, as can be concluded from the absence of the accusative case on the nominal under consideration. In (20c), however, there is no evidence that advancement has taken place. The presence of the accusative case on the nominal *Paul* attests to this conclusion. Of crucial importance regarding (20c) is that in this construction the accusative nominal *Paul* is the only eligible antecedent of the reflexive *apnee*. Recall that in Hindi *apnee* can have nothing else but a subject as its antecedent. It follows that the accusative nominal *Paul*, the only antecedent of *apne*, is the subject of the passive construction in (20c). In a sense, then, it can be concluded that in (20c) the subjecthood of the accusative-marked nominal *Paul* is not dependent on its promotion to 1, and that promotion of this nominal to 1 is actually optional. Hock (1985:66) draws similar conclusions regarding advancement in Sanskrit of non-terms and terms to direct object and subject, respectively. He notes (p. 66) that ... 'if a non-term, adverbial constituent shows case variation between, say, locative and accusative, promotion to direct object status is possible only if there is no other direct object ... and that even under these conditions, promotion of that accusative-marked NP to subject of the passive is only optional.' The point here is to show that the facts of Hindi presented in (20c) are not an isolated case, and that they obtain in other Southeast Asian languages as well, such as Sanskrit. While these facts accord well with Hindi syntax they, obviously, conflict with the 2>1 analysis as well as with some of RG laws presented earlier in this paper, such as the Stratal Uniqueness Law. Again, by virtue of this law, each term bears one and only one grammatical relation to the predicate (Perlmutter 1983:88). Now, reconsider the construction in (20c). As was pointed out above, in this construction the term *Paul* bears not one but two grammatical relations to the predicate: First, *Paul* is a direct object because of its case, it bears the accusative case; second, *Paul* is the grammatical subject in the construction under consideration because it is the only eligible antecedent of the reflexive *apne*: in Hindi, only a nominal that is a subject can be the binder of the reflexive *apne*.

(20) a. John-ne Paul-ko apne kamre me dekhaa
John-Erg / Paul-Acc / self / room / in / saw-Perf
'John$_i$ saw Paul$_j$ food in self$_{i/*j}$ home.'

b. Paul apne kamre me dekhaa gayaa
Paul / self / room / in / seen / was-Perf
'Paul$_j$ was seen (by John$_i$) in self$_j$/(*$_i$) home.'

c. Paul-ko apne kamre me dekhaa gayaa
Paul-Acc / self's / room / in / seen / was-Perf
'Paul$_i$ was seen in self's$_i$ home.'

For Ciluba, the data in (21)-(22) suggest that accusative nominals can undergo 2>1 advancement, much as they can in English and other languages. For dative nominals, however, the data show that when such nominals advance to subject, they do so in one leap only that is, from their initial relation as dative to subject relation, and not through 3>2>1 advancement. Any attempt to advance a dative nominal for instance to 2 first and then to 1 results in an ungrammatical structure, as can be concluded from (21b).[2]

(21) a. Jean u-aku-p-a **ba-ana** ci-akudia
John / Ag-PTs-give-FV / pf.pl-child / pf-food
'John gave the children food/food to the children.'

b. *Jean u-aku-p-a ci-akudia **ba-ana**
Jean / Ag-PTs-give-FV / pf-food / pf.pl-child

c. **Ba-ana** ba-aku-p-ibw-a ci-akudia kudi Jean
Pr.pl-child / Ag-PTs-give-PSV-FV / pf-food / by / John
'The children were given food by John.'

Similar conclusions obtain also for locative advancement in this language, as shown in (22). Here, note that the locative *mu-cikuku* 'in the kitchen' advances directly to 1, as in (22d), and that attempting to advance it to 1 via Loc>3>2>1 advancement would yield unacceptable sentences, as evidenced by (22b,c).

(22) a. Jean u-aku-p-a ba-ana ci-akudia **mu-cikuku**
John / Ag-PTs-give / pf.pl-child / pf-food / Loc.in-kitchen
'John gave the children food in the kitchen.'

b. *Jean u-aku-p-a ba-ana **mu-cikuku** ci-akudia
John / Ag-PTs-give / pf.pl-child / Loc.in-kitchen / pf-food
'John gave the children in the kitchen food.'

c. *Jean u-aku-p-a ***mu-cikuku** ba-ana ci-akudia
John / Ag-PTs-give / Loc/in-kitchen pf/pl-child / pf-food
*'John gave in the kitchen the children food.'

d. **mu-cikuku** mu-aku-p-ibw-a ba-ana ci-akudia kudi Jean
Loc.in-kitchen / Ag-PTs-give-PSV-FV / pf.pl-child / pf-food / by / Jean
'Lit: In the house was given the children food by John.'

The fact that locatives in Ciluba can advance to 1 via loc>1 rather than Loc>3>2>1 advancement is not an isolated case. Dalgish 1976 makes a similar claim with respect to Olutsootsoo, a Bantu language of Kenya, and so does Kimenyi 1974 with respect to Kinyarwanda, a Bantu language of Rwanda. Both Dalgish and Kimenyi show respectively that in Olutsootsoo and Kinyarwanda advancement to 1 is not limited to terms and that locatives can advance to 1 as well. That locatives can advance to 1 is not unique to Bantu languages, but it is also attested to in non-Bantu languages. For instance, quoting Bell 1974 on advancement in Cebwano, a language of the Philippines, Perlmutter and Postal (1984:90) acknowledge that Cebwano allows with great freedom advancement[3] to 1 not only

of 2s and 3s but also of instrumentals, locatives, benefactives, temporals, etc. This freedom of advancement of both terms and non-terms to 1 is also evident in Ciluba, as can be seen from the data in (21) and (22). But what are the implications of such advancement of non-terms to 1 for the relational distinction between terms and non-terms, and for relational laws such as the Agreement law.

Consider, for instance, the Agreement Law. According to this law, which was stated earlier in (7a), only Terms can trigger subject-verb agreement in a clause. Now, consider agreement in (22d), above. This clause shows that the locative nominal *mu-nzubu* 'in the house' agrees with the verb *-pa* 'give' by means of the locative agreement prefix *mu-*. The question that arises here is whether locatives should be treated as Terms. Based on the data presented here, I would like to suggest that the scope of termhood in RG be extended so as to include locatives in languages such as Ciluba, since locatives are shown to behave like subject and especially so with respect to the Agreement Law.

5. Conclusions

In this paper I have been concerned with one of central claims in RG regarding passive, viz. the claim that the relational network of every passive clause in any human language has a nominal bearing the 2-relation and the 1-relation in successive strata. While this claim receives support from languages such as English and other languages around the world, it fails to accommodate data from some Asian and African languages, and from Hindi and Ciluba in particular.

The literature on Hindi provides evidence that in Hindi, there are cases where an accusative nominal can be the subject of a passive clause without necessarily having undergone 2>1 advancement. As a result, contrary to the Stratal Uniqueness Law, in Hindi it is possible that a term bear two grammatical relations to the predicate: the direct object relation on the one hand, and the subject relation on the other, as illustrated in (20c).

Unlike Hindi, accusative advancement in Ciluba accords well with the above RG claim about the relational network of a passive clause. However, the challenge to this claim comes from dative/locative advancement. I have shown that in Ciluba passivization of dative/locative nominals is a one-step process, 3>1 /loc>1 advancement, and that these nominals do not need to undergo 3>2 /loc>3>2 advancement prior to advancing to 1, as claimed in Relational Grammar. Since in Ciluba it is not just dative nominals that passivize directly from 3 to 1, and locatives behave the same way as well, the question is whether taking into account RG laws such as the Agreement Law passivizing locatives should be treated as Terms. In light of the available evidence I have suggested that they should: In Ciluba and related Bantu languages locative nominals behave like Terms not only in terms of their ability to govern subject-verb agreement but also, as I have shown elsewhere (e.g. Kamwangamalu 1985), in terms of their ability to do other things that Terms can do, including the ability to relativize, the ability to passivize, the ability to cleft, to list just a few.

NOTES

*An earlier version of this paper was presented at the 13th South Asian Languages Analysis (SALA) Roundtable, University of Illinois at Urbana-Champaign, Illinois, U.S.A., 25-27 May 1991. I would like to thank the National University of Singapore, where I was then working as a Visiting Lecturer, for sponsoring my participation in the SALA Roundtable. Also, I would like to thank Yamuna Kachru and Tara Mohanan for providing me with the Hindi data which have made this paper possible. Finally, I would like to acknowledge with gratitude the comments of Yamuna Kachru and Hans H. Hock on an earlier version of this paper. I alone am responsible for any remaining errors of interpretation or analysis of the data presented in this paper.

[1] Below are the abbreviations used in this paper:

Dat	= dative;	Erg	= ergative	Nom	= nominative
Loc	= locative;	pf	= prefix	Perf	= perfective
PSV	= passive	pl	= plural	sg	= singular
PTs	= past tense	Gen	= genitive	Ag	= agreement
FV	= final vowel	pres	= present tense		
CP	= conjunctive participle				

[2] Besides, it should be pointed out that in Ciluba a dative (i.e. indirect) object has prominence over an accusative (i.e. direct) object. This explains why in ditransitive constructions a dative object must always be close to the verb, regardless of whether the accusative object is animate or inanimate. For further details, see Kamwangamalu 1985.

[3] Perlmutter and Postal 1984:90 do not specify whether in Cebwano advancement of instrumentals, locatives, benefactives, temporals and other non-terms to subject is done via intermediate stages, such as Inst/Loc/Ben/Temp >3>2>1, or whether it is done in one leap, e.g. Inst/Loc/Ben/Temp >1, as is the case in Ciluba and related Bantu languages (e.g. Lingala, Swahili, Kikongo).

REFERENCES

BRESNAN, Joan, & J. M. Kanerva. 1989. Locative inversion in Chichewa: a case study of factorization in grammar. Linguistic Inquiry 20:1.1-50.

DALGISH, Gerard M. 1976. Passivizing locatives in Olutsootsoo. Studies in the Linguistic Sciences 6:1.57-68.

FRANZ, Donald G. 1981. Grammatical relations in universal grammar. Monograph, Summer Institute of Linguistics: University of Lethbridge.

HOCK, Hans Henrich. 1982. The Sanskrit passive: Synchronic behavior and diachronic development. South Asian Review 6.127-37.

——. 1985. Sanskrit Double-Object Constructions: Will the real object please stand up? Indian Journal of Linguistics 12.50-70.

KACHRU, Yamuna. 1980. Aspects of Hindi grammar. New Delhi: Manohar Publications.

——. 1981. Transitivity and volitionality in Hindi-Urdu. Studies in the

Linguistic Sciences 11.2.191-93.
——. 1990. Experiencer and other oblique subjects in Hindi. Experiencer subjects in South Asian languages, ed. by M. K. Verma and K. P. Mohanan, 59-75. Stanford: The Center for the Study of Language and Information.
——, Braj B. Kachru, & Tej K. Bhatia. 1976. The notion 'subject': A note on Hindi-Urdu, Kashmiri and Panjabi. The notion of subject in South Asian Languages, ed. by M. K. Verma, 79-108. University of Wisconsin, Madison: South Asian Studies.
——, & Tej K. Bhatia. 1977. On reflexivization in Hindi-Urdu and its theoretical implications. Indian Linguistics 38:1. 21-38.
KAMWANGAMALU, Nkonko. 1985. Passivization in Bantu languages: implications for relational grammar. Studies in the Linguistic Sciences 15:1.109-133.
KIMENYI, Alexandre. 1976. A relational grammar of Kinyarwanda. UCLA Ph.D. dissertation in Linguistics.
MISHRA, Mithilesh K. 1990. Dative/experiencer subjects in Maithili. The notion of subject in South Asian Languages, ed. by M. K. Verma, 105-117. University of Wisconsin, Madison: South Asian Studies.
MOHANAN, Tara W. 1990. Arguments in Hindi. Stanford University Ph.D. dissertation in Linguistics.
PANDHARIPANDE, Rajeshwari. 1981. Syntax and semantics of the passive construction in selected South Asian languages. University of Illinois at Urbana-Champaign Ph.D. dissertation in Linguistics.
PERLMUTTER, David. 1983 (ed.). Studies in relational grammar 1. Chicago: The University of Chicago Press.
——, & P. Postal. 1984. The 1-Advancement Exclusiveness Law. Studies in relational grammar 2, ed. by D. Perlmutter and Carol G. Rosen, 81-125. Chicago: The University of Chicago Press.

Studies in the Linguistic Sciences
Volume 23, Number 2, Fall 1993 [publ. October 1996]

ANIMACY DISTINCTIONS IN AKAN GRAMMAR*

Emmanuel Kweku Osam
University of Oregon

Animacy distinctions has never been considered to be one of the outstanding features of the grammar of Akan. However, based on the form and distribution of nominal prefixes in the language and the nature of the pronominal system, it is concluded that the notion of animacy distinction is relevant to the grammar of Akan.

0. Introduction

The purpose of this paper is to show that the grammar of Akan is sensitive to animacy distinctions, and to some limited extent, we could even talk about the presence of animacy hierarchy in the language. Animacy distinctions in Akan appear mainly in nominal affixes and the forms and behavior of pronouns.

1. Nominal affixes

One of the areas in which the distinction between animate and inanimate nouns is shown is in the nominal affixes in the language. The fact that Akan has a nominal prefix system has long been recognized (see for example Christaller 1875, Balmer and Grant 1929, Akrofi 1935, Welmers 1971, 1973, Essilfie 1977, Dolphyne 1988, Dolphyne and Dakubu 1988, Osam 1993 and 1994). In Osam 1993 and 1994, in agreement with Welmers 1971 and 1973, the argument is made that the current noun prefixes are the historical remains of the old noun class system that must have existed in Proto-Akan. The prefixal system as it exists currently in the language is illustrated in (1).

(1)			
	Class 1	o-nua	'sibling'
		o-nipa	'person/human being'
		o-birempon	'a great person'
		ɔ-baa	'woman'
		ɔ-bɔdɔm	'dog'
		ɔ-pɔnkɔ	'horse'
	Class 2	a-bofra	'child'
		a-noma	'bird'
		a-berewa	'old woman'
		e-wi	'thief'
	Class 3	i-dua	'tree'
		i-kur	'sore'
		ɨ-dan	'building'
	Class 4	ɛ-boɔ	'stone'
		ɔ-dan	'building'
		ɛ-woɔ	'honey'

Class 5	n-dua	'trees'
	n-kura	'mice'
	n-twɛr	'frogs'
	n-dowa	'bees'
	m-bofra	'children'
	m-bowa	'animals'
	m-fe	'years'
Class 6	a-ka	'debts'
	a-dan	'buildings'
	a-hɔho	'visitors'
	e-kunyin	'great men'
	e-din	'names'
	e-kuw	'clubs/associations'

As indicated in (1), the prefixes are either vowels or homorganic nasals. Prefixes involving vowels, with the exception of the prefix labeled Class 4, are paired on the basis of vowel harmony (2). The first member of each pair has the advanced tongue root feature (+ATR), and the second member is minus the advanced tongue root feature (-ATR). The nominal prefixes are also distinguished according to number. The prefixes marked Classes 1-4 indicate singular nouns whereas those marked Classes 5 and 6 identify plural nouns.

(2)

SINGULAR		PLURAL	
Class 1	o-/ɔ-	Class 5	n-
Class 2	e-/a-	Class 6	e-/a-
Class 3	i-/ɨ-		
Class 4	ɛ-		

The noun prefixes currently in the language, in various ways, reflect the distinction between animate and inanimate nouns.

1.1 Singular prefix

One of the features of the prefixal system is that there is a semantic motivation associated with the nouns that take a particular prefix. For example, in the singular, the nominal prefix *o-/ɔ-* can be found, with some exceptions[1], on animate nouns. On the other hand, the prefix *e-/ɛ-* goes on inanimate nouns only. This distinction is illustrated in (3). In (3a), the nouns which have *o-/ɔ-* are animate. However, in (3b), only inanimate nouns are shown to have the prefix *e-/ɛ-*. The main point here is that whereas the *o-/ɔ-* are predominantly animate prefixes, the *e-/ɛ-* are inanimate prefixes without exception. In other words, only inanimate nouns occur with *e-/ɛ-* .

(3) a.	o-panyin	'elder'
	ɔ-hɔho	'visitor'
	ɔ-kɔdeɛ	'eagle'
	ɔ-kɔtɔ	'crab'

b.	ɛ-boɔ	'stone'
	ɛ-dan	'building'
	ɛ-woɔ	'honey'
	e-tuo	'gun'

1.2 Loss of nominal prefixes

In Osam 1993 and 1994, the observation is made that one of the reasons for considering the noun class system in Akan as a decayed one is the loss of nominal prefixes. This loss may affect the singular only, or in some instances affects both the singular and the plural. In (4), the loss has affected only the singular nouns, but in (5) there is complete loss of the nominal prefixes in the singular as well as the plural.

(4)	SINGULAR	PLURAL	
	tɛtea	n-tɛtea	'ant'
	prako	m-prako	'pig'
	pataku	m-pataku	'hyena'
	bokiti	m-bokiti	'bucket'
	dadewa	n-dadewa	'nail'
	siw	e-siw	'hill'

(5)	SINGULAR	PLURAL	
	koraa	koraa	'calabash'
	kyɛw	kyɛw	'hat'
	sundzi	sundzi	'pillow'
	dwomba	dwomba	'pestle'
	wodur	wodur	'mortar'
	kuntu	kuntu	'blanket'

The complete loss of nominal prefixes as illustrated in (5) is another evidence that the grammar of Akan is sensitive to animacy distinctions. The observation here is that inanimate nouns are more likely to lose their nominal prefixes than animate nouns. There are definitely more inanimate nouns in the language without prefixes than animate nouns without them.

1.3 Double plural marking

As stated in Osam 1993:100, one of the characteristics of the decayed noun class system of Akan is that certain nouns have double plural marking. This is the process in which certain nouns mark their plurals by using a prefix and a suffix at the same time. When we examine those nouns which behave this way, we find that they are all human nouns. Non-human nouns do not undergo double plural marking. This process, therefore, serves to distinguish human nouns from non-human animate nouns as well as from inanimate nouns.

(6)	SINGULAR	PLURAL	
	ɔ-hen	a-hen-fo	'chiefs'
	ɔ-saman	n-saman-fo	'ghosts'
	o-banyin	m-banyin-fo	'men'
	o-panyin	m-panyin-fo	'elders'
	a-sew	n-sew-nom	'in-laws'
	o-nua	e-nua-nom	'siblings'

1.4 Numeral modifiers

Modification by numerals is not part of the nominal affixes but since it has to do with noun modification it is appropriate if it is discussed under this section. The behavior of numeral modifiers in Akan distinguishes human nouns from non-human ones. Before discussing this difference I give the numerals from 1 to 9 in Akan using the Fante dialect (7).

(7)	FANTE	ENGLISH
	kor	one
	ebien	two
	ebiasa	three
	anan	four
	enum	five
	esia	six
	esuon	seven
	awɔtwe	eight
	akron	nine

In all the dialects of Akan, when the numerals from 1 to 9 are used to modify human nouns, the form of the numeral is different from when they are used to modify non-humans. When these numerals modify non-human nouns, the forms are the same as given in (7). But when the modified noun is a human noun, the prefix *ba-* is attached to the numeral. This prefix derives from the noun *ba* 'child'. I should also point out that the vowel of the prefix will harmonize with the vowel of the root numeral. This distinction is illustrated in (8-10) with examples from the Fante dialect. As (9b) and (10b) show, it is wrong to put the prefix *ba-* on a numeral that modifies non-human nouns. On the other hand, when a numeral modifying a human noun does not have the *ba-* (8b), the result is only a questionable construction. It is possible to predict that over a period of time speakers may drop the *ba-* and regularize the form of the numeral modifiers.

(8) a. Nyimpa ba-anan
people child-four
Four people.

b. ?Nyimpa anan
people four
Four people.

(9) a. N-dua anan
CLASS 5-tree four
Four trees.

b. N-dua *ba-anan
CLASS 5-tree child-four
Four trees.

(10) a. M-bɔdɔ m anan
CLASS 5-dog four
Four dogs.

b. M-bɔdɔ m *ba-anan
CLASS 5-dog child-four
Four dogs.

It is also possible to use the human noun numeral modifier as the head of a noun phrase. This is illustrated in (11). These sentences are taken from the New Testament of the Bible. In this usage it is impossible to replace the (*ba*+numeral) with the plain numeral. So whereas in its function as the modifier the numeral could possibly occur without the *ba-* (see 8b), when it functions as the head of the noun phrase it is grammatically wrong to have a bare numeral. In (11), therefore, *ɔbaako* 'one (person)' and *baakron* 'nine (people)' cannot be replaced with *kor* 'one' and *akron* 'nine' respectively.

(11) a. Na ɔ-baa-ko no bua-a no dɛ
and CLASS 1-child-one the reply-COMPL 3SG OBJ COMP
"Ana i-nn-suro Nyankopɔn?" (Luke 23:40)
why 2SG SUBJ-NEG-fear God
And the (other) one replied "Don't you fear God?"

b. Na ba-akron no wɔ hen? (Luke 17:17)
and child-nine the be where
And where are the (other) nine (people)?

2. Pronouns

Another source of evidence for the animate-inanimate distinction is the pronominal system of the language. The evidence is based on the forms and the behavior of pronouns.

2.1 Subject pronouns

The subject pronouns in the language are distinguished on the basis of animacy. This distinction exists only in the Twi dialects, and it is relevant only in the 3SG subject pronoun. The 3SG subject prefix for animate nouns is *o-/ɔ-*, but that for inanimate nouns is *e-/ɛ-*. This is illustrated by examples (12) and (13).

(12) a. Abofra no bɛ-yera
child the FUT-be lost
The child will be lost.

b. ɔ-bɛ-yera
s/he-FUT-be lost
s/he will be lost

(13) a. Dua no bɛ-yera
tree the FUT-be lost
The tree will be lost.
b. ɛ-bɛ-yera
it-FUT-be lost
it will be lost

It is clear that these prefixes derive from the old noun class system (Osam 1993). The animate subject prefix is a reanalysis of the old noun Class 1 marker whereas the inanimate subject prefix derives from the old noun Class 4 prefix. In the Twi dialects, the distinction in animacy as reflected in the 3SG subject prefixes is strictly maintained. In Fante, however, this distinction is neutralized since the same pronominal form is used irrespective of the animacy status of the antecedent noun. The Fante equivalent of (12b) and (13b) would be (14). Note that Fante uses the lexical form *yew* 'be lost' in place of the Twi verb *yera* 'be lost'. So the verbs *yew* and *yera* 'be lost' are dialect variants.

(14) ɔ-bɛ-yew
s/he/it-FUT-be lost
s/he/it will be lost

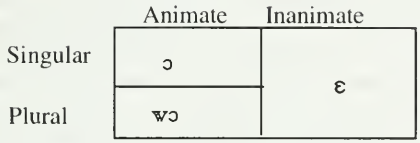
2.2 Lack of number distinctions

A further evidence of the grammatical differentiation between animate and inanimate nouns is that whereas anaphoric animate pronouns distinguish between singular and plural, inanimate pronouns do not make such distinctions. This distinction applies more in the case of Asante and Akuapem and related dialects than in Fante, even though as I will show below, in the speech of some Fante speakers this distinction is available. In (15a), the subject is a plural animate noun and in (15b) it is replaced by the anaphoric pronominal prefix, *wɔ* 'they'. In (16a), the subject is an inanimate plural noun but the pronominal prefix replacement in (16b) is the same form used for singular as illustrated earlier in (13b).

(15) a. Mbofra no bɛ-yera
children the FUT-be lost
The children will be lost.
b. Wɔ-bɛ-yera
3PLU-FUT-be lost
They will be lost.
(16) a. Ndua no bɛ-yera
trees the FUT-be lost
The trees will be lost.
b. ɛ-bɛ-yera
3PLU-FUT-be lost
They will be lost.

The relationship between the notion of animacy and subject prefixes can be summarized as in (17).

(17)



This lack of number distinction in the inanimate pronominal prefixes reflects the presence of animacy hierarchy in the language. In his discussion of animacy hierarchy, Comrie states the following: '... having distinct singular and plural forms is again a characteristic of noun phrases with high animacy ... another opposition that correlates closely with animacy is the existence versus non-existence of a number distinction, the split invariably being that noun phrases higher in animacy have the distinction while those lower in animacy do not.' (1981:180, 182) Even though the evidence for such a hierarchy in Akan is minimal compared to languages like Russian, Chukchi, Dyirbal and many others, we still have something that validates the argument that the notion of animacy hierarchy is relevant in the language.

2.3 Possessive pronouns

The distinction between animate and inanimate nouns is also demonstrated, in the Twi dialects, in the marking of certain possessive constructions. These are the constructions in which the possessed noun indicates some kind of relation, for example self, inside, bottom. In Akan and other languages, these are the nouns which are the sources of postpositions. In such constructions, when the possessor noun is animate, a full pronoun is used; but when it is inanimate we only get a pronominal prefix which incidentally is of the same form as the subject pronominal prefix. This difference is shown in (18) and (19).

(18) a. Kofi ho a-yɛ fi
Kofi self PERF-be dirty
Kofi is dirty.

b. Ne ho a-yɛ fi
3SG POSS self PERF-be dirty
He is dirty.

(19) a. Adaka no ho a-yɛ fi
box DEF self PERF-be dirty
The box is dirty.

b. ɛ-ho a-yɛ fi
it-self PERF-be dirty
It is dirty.

2.4 Behavior of 3SG object pronoun

Another evidence that the grammar of Akan is sensitive to animacy has to do with the behavior of the 3SG object pronoun. An aspect of Akan grammar that has been noted by various writers (including Christaller 1875, Stewart 1963, Boadi

1976, Lord 1982, Saah 1988, 1992) is that if the antecedent of the 3SG object pronoun is an inanimate noun, the pronoun is not overtly coded. Boadi 1976 refers to this as the "Pronoun-3-Object Deletion Rule". Examples (20) and (21) illustrate this phenomenon. In (20a) the direct object is an inanimate NP, whereas in (21a) it is an animate NP. If the inanimate NP of (20a) is replaced by a pronoun, even though the sentence is grammatical, semantically, it implies that the antecedent of the object pronoun is an animate entity, not an inanimate one. The only way (20b) can be formed to mean that the direct object is an inanimate entity is to have a zero pronoun in object position, as shown in (20c). Similarly, in (21c)where the object pronoun is covertly coded, the implication is that the direct object is an inanimate entity. In order to have the meaning that the direct object as an animate entity, the object pronoun has to be overtly coded as in (21b).

(20) a. Kofi bɔ-tɔn dua no
Kofi FUT-sell tree the
Kofi will sell the tree.

b. Kofi bɔ-tɔn no
Kofi FUT-sell 3SG
*Kofi will sell it.

c. Kofi bɔ-tɔn Ø
Kofi FUT-sell 3SG
Kofi will sell (it).

(21) a. Kofi bɔ-tɔn abofra no
Kofi FUT-sell child the
Kofi will sell the child.

b. Kofi bɔ-tɔn no
Kofi FUT-sell 3SG
Kofi will sell him/her.

c. Kofi bɔ-tɔn Ø
Kofi FUT-sell 3SG
*Kofi will sell him/her.

Even though Akanists are aware of this process, not enough has been done in terms of an explanation. One paper that has tried to deal with the issue is Saah 1992. This paper sets the rule in Akan within the framework of the Government and Binding Theory by treating it as an example of null object in Akan. Another paper that also attempts to explain the process is Boadi 1976. Boadi's approach is to determine the historical source of the phenomenon. His conclusion is that this phenomenon must have been borrowed into Akan from Ga, a neighboring Kwa language spoken in coastal Ghana.

Irrespective of the historical source of this phenomenon, it is my opinion that it has important functional implications in synchronic Akan. This view becomes relevant when we try to answer the question: Why is the object pronoun overt when its antecedent is an inanimate noun? The answer to this question is that it does so in order that the hierarchical ordering of animate and inanimate nouns is not subverted. It has to be understood that the form of the 3SG object pronoun is the same irrespective of the animacy status of the antecedent noun. Furthermore, as

mentioned above, since inanimate nouns do not make number distinctions in the pronoun, this same pronoun, *no*, is used for singular and plural antecedent inanimate nouns. The covert coding of the object pronoun when its antecedent is inanimate is built into the language to avoid the danger of hearers confusing an inanimate noun with an animate noun. It is a way of telling the difference between animate and inanimate nouns. In another sense the fact that animate nouns get replaced by pronouns but inanimate nouns are replaced by zero demonstrates that in a hierarchical ordering, animate nouns occupy a higher level than inanimate ones. It is for this reason that we can talk of animacy hierarchy in Akan.

Having discussed this process and the motivation for it, it is necessary to point out that there are two conditions under which this process is compromised. As it will be shown below, these exceptions can be functionally accounted for. One of the exceptions is that when the direct object in the sentence is followed by an adverbial element indicating time or location, the rule does not apply. In other words, for a third person pronoun whose antecedent is inanimate to be covert, as Boadi 1976 puts it, the inanimate direct object has to "occur utterance finally". This process is illustrated in (22). As shown in (22c), the presence of the adverbial *ɔkyena* 'tomorrow' requires that the object pronoun be overtly coded. That the overtness of the pronoun is conditioned by the presence of the adverbial element is supported by the fact that when the adverbial item is fronted in a focus construction as in (22d) so that the direct object is in utterance final position, the pronoun is not overt as expected. I should mention that when uttered without a context, the animacy status of the antecedent noun of the object pronoun in (22b) is not clear; it could refer to an animate or inanimate noun. Similarly in (22e), the sentence is ungrammatical if the utterance final pronoun has an inanimate noun as its antecedent.

(22) a. Kofi bɔ-tɔn dua no ɔkyena
Kofi FUT-sell tree the tomorrow
Kofi will sell the tree tomorrow.

b. Kofi bɔ-tɔn no ɔkyena
Kofi FUT-sell 3SG tomorrow
Kofi will sell it tomorrow.

c. *Kofi bɔ-tɔn Ø ɔkyena
Kofi FUT-sell 3SG tomorrow
Kofi will sell (it).

d. ɔkyena na Kofi bɔ-tɔn Ø
tomorrow FOC Kofi FUT-sell 3SG
It is tomorrow that Kofi will sell (it).

e. *ɔkyena na Kofi bɔ-tɔn no
tomorrow FOC Kofi FUT-sell 3SG
It is tomorrow that Kofi will sell (it).

We can offer a functional explanation as to why the presence of an adverbial element requires the inanimate object pronoun to be overtly coded. This explanation has to do, specifically, with the pragmatic notion of topicality. It has been established that, at the clausal level, the NP that codes the subject relation is more topi-

cal than any other entity in the clause; this is followed by the NP that codes the direct object relation (Givón 1984). In the Givonian functional framework, the subject is the "primary clausal topic" while the direct object is the "secondary clausal topic". Topicality hierarchy involving grammatical relations can be represented as follows:

Subject > Direct Object > Adverbial.

In Akan, the immediate postverbal position is one of the crucial defining characteristics of direct object. This means that an NP which bears the direct object relation necessarily has to occur immediately following the verb.

The reason the presence of an adverbial element in the post object position as in (22a) triggers the presence of the inanimate object pronoun is that since the direct object is more topical than an adverbial item, and since the immediate postverbal position defines direct objecthood in Akan, if the pronoun is not overtly present it would create the impression that the adverbial element is more topical than the direct object NP. It is as if the inanimate object pronoun finds its topicality status threatened and so it has to make a physical appearance in order to assert its status. With this explanation it is understandable why when there is no adverbial in sentence final position the pronoun is covert. Under that condition, there is no threat to its topical status.

The second condition which dictates the overt coding of the inanimate object pronoun is that there are a class of verbs which when used in the clause requires the presence of the pronoun. Example (23) illustrates this. In (23c) the absence of the inanimate object pronoun changes the meaning of the sentence. It should be noted that (23b) is ambiguous. If it is uttered in isolation from a context, the referent of the pronoun *no* could be either animate or inanimate.

(23) a. Kofi bɛ-hyew edziban no
Kofi FUT-burn food the
Kofi will burn the food.

b. Kofi bɛ-hyew no
Kofi FUT-burn 3SG
Kofi will burn it.

c. Kofi bɛ-hyew Ø
Kofi FUT-burn 3SG
*Kofi will burn (it).

d. Kofi bɛ-hyew
Kofi FUT-burn
Kofi will get burnt.

Other verbs in this class are: *sɛe* 'destroy', *bu* 'break', *hyew* 'burn', *kyea* 'bend', *tsen* 'straighten', *tsew* 'tear', *moa* 'crumple', *yew* 'lose', *koa* 'bend', *monkyem* 'crumple', *butuw* 'overturn', *bɔ* 'break'. One feature of these verbs is that they belong to the class of middle verbs, that is those verbs which normally take Theme direct object in transitive clauses, but also permit Theme subject in intransitive constructions. Why these verbs condition the overt coding of the inanimate object pronoun is not very clear to me at this stage. However, a possible reason may be

because they allow Theme subject in intransitive constructions. In such constructions, the subject entity is the one which undergoes the event indicated by the verb. Now, if the clause is supposed to have an object entity but this entity is not overtly coded, the only interpretation we can assign to such a clause is that it is the subject, the Theme, which undergoes the change in state. So in (23c), without the overt pronoun, the sentence cannot be interpreted as 'Kofi will burn it'. With no pronoun following the verb *hyew* 'burn', we have to interpret it as being used intransitively (23d). There is a pragmatic constraint on this analysis and it is that the analysis is legitimate only where the verb is capable of taking animate Theme subjects. For example, the verb *bɔ* 'break' cannot be given the same analysis because by its semantics animate entities cannot be its Theme. Another possible reason is that these are change of state verbs which have drastic effect on the state of the entities which undergo the change of state. Since the change is drastic a way has to be found out to show the entity that has been so affected. In terms of notion of transitivity, the extent of the affectedness of the Theme NPs of these verbs make clauses that involve these verbs very high in transitivity. This is because the extent of the affectedness of a Theme entity is one of the indicators of high transitivity (Hopper and Thompson 1980).

3. Conclusion

In this paper I have shown that the conceptual distinction between animate and inanimate nouns in Akan has reflections in the grammar of the language. As mentioned at the beginning of the paper, Akan is not one of the languages noted for having an animacy hierarchy. Nevertheless, there is an extent to which we can say that the animacy distinctions instantiated in the language form a basis for such a hierarchy. This comes out in the pronominal coding of animate and inanimate entities. I have shown that whereas animate entities have coding forms which reflect differences in number, the same is not true of the coding forms of inanimate nouns. Furthermore, the system of plural suffixation in addition to the regular prefixation, sets human animate nouns apart from nonhuman animate and inanimate entities. The strongest manifestation of an animacy hierarchy is in the behavior of the inanimate object pronoun. Based on the evidence, we can say that in Akan, the following hierarchy exists:

Human > Animate Nonhuman > Inanimate.

This hierarchy, as has been shown above, is manifested in various aspects of the grammar.

NOTES

* An earlier version of this paper under the title *Animacy distinctions in Akan* was presented at the 25th Annual Conference on African Linguistics, Rutgers University, March 25-27, 1994. I appreciate comments made on earlier drafts of this paper by Colette Craig, Florence Dolphyne, and Kofi Saah.

[1] Even though the nouns which occur with the prefix *o-/ɔ-* are predominantly animate entities, there are some inanimate entities that take this prefix. Examples include *ɔbotan* 'rock', *ɔman* 'country, nation', *ɔdan* 'house', *ɔtan*

'hatred'. There are other cases like *owu* 'death', *ohia* 'poverty', *Osaman* 'ghost' which are biologically not animate but are considered as such based on cultural beliefs.

REFERENCES

AKROFI, C. A. 1937. Twi kasa mmra [Twi grammar] (1965 edition). Accra: Waterville Publishing House.

BALMER, W. T., & F. C. F. Grant 1929. A grammar of the Fante-Akan language. London: Atlantis Press.

BOADI, L. 1976. A note on the historical antecedents of the obligatory pronoun -3- deletion rule in the Akan dialects. Acta Linguistica Hafniensia 16:1.1-10.

CHRISTALLER, J. 1875. A Grammar of the Asante and Fante language called Twi. Basel: Basel Evangelical Missionary Society.

COMRIE, Bernard. 1981. Language universals and linguistic typology. Chicago: University of Chicago Press.

DOLPHYNE, Florence. 1988. The Akan (Twi-Fante) language: Its sound systems and tonal structure. Accra: Ghana Universities Press.

——, & M. E. K. Dakubu. 1988. The Volta-Comoe languages. The languages of Ghana, ed. by M. E. K. Dakubu, 50-90. London: Kegan Paul International.

ESSILFIE, T. 1977. Serialisation in Fante. Leeds University Ph.D. dissertation in Linguistics.

GIVON, Talmy. 1984. Syntax — a functional-typological introduction. Amsterdam: John Benjamins.

HOPPER, Paul., & Sandra. Thompson. 1980. Transitivity in grammar and discourse. Language 56.251-299.

LORD, C. 1982. The development of object markers in serial verb languages. Studies in transitivity, ed. by P. Hopper & S. Thompson, 277-299. New York: Academic Press.

OSAM, Emmanuel. K. 1993. The loss of the noun class system in Akan. Acta Linguistica Hafniensia 26.81-106.

——. 1994. Aspects of Akan grammar — a functional perspective. University of Oregon Ph.D. dissertation in Linguistics.

SAAH, Kofi. K. 1992. Null object construction in Akan. MIT Working Papers in Linguistics 17.219-244.

——. 1988. WH- questions in Akan. Journal of West African Languages 18:1.17-28.

STEWART, J. M. 1963. Some restrictions on objects in Twi. Journal of African Languages 2.145-149.

UNITED BIBLE SOCIETIES 1982. AhyEmu Fofor [The New Testament]. Accra: The Bible Society of Ghana.

WELMERS, W. 1971. The typology of the Proto-Niger-Kordofanian noun class system. Papers in African Linguistics, ed. by C.-W. Kim, & H. Stahlke, 1-16. Edmonton: Linguistic Research, Inc.

——. 1973. African language structures. Berkeley: University of California Press.

Studies in the Linguistic Sciences
Volume 23, Number 2, Fall 1993 [publ. October 1996]

THE INTERPLAY BETWEEN TONE, STRESS, AND SYLLABIFICATION IN THAI[1]

Uthaiwan Wong-opasi
University of Illinois at Urbana-Champaign
and Syracuse University

In this study, tone is analyzed as an autosegment closely related to individual phonetic segments and suprasegments such as stress and syllable structure. Tone assignment in Thai is cyclic throughout word formation, observing two well-formedness conditions, namely—syllable structure and the one-tone-only association. The hypothesis that tone is an autosegment, separated from, but correlates with stress and syllable structure will be proven via a historical linguistic account of tonogenesis. Finally, the influence of Thai stress on tone distribution and syllable structure will be treated for the first time in the literature via an examination of language games involving foreign loanwords.

1. Background

Thai has FIVE lexical tones in the unmarked cases (cf. (1))

(1) Mid: khaa[2] 'a grass (imperata cylindrica)'
Low: khàa 'galangal, a rhizome' High: kháa 'to engage in trade'
Falling: khâa 'to kill; a servant' Rising: khǎa 'a leg'

However, not all five tones occur freely and various researchers, of both the segmental and autosegmental theories, have attempted to describe the tone occurrence restrictions. We will first outline some of the arguments supporting the segmental analyses (Henderson 1949; Leben 1971, 1973; Gandour 1974) which Yip 1980, 1982 rejected in favor of an autosegmental approach.

Leben argues for a segmental nature of Thai tones because tone may be dependent on vowel length. He points out one tone change phenomenon in Thai which is triggered by a vowel shortening rule. Assuming contour tones to be derived (Leben 1971; Gandour 1974), from the evidence that as a result of vowel shortening, the contour tone in the first element is neutralized to mid in the combinative speech style (cf. (2a)); while tone remains unaltered when vowel shortening does not apply (2b)[3], Leben concludes that there is a vowel shortening rule (VV->V) and a convention that simplifies HL or LH to a "compromised" mid tone, phonetically. Furthermore, in the cases of underlying level tone, e.g. a high level tone, no tone change results despite vowel shortening (2c).

(2) a.	thîi	nǎy	→	thi'	nǎy	'where?'
	sìi	khǎaw	→	si'	khǎaw	'white'
	sǎaw	sǎaw	→	saw	sǎaw	'young girls'
	wâaŋ	wâaŋ	→	waŋ	wâaŋ	'at your leisure'

b.	thâw	ray	→	thâw	ray	'how much?'
c.	náam	chaa	→	nám	chaa	'tea'

According to Yip, if tone is segmental, when a vowel deletes, tone should be deleted as well (LH->L or H) rather than a compromised mid tone (*M). She also cites Gandour's observation that elsewhere in the language, LH is converted to H, as in *chăn →chán* 'I'; *phŏm →phóm* 'I, male', and *khăw →kháw* (third person pronoun). Moreover, upon a close examination of Henderson's data, Yip found examples of HL tone on short vowels followed by a glottal stop or zero in sentence intonation such as *bâ(ʔ), lâ(ʔ)*. Furthermore, Yip cites counterexamples which include cases of tone change accompanying vowel shortening with underlying LEVEL tones.

(3)	yàaŋ	rai	→	yaŋ	ŋai 'how?'
	yàaŋ	níi	→	yaŋ	ŋíi 'this way, like this'

Due to the gaps in Henderson's data for the discussion in her article, Yip cannot offer a firm conclusion but suggests that the tone change in (3) may be morphological rather than phonological or it may be stress conditioned (not by complete absence of stress, like the neutral tone, but rather by secondary as opposed to main stress). This view is supported by Henderson's discussion of the neutral tone in unstressed syllables (1949:37, fn. 27)

> The actual pitch of the neutral tone may vary according to context, but is most commonly mid level.

Also, Henderson 1967 affirms that the stress in forms like (2a), such as *thîi năy →thi 'năy*, falls on the second syllable.

At this point, we would like to suggest treating the vowel shortening and tone change in (2) - (3) as postlexical phenomena which are caused by the stress pattern in Thai. Specifically, Thai is stress-final within a given phonological domain, i. e. a word, a compound, a phrase, or a sentence. Although the two processes usually cooccur as they usually take place in unstressed syllables, they are independent from each other as we may find tone change without vowel shortening (4a) and vice versa (4b).

(4) a.	nă̆ŋ	s~~ŭu~~		→	naŋ	s~~ŭu~~		'book'
	thá?	lee		→	tha	lee		'sea'
	má?	rá?		→	ma	rá?		'bitter melon'
	krà?	còk		→	kra	còk		'mirror'
	thoo	rá?	thát	→	thoo	ra	thát	'television'
	kaa	lá?	mɛɛ	→	kaa	la	mɛɛ	'a kind of sweet'
	pray	sà?	nii	→	pray	sa	nii	'post office'
b.	taa	plaa		→	ta'	plaa		'callous(lit. eye+fish)'
	pàak	kaa		→	pàk	kaa		'pen(lit.mouth+crow)'
	yàaŋ	nán		→	yàaŋ	ŋán		'that way, like that'
	yàaŋ	níi		→	yàŋ	ŋíi		'this way, like this'

The last example, *yàaŋ núu → yàŋ ŋúu,* is a variant of *yàaŋ núu → yaŋ ŋúu* according to Henderson (cf. (3) above). It is given as a counterexample in Yip 1982.

The preceding discussion of tone neutralization is disapproved of by Gandour who claims that acoustic results showed no evidence of such neutralization (though he eventually admits it in Gandour 1979:140). However, native speakers of Thai (Warotamasikhadit 1967; Surintramont 1973) can hear the tone change which approximates the mid tone, although not exactly identical to underlying mid tone, and the shortened vowels which are not exactly equivalent to underlying short vowels, either. Given the nature of these postlexical rules which can be stylistic variants, and therefore optional, the inconclusiveness of the data is expected. Therefore, it may be safer to consider, as a point of departure, the nature of Thai tones in isolative speech. That is, we should concentrate on the tone on each syllable when pronounced in isolation or emphatically since these tones are truly lexical tones which occur at the lexical level in Lexical Phonology's terms.

We turn now to isolative speech style. Gandour 1974:138 proposes that the tone domain for each tone is a sonorant segment in syllable rime. That is, a syllable with $CV(V)C_s$ structure, where C_s = a sonorant coda (traditionally called LIVE syllable), can have all 5 tones, while syllables checked with a stop (henceforth, CHECKED or DEAD syllables) have the tone distribution, formulated in Yip 1982:89, as follows.

(5)						
	CVC	H	*HL	*LH	L	*M
	CVVC	*H	HL	*LH	L	*M

A segmental analysis like Gandour's clearly explains the absence of LH and HL on short CVC, and the presence of HL on CVVC. However, he attributes the absence of LH on CVVC as due to universal performance reasons. Yip, on the other hand, contests that HL on short CVC, and CVVC with high level tone are found in words other than loanwords and onomatopoeia (although they are less common), and that LH on CVVC is found in other languages.

(6) kháat 'card' (Eng. loanword) but also pɔ́ɔt 'a woman's nickname'
khlâk 'crowded, tightly packed'

The only real gaps, according to Yip, are LH and M. (She also notes that the few apparent examples of LH are all intonationally derivable.) It must be noted that neither Gandour nor Yip can account for *M. Regarding the absence of LH, Yip proposed a universal condition that prohibits the configuration *LHʔ which applies to Thai, Zahao, and Cantonese. She claims a laryngeal feature such as syllable-final glottalization as the (p, t, k, ʔ) stops in Thai involve a glottal closure. Initially, these stops and (c) are glottalized. According to Yip, a segmental analysis cannot account for the existence of words like *khlâk* which bear two tones on one vowel. Neither can the segmental approach describe the lack of three tone sequences in forms with three sonorant segments. To circumvent this fact, Gandour posits a requirement that either the first and second segments be the same, or the second and third. In an autosegmental solution, however, the restriction is stated with less complications—only two tones are permitted per syllable.

Another argument for a segmental account given by Leben (1973:34) is the fact that tones move with their vowels. Evidence is drawn from a Thai word game where rimes interchange between syllables.

(7) kôn yày → kày yôn 'big bottom'

Yip (1980:11), however, argues for floating tones which remain stable despite segmental change, i.e. transposition of rimes, in another Thai word game (following Gandour's observation).

(8)	klûay	hɔ̌ɔm	→	klɔ̂ɔm	hŭay	'banana'
	tên	ram	→	tâm	ren	'dance'

A segmental suggestion will fail to account for the lack of tone movement. An autosegmental proposal, on the other hand, can account for both phenomena. In (7), rime movement occurs before the mapping of the suprasegmental tones, whereas (8) shows segmental change after tone association.

Although two distinct rule orderings can give rise to two different language games (as in (7) and (8)), upon a scrutiny of more data, we find tone patterns that cannot be accounted for by Leben's or Yip's proposal since they show different tone patterns from those existing in the base forms.

(9)	chí	khaa	kôo	->	a.	ʔchôo	khaa	kí	'Chicago'
	H	M	HL			*HL	M	H	(rime change, then tone mapping)
					b.	*chóo	khaa	kî	
						H	M	*HL	(tone mapping, then rime change)
					c.	choo	khaa	kì	
						M	M	L	(resulting tones not in base form)
	khaa	buu	kì	->	a.	khì	buu	kaa	'Kabuki'
	M	M	L			L	M	M	(rime change, then tone mapping)
					b.	*khi	buu	kàa	
						*M	M	*L	(tone mapping, then rime change)
					c.	khí	buu	kàʔ	
						H	M	L	(resulting tones not in base form)

Following Burzio 1991, who proposes a well-formedness condition for English in cases where the ordering of certain phonological rules cannot yield a correct result, the situation in Thai here can be explained by resorting to tone well-formedness conditions (cf. details in section V below). The proposed well-formedness conditions on tone advocated in the present study, in turn, find support in tonogenesis and syllable structure (cf. sect. III).

2. Tonal onsets

Tumtavitikul 1991 argues for a combined segmental and suprasegmental approach. She demonstrates the inadequacy of positing floating tones alone by proving the fact that certain Thai consonants, in fact, carry underlying tones.[4] This is illustrated by suffixed examples such as *leekh + aa →leekhǎa* 'secretary' vs. *rookh + aa →rookhaa.* 'disease, illness'. The two forms have identical syllable structures in the base forms, i.e. a long syllable closed by an aspirated stop, plus an *aa* suffix. The two forms also exhibit tonal differences when suffixed by a short vowel plus a stop consonant, e.g. *leekh + aʔ→leekhàʔ* vs. *rookh + aʔ→rookháʔ.* According to her, even though we may posit a floating H and L tone on the two forms, respectively, at the right edge of the base morpheme, or at the left edge of the suffixes, we still cannot account for such variations. Thus, she spells out tone assignment rules in (10) (present author's emphasis).

(10) a. For ALL voiceless fricative onsets and SOME voiceless aspirated stop onsets, Rising tone surfaces on unchecked syllables.
b. Elsewhere, unchecked syllables surface with Mid tone.
c. For all resonant onsets and SOME voiceless aspirated stop onsets, short checked syllables surface with High tone.
d. Elsewhere, checked syllables surface with Low tone.
e. Rising tone in (a) is derived.

The problems which arise in the above analysis are that various issues are left out unanswered. First of all, Tumtavitikul does not account for the fact that the long and short versions of the suffix, i.e. aa vs. aʔ, are merely the Thai phonological variants of the Indic short nominal suffix -a. Secondly, the tonal rules in (10) seem to be arbitrary. Moreover, there are no explicit ways to distinguish the two types of voiceless aspirated stop onsets in (10a) and (10c) or two distinct classes of voiceless fricative onsets. Failure to define these distinctions results in arriving at wrong predictions. For example, we would not have a monosyllabic minimal pair such as *fǎn* 'to dream, a dream' vs. *fan* 'tooth'; or a tonal contrast on the second syllable in multisyllabic words such as *chàʔ nǔan* 'an electric insulator' vs. *cháʔ nuan* 'a slate for writing on, a fuse', and *chàʔ lǐaŋ* 'a verandah, a balcony' vs. *càʔ riaŋ* 'woman's name'. Finally, the application of too many complicated phonological rules which interact with each other does not facilitate language acquisition.

Alternatively, an analysis which is based on phonetic-phonological correlations and tonogenesis as well as the morphology-phonology interface, utilizing well-formedness conditions on syllable structure and tone at each stage of word formation will be argued for to which we now turn.

3. Tonogenesis and phonetic-phonology correlations

It could well be argued for that each Thai word is morphologically marked for tone in the lexicon. However, regarding loanwords from nontonal languages such as Pali, Sanskrit, English, etc., how the borrowed words get tones is of par-

ticular interest. An examination of the Indic and Thai consonantal systems reveals the following facts :

i) There was a loss in the voicing contrast so that both the Indic voiceless aspirated stops and voiced unaspirated stops were neutralized to voiceless aspirated stops, i.e. kh, g →kh; ch, j→ch; ṭh, ḍ→th; th, d→th; ph, b→ph. The original distinction between the two classes of consonants, however, is preserved in the resulting distinct tonal consonant classes (cf. appendix and details below).

ii) Voiced aspirated stops were lost in Thai pronunciation, i.e. gh, jh, ḍh, dh, bh→ Ø.

iii) The asymmetrical distribution of Thai voiced unaspirated stops, i.e. *g but d, b can be explained as due to a later Thai creation and thus, was not involved in tonogenesis.

iv) Both the retroflex and the dental consonants are neutralized to dentals, t̪, t → t; ṭh, th → th; ḍ, d → d (→ th); ḍh, dh → dh (→th); ṇ, n → n.

v) Moreover, laboratory experiments show that tone change can be induced by onsets:

> ....voiceless oral obstruents produce high tone (or a higher variant of a tone) on the following vowel, whereas voiced oral obstruents produce low tone (or a lower variant) on the following vowel. (Haudricourt 1961, Cheng 1973, quoted in Ohala 1978:25)
>
> ....high air flow after voiceless, especially voiceless aspirated, obstruents, and low air flow after voiced obstruents caused the high and low pitches, respectively. (Ohala 1973, and Ohala 1978:26)[5]
>
> The available data suggest that pitch following voiced stops is substantially similar to that following sonorants and that it is the pitch following voiceless stops that is perturbed upwards. (Lea 1972, Hombert 1975, Jeel 1975, quoted in Ohala 1978:29)

The lost voicing and aspiration contrasts are preserved in the different tones assigned to the consonants in question. Thus, during the evolution from a voiced unaspirated stop to a devoiced aspirated counterpart, two features were encoded as one extrapolated tone, namely, [-voice] → H and [+spread] → H, equivalently HH→ H. To distinguish original voiceless aspirated stops from devoiced aspirated stops which were already assigned a high tone (in addition to unmarked [-vc., -sprd] onsets → default M), the single feature [+spread] of the original voiceless aspirated stops must take on a low tone as they are a lower variant between the two types of consonants. This is as outlined in (11) and (12), respectively.

(11) [-voice] [+spread] → HH → H

\ /

O

That is, high tone in Thai onsets was caused by the devoicing and the aspiration of voiced unaspirates (e.g. g → kh). This also applies to sonorants since sonorants pattern like original voiced stops in tone assignment (cf. point (v) above), and because of the loss of voicing contrast in voiced and voiceless sonorants in Thai.

(12) [+spread] → L
|
O

(12) should include also all aspirated voiceless segments, e.g. [kh], sibilants of Indic origin [s̩, s̈, s → s], and segmentless voiceless glottal fricative [h].

Contrary to Hombert 1978, Hock (1985:98) states that "Both onset and post-nucleus consonants can induce tone, but onset consonants have a greater effect." The Thai data confirms Hock's claim in unmarked cases. For marked cases such as in checked syllables, we postulate that the glottalized feature of final stops (p, t, k, ʔ) in Thai has an effect on tone because these stops involve a glottal closure. Syllable-initially, these stops and (c) are glottalized while at the end of a syllable, they are unreleased stops, and therefore can increase syllable weight which in turn, can cause a low tone.[6] The formulation is as follows.

(13) [-continuant] → L
|
R

4. Objectives

This paper seeks to answer the following questions via the use of phonetic-phonological explanations in the Feature Geometry Theory (Clements 1985; Sagey 1986; Bao 1990; Duanmu 1990), the Autosegmental approach, and Burzio's 1991 Well-formedness Condition as an amendment to previous research.

i) How is tone influenced by syllable structure? Specifically, why only heavy (or checked) syllables have tone occurrence restrictions as in (5), reproduced in (14).

(14) CVC H *HL *LH L *M
CVVC *H HL *LH L *M

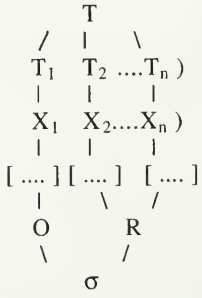
In addition to Yip's counterexamples to *HL on CVC and *H on CVVC, explanations will be given as to why these occurrences are marginal. Moreover, Yip's universal prohibition of the *LHʔ configuration will be explained. Finally, *M will also be accounted for with the same principle.

ii) Why is a modular approach to Thai tonology preferred? This is because within Feature Geometry Theory, the segmental analysis can be restated in autosegmental terms. With each phonetic feature constituting a separate tier from, but interactive with, the segmental and the tonal tiers, we can achieve an optimal analysis of tone assignment and tone change in the formation of native Thai words along with foreign loans, language games, and sentential tonology. That is, Thai tone must be assigned and reassigned according to the syllable structure of Thai words at every stage of derivation, from isolated lexical words to units larger than words or postlexical units.

iii) How can stress interact with tone assignment? Tone assignment and reassignment will be seen as an adaptation of Thai lexical and prosodic stress to accommodate different stress patterns of the donor languages.

5. Well-formedness condition, syllable structure and tone assignment

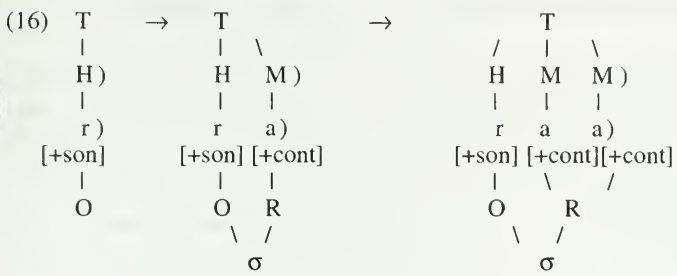
In this analysis, Thai syllables are proposed to be of two types:—LIGHT vs. HEAVY syllables—with the former containing [+continuant] rime, and the latter, [-continuant]. All short vowels are treated as followed by a glottal stop and long vowels are represented as sequences of two identical vowels. It will be argued for here that the tone bearing units (TBU's) in Thai can be any segment in the syllable. Moreover, unlike previous proposals which utilize a noncyclic approach, this study will show that tone assignment in Thai is cyclic both at the syllabic and the morphological levels (cf. (15) and section 6). That is, tone is first assigned according to the phonetic properties of the onset of a syllable which constitutes the first tonal domain, represented here by a right bracket. Unspecified onsets induce mid tone which is also an unspecified or default tone. Devoiced aspirated onsets, on the other hand, raise the tone of the syllable to H (rule (11)) whereas original voiceless aspirated onsets are assigned L tone (rule (12)). Upon incorporation of each new segment in the syllable, a subsequent tonal domain is created (i.e. from tonal domain 1, 2... to T_n), while previously assigned tones are erased and a new tone is assigned in accordance with the [+/- continuant] feature of the new rime segment. That is, in marked cases, a rime-final [-cont] segment induces L tone (rule (13)) while M is the default tone for unmarked [+cont] rime. The generation and regeneration of tone is seen in (15) where the association line is truncated and readjoined between the tonal tier (T's) and the segmental tier (X's) when each element is integrated in the syllable, from onset (O) to rime (R), thereby creating a gradually larger tonal domain. We show the tone assignment of a Thai native word *raa* 'mold' in (16) (next page). [r] is first assigned a H tone. Then, [a] is integrated with a new tone M assigned because [a] is [+cont]. The preceding L tone is thus dissociated. When the final segment is incorporated into the syllable, the second tone is replaced by the final tone.

(15) Tone Assignment: T → 

Hence, for light syllables which have unmarked syllable structures, the default tone is mid unless specified otherwise in the lexicon. Consequently, the unmarked syllable structure allows its association with any one of the five lexical tones, the choice of which is lexically governed, and is reflected in the use of an explicit tone marker in the language. Heavy syllables, on the contrary, are subject to a well-formedness condition in Thai which prohibits mid tone (cf. (17a)). This is due to the fact that unlike sonorant codas which pattern like vowels, stop codas

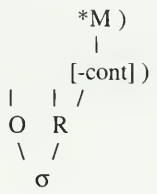
constitute heavy syllables and for the purpose of tone assignment, this heavy weight exerts itself in a low tone. Another well-formedness condition (17b) prohibits that no Thai segments contain more than one tone.

(16)



(17) Well-formedness Conditions:

a. Prohibition of Mid Tone:



b. Tone Restriction:only ONE tone can be associated to a segment : *T T)

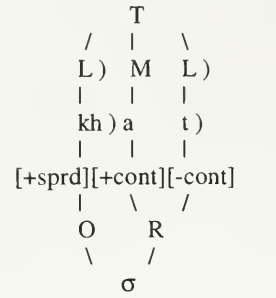
The fact that contour tones exist in Thai syllables is due to the preservation of previously assigned tone when the onset contains either the dual [-voice] and [+spread] feature or [+spread] alone. That is, the bracket erasure convention does not apply in these cases and therefore, two tones are assigned, one to the onset, and the other eventually to the coda. For low tone, since both the onset and the final segment bear non-distinct low tones, the two low tones are fused into one— LL → L (cf. (18a) on the next page). For rising tone, there is an additional contour tone exaggeration effect whereby the LM tone sequence is enhanced to LH for maximal audibility (cf. (18b) on the next page). Thus, the cyclic tone assignment (15) and the well-formedness condition (17b) account for contour tones in short and long syllables.

We turn now to the high tone in CVC and the falling tone in CVVC structures. High tone is first assigned to a devoiced aspirated onset. Because of the highly marked feature [-voice, +spread], the already assigned high tone is kept while integration of the remaining segments of the syllable generates an additional low tone in agreement with the stop coda. The outcome is HL but due to the shortness of the heavy syllable, the final tone is not integrated, except in very few intonationally derivable lexical words. However, in long heavy syllables, either option is available. With the preservation of the final low tone, the syllable contains a falling tone as found in most Indic loanwords. In most English loans, in contrast,

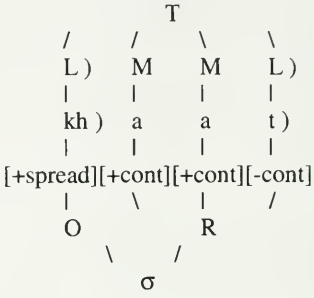
in response to the need to distinguish English words from already established Thai words, including Indic loans, the final low tone is simply dropped.

(18) a. Low Tone in CV(V)C Structures: Tone Conflation: LL → L.

i. khàt 'scrub'

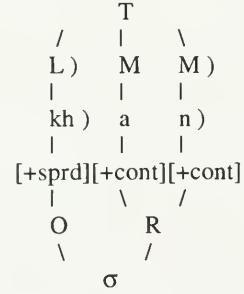


ii. khàat 'broken'

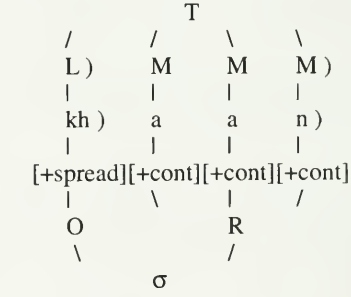


b. Rising Tone in CV(V)C Structures: Contour Exaggeration: LM → LH.

i. khăn 'bowl or basin used for dipping water'

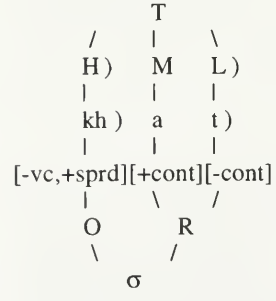


ii. khăan 'to call in reply'



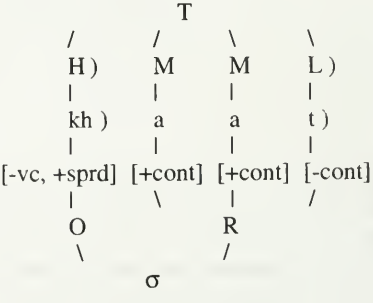
(19) a. High Tone:

i. khát 'select'



b. Falling Tone:

ii. khâat 'to gird'



6. Evidence

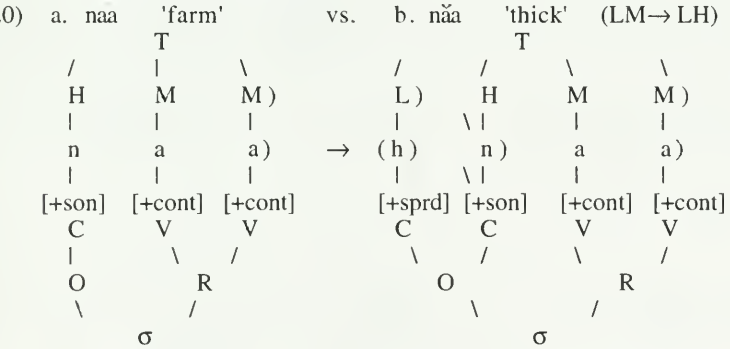
Tone forms an autosegmental tier as it can be separated from the segmental tier. However, the tone autosegment is linked to individual segmental features as in the Feature Geometry Theory. Evidence is found both on phonological and

morphological grounds. We will first look at three types of cluster onsets:a) clusters with a segmentless tonal onset [h]; b) true cluster onsets; and c) unparseable cluster onsets.

6.1 Phonological evidence

Thai has one segmentless tonal onset which is a voiceless glottal fricative [h] with the feature [+spread]. Like all other aspirates which are [+spread], it bears a L tone. If tone were segmental, it would be impossible to convert the high tone to low in short heavy syllables with a sonorant onset just by adding this segmentless [h] to the onset, e.g. *mát* (H) 'to tie' vs. *hmat* (L) *[màt]* 'fist' because this segment does not manifest itself phonetically. A rising tone obtained by adding this segmentless onset to a sonorant within the initial cluster would be impossible, too, e.g. *naa* (M) 'farm' vs. *hnaa* (LH) *[nǎa]* 'thick'.[7] (20) shows the elimination of the stranded consonant [h] from an unpronounceable Thai cluster [hn]. Despite the suppression of [h] from the segmental tier, its tone is not deleted and has the spreading effect onto the following onset. The well-formedness condition (17b) ensures that only one tone is assigned to the sonorant after spreading; thus, the tonal feature [H] of the sonorant must be dropped

(20) a. Segmentless Tonal Onset:

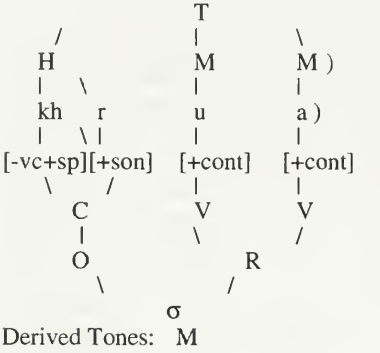


b. True cluster onsets: True cluster onsets are those which share the same C-slot underlyingly and therefore, do not permit vowel insertion to break the cluster. As such, only the first member in the cluster can attract tone while no tone can be assigned to the second member. In (21a) (next page), high tone is assigned to [kh] according to the features [-voice, +spread]. Because [r] is not provided with an independent C-slot, it cannot bear tone but receives the high tone from its partner through spreading. Likewise, the feature [+spread] triggers low tone in (21b) (next page) and spreads onto [r]. If tone were segmental alone, there would be no way to explain the single tone assigned to true cluster onsets.

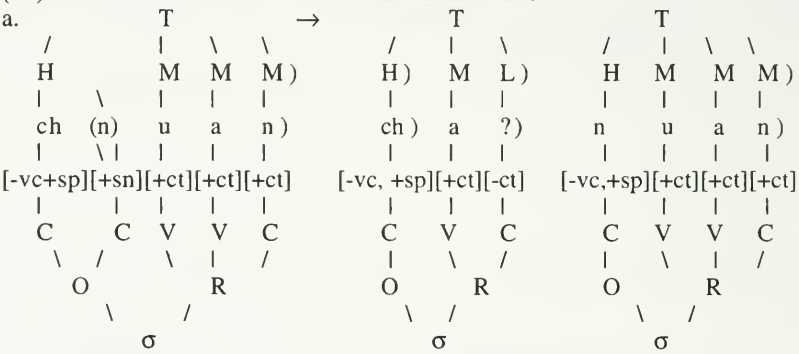
c. Unparseable cluster onsets: For consonants that cannot form an initial cluster in Thai, then, either the second member is dropped, e.g. sr → s; cr → c, or a default vowel [aʔ] is inserted to break the cluster to accommodate the Thai pronunciation, e.g. *chn* → *chàʔ-n*. Before vowel epenthesis, however, the features of the first member of the cluster are collocated onto the second member,

i.e. [-voice, +spread] of [ch] on [n] in (22a) and [+spread] in (22b), respectively. Thus, H and L tones are assigned to (22a) and (22b), in that order.

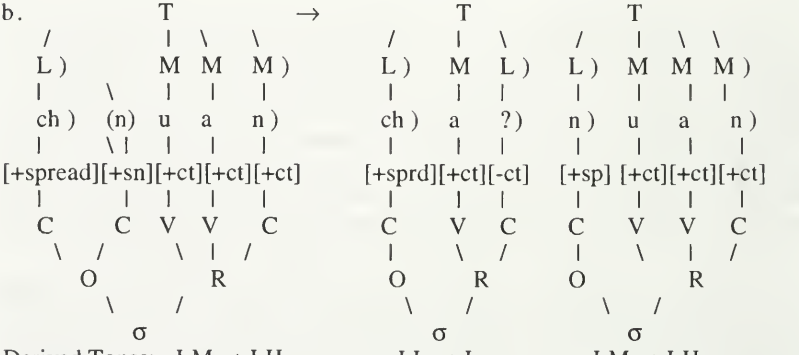
(21) a. khrua 'kitchen' vs. b. khrŭa[8] 'aged Buddhist priest'



(22) a. chnuan → chá? nuan 'fuse'



b. chnuan → chà? nŭan 'insulator'

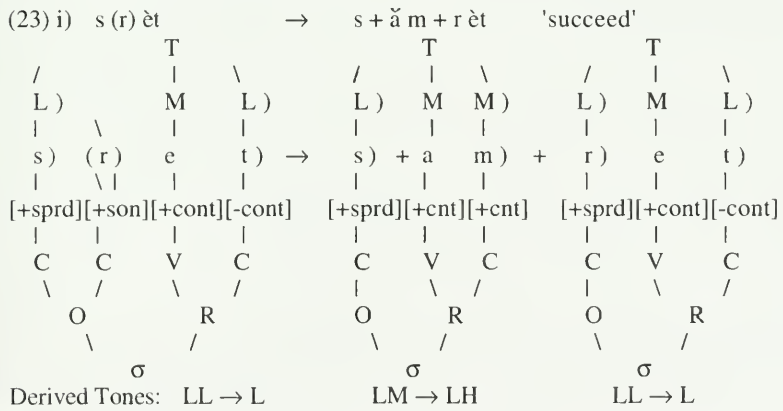


6.2 Morphological evidence

Tonal consonants clearly manifest themselves in derived words through the infixation of *am (n)* (23), lexical derivation (24-26), and language games (32).

6.2.1 am (n) infixation

A limited number of monosyllabic words become bisyllabic via *am (n)* infixation. Such an infixation is no longer operative in Thai and its idiosyncrasies must be recorded in the lexicon since the base for the *am (n)* infixation can be of any lexical category and the derived form usually maintains a synonymous or related meaning to that of the base form while the lexical category of the base may or may not be altered in the derived form. Examples are: *s(r)èt* 'to finish'→ *s+ǎm+ rèt* 'to succeed'; *t(r)ùat* 'to check'→ *t+am+ rùat* 'police'; *dəən* 'to walk'→ *d+am+ nəən* 'to proceed; (Royal vocabulary) to walk'; *khróp* 'to complete'→ *k+am+ róp* 'time (as in "cycle")'; *pràap* 'to get rid of'→ *b+am+ràap* 'to subdue'; *pruŋ* 'to improve'→*b+am+ruŋ* 'to support'; and *tràt* 'to speak (Royal vocabulary)' → *d+am+ràt* 'to speak (Royal vocabulary)'. For the first example, the cluster [sr] does not form true cluster onsets in Thai since they each take a separate C-slot. Because the [r] is deleted from the cluster, it allows the spreading of [+spread] from the first member of the cluster onto the second member before *am(n)* infixation. The tone in the infixed form is thus as predicted. (The changes p→b and t→d are explained in section III pt. (iii) above.)



Only through an autosegmental analysis which permits spreading can we arrive at a consistent explanation for retaining the tone of the base forms in the second syllable of the derived words while the first syllable acquires a new tone according to its phonological constructs. It also allows for the lack of tone sandhi in forms such as *sǐaŋ* → *s+ǎm+niaŋ* 'voice' and *sr̥ǔam* → *s+ǎm+ruam* 'laugh'. In the latter cases, we can claim that *am(n)* infixation applies first and therefore, bleeds spreading. Finally, for words that are clearly marked with a lexical tone via the use of a tone marker such as *càay* 'pay' → *c+am+ nàay* 'distribute', it is evident that since the lexical tone is not part of the syllabic structure, the lexical

tone must then be retained in the second syllable, leaving the first syllable free to obtain tone as conditioned by its own syllable structure.

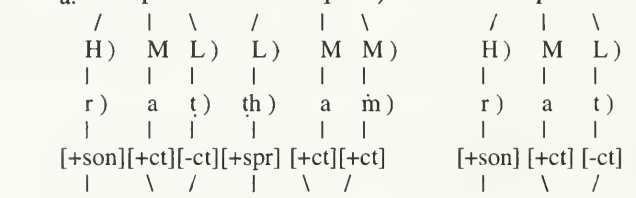
6.2.2 Derived lexical items

Most Indic loanwords suffered the following morpho-phonological changes: i) final vowel (plus nasal) truncation[9] (followed by cluster simplification, when applicable, and final stop neutralization to unreleased p, t, k stops; or l, r → n); ii) otherwise, the final vowel is retained, either as a long vowel or as a short vowel accompanied by a glottal stop, primarily in suffixed forms and compounds.

In general, the Indic -a suffix is truncated in Thai. The following examples are from Pali (unless indicated otherwise) although the same set of rules applies also to Sanskrit:*kul+a* (→ *kul*) → *kun* 'family'. (The intermediate stage of derivation is given in parentheses.) The distinct tone patterns are regulated by Thai syllable structure. Specifically, tone evolved from the loss of certain Indic phonetic properties of the onset and the new rime structures in the Thai language as discussed above, e.g. *raṭṭh+aṁ*→ *rát* 'country'; *raaṣṭr+a* (Skt.)→ *râat* 'country'; *śaastr+a* → *sàat* 'science/art'; *narak+a* → *na rók* 'the great hell'; *jan+a* → *chon* 'the public; people', the last two examples exhibit an additional vowel change [a→o] word-finally in most Indic loans. (24) shows the effect of syllable structure on tone in simple loanwords with final vowel truncation.

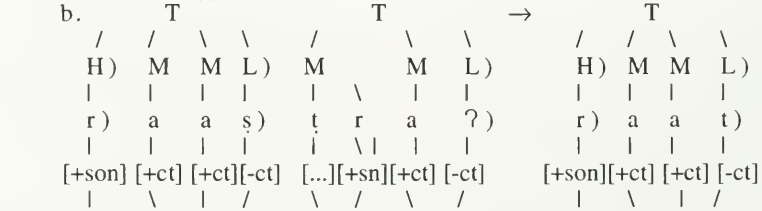
i) Final Vowel Truncation:

(24) a. raṭṭhaṁ[10] (Pali) → rát 'country'



Derived Tones: HL→H and LM → LH vs. HL→H

b. raaṣtra (Skt.) → râat 'country; citizen'



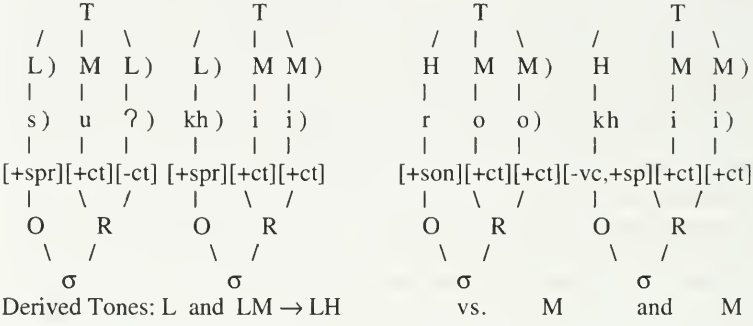
Derived Tones: HL and L vs. HL

The preservation of the suffix -a is less common and requires various adjustments. Despite the fact that Thai does not have comparable stress to the type of stress found in stress languages like English, Pali, and Sanskrit, Thai stress can be detected solely via its influences on stress-related phenomena such as tone and vowel length. Stated differently, given that the tone and the vowel length of each word in Thai are lexical as they constitute semantic differences, these tone and vowel length contrasts must be encoded in the lexicon. In addition, each lexical word has a potential to bear stress in a slow, emphatic speech. The capability to bear tone, vowel length contrast, and stress applies both to mono- and poly-syllabic words, in the optimal cases. By implication, this means that each syllable in poly-syllabic words can bear the tonal, vocalic, and stress distinctions as well. However, in normal speech, not all syllables or words are pronounced with equal force or stress as regulated by the prosodic rhythmic patterns. We postulate that within a phonological domain, Thai primary stress falls on the last syllable, and the secondary stresses alternate with unstressed syllables. When a non-final syllable is destressed, it can cause vowel shortening and/or mid-tone neutralization as seen in (4a) above. The shortening of the short vowel [a] which, in Thai, must be accompanied by a glottal stop, is carried out by the deletion of the glottal stop. Our hypothesis that Thai is stress-final within a given phonological domain is corroborated by the fact that a glottal stop can be deleted in non-final, unstressed position, as in *má? rá?* → *ma rá?* (with an additional mid-tone neutralization on the unstressed syllable) but it can never be dropped word-finally as shown in the ill-formedness of *má? rá?* → **ma? ra* or **ma ra*. In effect, stress in Thai falls on the final position of a phonological domain, i.e. on the final syllable of a word, at the lexical level where lexical words are being constructed, or on the last syllable of a phrase at the sentential or postlexical level. This is due to the fact that the final syllable of the stress domain in Thai is heavy by position. The final stress in Thai overrides the stress patterns in the donor languages, which for Pali and Sanskrit may fall either on the stem or the suffix; and for English, either on the final or non-final syllable. As such, the domain-final syllable in Thai must contain at least two moras or a two-segment rime, so that the final syllable can bear stress. To render the short Indic suffix with a plain short vowel in Thai would violate the syllable-structure requirement. Consequently, the Indic short vowel must be lengthened or a glottal stop must be added to preserve the heavy weight assigned to the stressed position. This effect is attested in many doublets in Thai. For example:*raaj+a* → *raa ch+aa / raa ch+á?* (or *râat*) 'king'; *kaay+a* → *kaa y+aa / kaa y+á?* (or *kaay)* 'body'; *geh+a* → *khee h+ăa / khee h+à?* 'house'. The choice between the final vowel truncation, its lengthening, and the glottal stop insertion is arbitrary and thus, must be marked in the lexicon. In certain instances, the resultant distinct forms are all employed to denote different lexical meanings, e.g. *likh+a* (cf. *likh+ati)* 'to write' → *lee kh+ăa* 'secretary (=one who writes)' vs. *lêek* 'arithmetic (=written numbers)'; *sukh+a* 'happiness' → *su kh+ăa* 'toilet (=happy place)' vs. *sùk* 'happiness'. Word-final glottal stop is retained only in emphatic speech or in Indicized contexts, e.g. *su kha?* in the Buddhist blessing, *?aa yú wan ná? su khà? phá? lá?* 'longevity, growth, happiness, and health'.

The effect of suffixal vowel lengthening is illustrated in (25).

ii) Indic Suffixation and Vowel Lengthening:

(25) a.leekh + aa (Skt.) 'secretary' vs. raaj+ aa (Skt.) 'king'
b.sukh + ii (P) '(the) happy, blessed (one)' vs. rookh + ii (P) '(the) sick'



The distinct tonal patterns caused by different syllable structures are most evident in the Thai reading of original Indic suffixed forms, e.g. Pali:*khattiya* → *khàt. tì? yá?* (m.) 'man of the warrior caste' vs. *khattiyaa* → *khàt. tì? yaa*, *khattiyaanii* → *khàt. tì? yaa nii* (f.) 'woman of the Khattiya clan'; *naaga* → *naa khá?* (m.) 'cobra, elephant; iron-wood tree; noble person' vs. *naagii* → *naa khii*, *naaginii* → *naa khí?* nii (f.) 'female cobra, female elephant; noble woman'.

Our postulation of syllable weight received from word-final stress is further corroborated by the fact that when simple words are compounded, neither an epenthetic long vowel nor a glottal stop are needed since the final syllable of the first member of the compounds is no longer stressed as the new stress domain is the last syllable of the entire compound. As a result, the tone of the unstressed syllable is neutralized to mid as observed by previous researchers (cf. (2-4) above). In emphatic speech, however, where all syllables are stressed, the glottal stop is retained and triggers an appropriate tone assignment according to the new syllable structure, e.g. *jan+a + pad+a* (people+path) 'country-side' → *chon+á + bòt* (emphatic speech) vs. *chon+a +bòt* (normal speech); or when the second member of the compound nouns begins with a vowel, e.g. *jal+a* (Skt.) 'water' + *aalay+a* (Skt.) 'dwelling, house' → *chá? laa lay* (emphatic speech) vs. *cha laa lay* (normal speech).

iii) Indic Compounds:

Thai has borrowed heavily from the Indic vocabulary through Thai nativization processes. We will first demonstrate simple word borrowing. Examples are provided in (26a-b). The adaptations of certain phonemic sounds to fit the Thai phonetic inventory, e.g. dental and sibilant neutralizations, aspiration and devoicing of stop consonants (cf. section III above), where applicable, are all shown in step ø of (26a-b) below, since no ordering relation is assumed among these phonological rules. For each borrowed lexical item, the Indic syllabification is shown in step 1. Being a tonal language, Thai assigns tone according to the syllable structure and the phonetic properties of both the onset and rime (step 2). Final nasals

are deleted (step 3). When a segment is deleted from a syllable, several adjustments are underway. First, as noted by Thai researchers, all syllables in Thai are closed by either a sonorant or a stop, including a glottal stop. This glottal stop may be deleted in unstressed position. However, since the stress domain of Thai words is the final syllable, a bare short vowel may not surface in word-final position. Consequently, no tone can be placed or retained on the ill-formed syllable structure (cf. the output after the word-final nasal is deleted **thằ* in *rát.*thằ* in step 3). To remedy the situation, a glottal stop is inserted (step 4) which in turn, calls for tone reassignment since the stressed, checked syllable conditions a low tone in unmarked cases (step 5), unless the onset is marked as [-voice, +spread].

(26) a. raṭṭhaṁ (Pali)	→ rát		'country'
ø. raṭṭhaṁ	→	raṭṭhaṁ	Dental Neutralization
1. raṭṭhaṁ	→	rat.thaṁ	Syllabification
2. rat.thaṁ	→	rát.thăm̀	Tone Assignment
3. rát.thăm̀	→	rát.*thằ	Final Nasal Deletion
4. rát.*thằ	→	rát.*thằʔ	ʔ-Insertion
5. rát.*thằʔ	→	rát.thàʔ	Tone Reassignment

The derivation in (26b) is in a similar fashion. Step ø denotes the neutralization of all sibilants. Syllable structure is assigned in step 1 followed by tone assignment in step 2 which succeeds only in the first syllable while the second syllable is devoid of any tone since its syllable structure is not permitted in Thai. Step 3 illustrates the non-applicability of the nasal deletion rule since the word in question does not contain a word-final nasal. Syllable adjustment by the insertion of a glottal stop takes effect in step 4 so that tone can be placed on the final syllable in step 5.

b. šaastra (Skt.)	→ sàat		'science/art'
ø. šaastra	→	saastra	Sibilant Neutralization
1. saastra	→	saas.*tra	Syllabification
2. saas.*tra	→	sàas.*tra	Tone Assignment
3. -		-	Final Nasal Deletion
4. sàas.*tra	→	sàas.traʔ	ʔ-Insertion
5. sàas.traʔ	→	sàas.tràʔ	Tone Reassignment

The outputs at this stage of derivation serve as inputs to compounding. Step 6 illustrates the juxtaposition of two or more simple words. The glottal stop of the first member of the compound is deleted in step 7 because it no longer stands at the edge of the word since the word boundary is now extended to cover the entire compound. Consequently, the previously-assigned tone is also removed as a consequence of loss of stress in non-final position. The elimination of glottal stop and mid-tone neutralization are also found at the postlexical or sentential level where words are put together to form phrases or sentences.

c. rát.thàʔ + sàas.tràʔ→ rát.tha sàat 'political science'
6. rát.thàʔ+sàas.tràʔ→ rát.thàʔ sàas tràʔ Compound Formation
7. rát.thàʔ+sàas.tràʔ→ rát.thà sàas tràʔ Non-final ʔ Deletion
8. rát.thàʔ+sàas.tràʔ→ rát.tha sàas tràʔ Mid-Tone Neutralization

Most simple words (cf. (26' a-b)) and compound lexical items (cf. (26' c) entered the Thai lexicon with the truncation of the final vowel (and glottal stop) (step 9), followed by the resyllabification of the remaining phonetic material (step 10), cluster simplification (step 11), final stop neutralization (step 12), and tone reassignment in accordance with a new syllable structure (step 13) in case a different syllable structure is derived.

(26') a. raṭṭhaṁ (Pali)	→ rát		'country'
9. rát.thàʔ	→	rát.th	Final Vowel (and ʔ-) Truncation
10. rát.th	→	rátth.	Resyllabification
11. rátth	→	rát	Final Cluster Simplification
12. -		-	Final Stop Neutralization
13. -		-	Tone Reassignment
b. šaastra (Skt.)	→ sàat		'science/art'
9. sàas.tràʔ	→	sàas.tr	Final Vowel (and ʔ-) Truncation
10. sàas.tr	→	sàastr.	Resyllabification
11. sàastr	→	sàas	Final Cluster Simplification
12. sàas	→	sàat	Final Stop Neutralization
13. -		-	Tone Reassignment

c. rát.thàʔ + sàas.tràʔ→ rát.tha sàat 'political science'

9. rát tha sàas.tràʔ	→ rát tha sàas.tr	F. V (& ʔ-)Truncation
10. rát tha sàas.tr	→ rát tha sàastr.	Resyllabification
11. rát tha sàastr	→ rát tha sàas	F. Cluster Simplification
12. rát tha sàas	→ rát tha sàat	F. Stop Neutralization
13. -	-	Tone Reassignment

6.2.3 Language games involving foreign loans

Interactions between morpho-phonological structures and stress are most evident in tone assignment to foreign loanwords in language games. We will attempt at a schematic characterization of Thai tone in language borrowing in the paragraph immediately below. (For a full discussion of Thai tonal adaptation of loanwords from both tonal languages such as Chinese and non-tonal languages such as English and Japanese, the reader is referred to Wong-opasi (in preparation).)

7. Tones in foreign loanwords

Gandour 1979 offers an insightful analysis of Thai tones in English loanwords in which he attributes the tonal development in English loans to the interactions between syllable structure and stress patterns of both English and Thai. Following Gandour, we exemplify the Thai tonal patterns in three classifications of words:monosyllables, bisyllables, and polysyllables.

7.1 Monosyllables

Tone is assigned based on the underlying syllable structures of English. That is, unchecked syllables receive the M tone while checked syllables acquire the H tone, e.g. *cream* → *khriim; share* → *chɛɛ; free* → *frii* vs. *jet* → *cét; soup* → *súp; card* → *kháat; cake* → *khéek.* The underlying structure of English syllables

is crucial to understand the counterexamples of a H tone on surface unchecked Thai syllables such as *bank* → *bɛ́ŋ; pump* → *pám* and *pipe* →*páy*; *mouse* (as in 'computer mouse') → *máw*. In effect, tone is assigned first according to the English syllable structures before cluster simplification to avoid violations of Thai syllable structure constraints which prohibit word-final clusters containing two consonants in a row (*CVCC) or a diphthong followed by a consonant (*CVGC) as the above examples reveal.

7.2 Bisyllables

The intervention of stress in tone assignment is more obvious in words with more than one syllables. For penultimate stress which produce a falling tone on the final syllable, e.g. *vísa* →*wii sâa; párty* → *paa tîi; dóctor* → *dɔ́k tə̂ə*, Gandour provides an explanation in (27).

(27) The stressed-unstressed English pattern correlates with a falling pitch contour. Since Thai rhythm requires that the last syllable in a phrase be stressed, it would appear that the falling pitch contour has been preserved in the Thai pronunciation, but that the point of the fall has been shifted to the final syllable in accordance with Thai rhythmic constraints. (Gandour 1979:137)

However, the HL pitch pattern is rendered in two different ways in final checked syllables, namely, as L or H, e.g. *crédit* → *khree dìt; pássport* → *pháat sa pɔ̀ɔt* vs. *bónus* → *boo nát; sándwich* → *sɛɛn wít*. To account for these discrepancies, Gandour attributes them to competing strategies in tone assignment, phonetic vs. non-phonetic motivations like ENGLISH orthographic influences.

In contrast, the correlations between ultimate stress and the mid tone, e.g. *shampóo* → *chɛm phuu; hotél* → *hoo ten*, are described as follows:

(28) These English source words have an overall rising stress pattern, the second syllables being comparatively higher in pitch than the first, and longer. If these words were to be adapted with the falling tone of the second syllable, the resultant tonal pattern would be considerably different from the perceived stress pattern. Thus, the final syllables of these bisyllabic words are assigned a mid tone which results in a closer approximation to the English stress pattern.

The tonal adaptation in (28) is not without exceptions as given in (29). These exceptional tonal behaviors are cited without explanations in Gandour (1979:139). (The comments in parentheses are made by the present author.)

(29) sóda	→ soo daa	*soo dâa
bílliard	→ bin lîat	*bin lìat, bin líat
nécktie	→ nék tháy	*nék thây (but also nék thay)
Chrístmas	→ khrít sa mâat	*khrít sa màat; *khrít sa máat

7.3 Polysyllables

Polysyllabic words are assigned tone primarily according to English syllable structure, irrespective of the pitch patterns of English stress.

(30) The rules for tonal assignment are based strictly on the interpretation of English syllable structure. Those syllables interpreted as smooth receive the mid tone in non-final position, the falling tone in final position; those syllables interpreted as checked receive the high tone in non-final position, the low tone in final position. Short open syllables in English source words that occur between a primary stressed syllable and a following syllable are assigned a mid tone in accordance with the tone reduction rule in Thai. Since the tonal patterns remain fixed in the adaptation of variable stress patterns found in polysyllabic English words, we cannot attribute the resultant tonal patterns to perceptual interpretation of the variable pitch contours associated with the English stress patterns. (Gandour 1979:140)

Again, alongside predictable tonal patterns in *compúter → khɔm phiw tə̂ə Chicágo → chí? khaa kôo; hámburger → hɛm bəə kə̂ə; lóttery → lɔ́ ta rîi*, we also encounter exceptions such as names of certain countries pronounced with the M tone on the final syllable, e.g. *América → ?a mee ri kaa; Swítzerland → sa wít sa lɛɛn*. According to Gandour, these exceptions are, again, due to a conventionalized reading pronunciation of ENGLISH orthography (present author's emphasis).

In spite of Gandour's keen observations of the influences from the donor language, we would like to elaborate an analysis which gives precedence to the stress pattern and syllable constructs of the host language. As seen in section VII above, for every generalization made on tone caused by stress and syllable structure of the mono-, bi-, and poly-syllabic English loans, there are exceptions which, following Gandour, suggests competing strategies in tone assignment. We will first comment on the influence of stress.

Concerning the stress patterns of English, although pre-final stress actually corresponds to the falling tone on final unchecked syllables in certain words, no conclusive evidence is found as we find both the HL and M tones assigned to the final unchecked syllable of bi- and poly-syllabic English words with non-final stress, e.g. *vísa →wii sâa* vs. *sóda → soo daa (*soo dâa); Chicágo → chí? khaa kôo* vs. *América → ?a mee ri kaa (*?a mee ri kaa)*. Neither can we find a uniformed tone assignment on final unstressed checked syllables, e.g. *crédit → khree dìt* vs. *sándwich → sɛɛn wít*. On the contrary, words with final stress in the English source always get the M tone on unchecked syllables, e.g. *shampóo → chɛm phuu (*chɛm phûu)* while the L tone is invariably assigned to checked syllables, e.g. *promóte → proo mòot*. This is due to the fact that the Thai final stress pattern is in agreement with that found in English, as pointed out by Gandour. We take this to reconfirm our already-proven claim for Indic loans that the domain-final syllable in Thai is heavy, and that only a long vowel or a short vowel with a glottal stop and L tone can surface in the unmarked cases. From this hypothesis, the discrepant tone assignments in English follow directly. A digression to a

discussion of tone in monosyllables within the framework of the present analysis is necessary at this point.

We claim in section V that the default tone for light or unchecked syllables is M and that the unmarked tone for heavy or checked syllables is L. This generalization covers both Thai native words and Indic loans as it is supported by Gandour's 1982 study of the frequency of tones in Thai words, ranking the frequency percentage as follows:

(31) M (39.98%) L (20.72%) F (17.33%) H (11.81%) R (10.16%)
(from a total of 61,222 tone occurrences in 25,000+ entries taken from Haas' 1964 dictionary)

The default tone assignment can also be extended to Non-Indic loanwords like English and Japanese. Examples are:(Eng.) *beer* → *bia; mile* → *may; film* → *fiim;* (Jap.) *zen* →*sen; yen* → *yen;* (Eng.) *date* → *dèet; gate* → *kèet; vote* → *wòot.* (Japanese checked monosyllablic loans are rare in Thai.)

However, there is an additional psychological factor in tonal development in foreign loans. That is, speakers of a borrowing language may employ a special tone to indicate that the word is clearly a foreign item in the lexicon of the donor language as reported in Kiu 1977. In the case of borrowing in Thai, unlike Indic loans which entered the Thai lexicon at a very early stage during the formation of the Thai language and have thus become an integral part of Thai vocabulary, for later borrowings from non-Indic sources, the Thai speakers feel the need to signal the foreign nature of these lexical words. Hence, there is a tendency to assign more marked tones to non-Indic loans. We hypothesize that these marked tones are the high and the rising tones (cf. the tone frequency scale in (31)). The H tone is an across-the-board assignment to all checked syllables as an attempt to neutralize the variable tones (i.e. L, HL, and H) conditioned by the underlying syllable structures in English, and by the effects of tone sandhi, i.e. tone change due to neighboring sounds, in the Chaozhou dialect of Chinese (see below). The rising tone, on the other hand, applies to unchecked syllables in Chaozhou, as another indicator, to distinguish the Chinese source since Thai and Chinese syllable structures are fairly similar.

Apart from external linguistic factors, the marked H tone is also based on phonetic grounds. As a result of tonogenesis from the Indic phonetic inventory, the majority of Thai consonants belong to the Low-consonants class, bearing the underlying [-vc., +sprd] phonetic features (cf. section III and the appendix). From the point of view of syllable structure in (5), we witnessed that both the CVC and CVVC structures can carry the L tone, while other tones are more restricted. The M tone is impossible due to the well-formedness condition in (17a), while the sequence LH never results in checked syllables because the tone imposed by the [-continuant] feature of the coda is the L tone (cf. (13)). The HL tone is barred from CVC syllables because the vowel is too short to bear a sequence of two tones in Thai. This leaves two possibilities, either the H or the L tone. While the Low-class consonants, including sonorants, have more restrictions than the other two classes of consonants, that is, they cannot bear a L tone when closed by a stop due to the

underlying [-vc., +sprd] features of the onset (cf. (11)). Thus, the H tone is assigned uniformedly to all checked monosyllables despite the three options L, HL, and H, e.g. *jet* → *cét; soup* → *súp; card* → *kháat; cake* → *khéek.* This psychological inclination also competes with the natural phonetic motivation, as evidenced in the conflicting tone assignments, L vs. H, in *AIDS* → *ʔèet. vs. Ed* → *ét; vote* → *wòot* vs. *Vogue* → *wóok.* The tendency for the high tone assignment to non-Indic loans is also reflected in Chinese loans, e.g. *pɛ̀ʔ* → *pɛ́ʔ* 'elder paternal uncle'; *kòk* → *kók* 'country', despite a possible direct tone transfer by preserving the original tone in Chinese. This is due to the fact that Chinese lexical words are subject to tone sandhi. Consequently, *kòk* 'country' is pronounced as *kók* in isolation but as *kók* in *kók uáŋ* 'king' (<- country+king). Regarding unchecked monosyllables, since some English syllable structures, e.g. peculiar onsets, fr-, fl-, sp-, st-, sk-, v-, or coda, -r, -l, -nk, -nt, -mp, etc., are sufficiently foreign to the Thai ears, the M tone is assigned without further needs to mark the non-native source. However, this strategy may compete with the preference for the H tone and both tones are found idiosyncratically, e.g. *cent* → *sen* vs. *saint* → *sén* (in 'Saint John School'), but also *sen* (in 'Saint Joseph School').

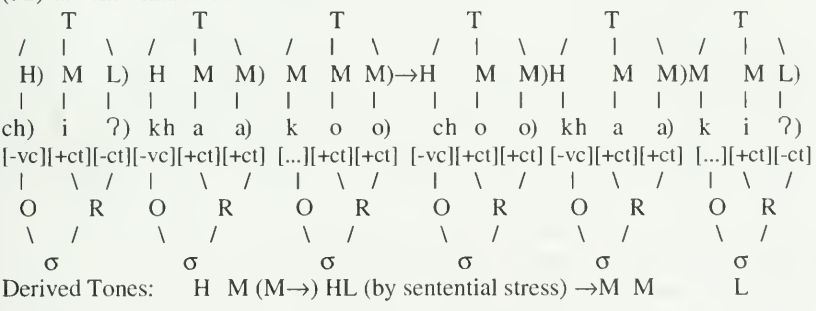
The strategy to differentiate foreign loans with a distinct tonal pattern from those found in the majority of words in the borrowing language persists in polysyllabic words as well. However, in polysyllabic words, there is another phonetic requirement, namely, stress. Contrary to Gandour, we claim that the Thai stress pattern plays a major role in tone assignment than the English stress patterns. Specifically, English final stress can fit into the Thai final stress pattern easily, resulting in M tone for CVV structure and L for CV(V)C, and no tone adjustment is needed. Pre-final English stress, in contrast, calls for certain adaptations. This is because unstressed syllables in English are normally shortened to a schwa. However, the final stress pattern in Thai requires that the vowel of the final syllable may not be reduced. As in the Indic cases, the vowel in question must be lengthened and assigned the M tone as regulated by the syllable structure requirements, e.g. *India* → *ʔin dia*, or be imposed a falling tone because of the psychological need to signal the foreign stress pattern, e.g. *chíʔ khaa kôo*, with the HL sequence combined on the last syllable. Otherwise, the shortened vowel in English must be assigned a glottal stop which entails the L tone, e.g. *khaa buu kìʔ*, with the split HL effect on two syllables, since the perceived overall pitch contour ML of the word approximates HL. The two competing strategies sometimes coexist as we find two pronunciations for proper names such as *Ithaca* → *ʔít tha kâa* vs. *ʔít tha kàʔ; Amko* → *ʔam kôo* vs. *ʔam kòʔ*, and for other not well-established loanwords. The proposition that the stress contour in English is of minor importance to the Thai stress pattern is also corroborated by the fact that the English stress pattern must be prohibited when it is in opposition to Thai phonology. Specifically, the HL pitch contour of English cannot be retained when the last syllable is CVC because of the shortness of the vowel. Therefore, only the L tone is retained (which still preserves the overall falling pitch of English), e.g. *promóte* → *proo mòot*, or in the case of a sonorant onset, only the H

tone is integrated on CVC, e.g. *ténnis* → *then nít,* while HL is not ruled out on CVVC, e.g. *Christmas* → *khrít sa mâat*, as in the case of Indic loans (cf. (19)).

8. Evidence from language games involving foreign loans

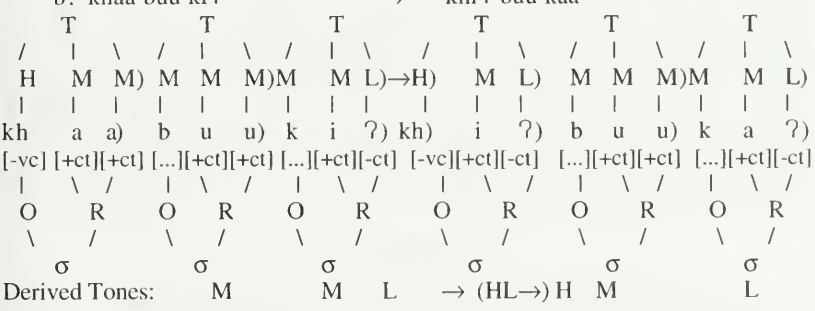
It is clear that tone is governed by syllable structure, stress, and the phonetic properties of each syllabic segment as some resulting tones in the language games may not be there underlyingly. In (32), all of the syllables in the source words are assigned tone according to rules (30). It is possible that the H M HL sequence of *chí? khaa kôo* be inverted following rime movement (i.e. *chôo khaa kí?*). However, HL tone preservation is impossible when the rime is turned into a short heavy syllable (*chóo khaa *kî?*). To preserve the tones, syllable rimes must be adjusted (*chó? khaa kîi*). However, the preferred form is *choo khaa kì?* since both the new tones and rimes are in agreement with the Thai stress patterns. Likewise, *khaa buu kì?* may be rendered *khì? buu kaa*. Nevertheless, the mid tone can never be retained on short heavy syllables (**khi? buu *kàa*) according to the well-formedness condition (17a). The preferred syllable structures for bearing M tone is a long light syllable (*khii*), and for L tone in final position, it is a short heavy syllable (*kà?*), i.e. *khii buu kà?*. The best preferred form, however, is *khí? buu kà?*.

(32) a. chí? khaa kôo → choo khaa kì?



Derived Tones: H M (M→) HL (by sentential stress) →M M L

b. khaa buu kì? → khí? buu kaa



Derived Tones: M M L → (HL→) H M L

9. Conclusion

We hope to have presented an extensively convincing analysis of tone assignment in Thai, under the assumptions that tonal assignment in Thai is modular, having a cyclic application, and that it must conform to the well-formedness conditions (17) at each stage of derivation. The modular tone assignment accounts for tone change at morpheme junctures, word boundaries, and within syllables. Through the cyclic application, Yip's stipulation of the prohibition of three-tone syllables can be dispensed with. The present study translates the three canonical tonal consonant classes (i.e. M, L and H) into morphophonological marking of underlying phonetic features resulting from historical sound change. We point out the importance of the recognition of tonal consonant classification as it facilitates the acquisition of tonal patterns in Thai. Without such marking in the lexicon, we would have no way to derive varying tonal patterns in Indic loanwords. We would also have to devise various complicated tonal rules which at times may fail to yield the correct results. Moreover, this morphophonological marking is needed independently for all lexical items in Thai since Thai tones are lexical but are constrained by the segmental properties and the structure of each syllable, and for loanwords, the stress patterns of both the donor and the borrowing languages as well, although the stress pattern of the borrowing language takes precedence.

Regarding syllable structure, the relevance of syllable weight in the operation of phonological rules is proven in the literature (for stress, see Halle & Vergnaud 1987; Wong-opasi 1987, etc.) For tone, the light syllable weight allows the syllable more flexibility to carry any of the five tones, including contour tones. Because syllables closed with a stop segment involve glottalization, it creates extra weight on these syllables, and hence it restricts HEAVY syllables from carrying certain tones. Unless intervened by some special phonetic properties of the onset, heavy syllables are assigned L tone; thereby, excluding the default M tone (*M). The shortness of vowels in heavy syllables accounts for their inability to accommodate both the extra syllable weight, imposed by the Thai final stress, and a contour tone (i.e. *HL or *LH on stressed CVC syllables), except when the contour tone is intonationally derived or in some marginal lexical words. Further, postulation of the [-cont] feature as causing L tone explains the absence of rising tone (*LH) on all Thai heavy syllables, long or short, despite its presence in other languages. This is due to the fact that the resulting tone would have the L tone, never the H tone, at the end of the contour tones. Aside from morphophonological factors, psychological factors do influence tone assignment as well since Thai has special tone patterns to mark English, Chinese, and Japanese loanwords from Indic loanwords and Thai native words.

NOTES

[1] I wish to thank Jack Gandour for kindly sending me his various papers on Thai. Special thanks are also due to Martha Ratliff for editing an earlier version of this paper. All errors of interpretation, however, are my own. Last, but not least,

this version of the paper has included substantial changes since its presentation at the first annual meeting of SEALS.

[2] The phonetic transcription in this study may be different from those employed in the papers being discussed. Specifically, off-glides are represented as [y, w] whereas on-glides are written as [i] and [u], respectively. All long vowels are transcribed as sequences of two identical vowels while short vowels are single vowels followed by a glottal stop except when eliminated by some phonological rule. All tones in the phonetic transcription are written on the first vowel irrespective of the gliding or vocalic properties of the high segment. Such tone values as M, L, H, HL, and LH on the examples are given as necessary.

[3] Another example of Leben's is *tɔ̂ŋ kaan* → *tɔ̌ŋ kaan* 'want' is questionable since underlyingly, the first lexical item contains a long vowel and it is shorthened *tɔ̂ɔŋ kaan* → *tɔ̂ŋ kaan*.

[4] For more details, the reader is invited to read the full discussion in her paper.

[5] The development of a high tone as in (12) is also attested in the evolution of Punjabi from Sanskrit (Bhatia 1975; Hock 1985).

[6] Phoneticians are invited to check out the validity of this hypothesis from laboratory analyses of the actual Thai isolative speech style.

[7] Along this line of hypotheses, David Strecker (in Comrie (1987:753)) discusses the development from Proto-Tai to the modern Tai languages. He states that the Thai words for 'face' and 'mother's younger sibling' had the same tone but different initials, namely a voiceless vs. voiced nasal, respectively, whereas in modern Thai they have the same initial but different tones as falling vs. high tone, in that order.

[8] As predicted, the bracket of the first tonal domain is erased in light syllables with devoiced aspirated onsets while the first tone is retained in light syllables with original aspirated onsets. This is due to the fact that tone is assigned in the first case before devoicing and aspiration apply while in the second case, it is assigned directly to the intrinsic [+spread] property of original aspirated onsets.

[9] These final vowels (plus an applicable nasal) are Indic declension endings. They were deleted in the majority of cases in the Thai adoption (Phanthumetha 1975:26-47).

[10] This lexical word is given with a final [ṁ] in HRH. Prince Kitiyakara Krommaphra Chandaburinarunath's dictionary. However, according to Phanthumetha (1975:68), the ending for Pali is [ŋ] while it is [ṁ] for Sanskrit.

Appendix: From Indic Consonant Inventory to Thai

Tonal Consonants	Mid	High	Low		
stops:	(-vc, -spr)	(-vc, +sprd)	(+vc, -sprd)	(+vc, +sprd)	(+nas)
Thai orthography:	ก	ข	ค	ฆ	ง
Indic velars:	k	kh	g	gh	ŋ
Thai velars:	k	kh		kh	ŋ
Thai orthography:	จ	ฉ	ช	ฌ	ญ
Indic palatals:	c	ch	j	jh	ñ
Thai palatals:	c	ch		ch	y
Thai orthography:	ฏ	ฐ	ฑ	ฒ	ณ
Indic retroflexes:	ṭ	ṭh	ḍ	ḍh	ṇ
Thai dentals:	t	th		th	n
Thai orthography:	ต	ถ	ท	ธ	น
Indic dentals:	t	th	d	dh	n
Thai dentals:	t	th		th	n
Thai orthography:	ป	ผ	พ	ภ	ม
Indic labials:	p	ph	b	bh	m
Thai labials:	p	ph		ph	m

Fricatives:	(+strident)			(+glottal/laryngeal)		(+labiodental)	
	High		Low	High	Low	High	Low
Thai ortho.	ศ ษ	ส	ซ	ห	อ	ฝ	ฟ
Indic	ś ṣ	s	-	h	-	-	-
Thai	s		s	h	h	f	f

Glides:	Low		Liquids:	Low	
Thai ortho.	ย	ว	ร	)	ฬ
Indic	y	v	r ṛ ṛh	l	ḷ
Thai	y	w	r	l	

Thai invention: ฎ ฑ (d) from Indic (ṭ), (ṭ), respectively; ɓ (b) from(p); and syllable initial อ (ʔ)

REFERENCES

BAO, Z. M. 1990. On the nature of tone. Massachusetts Institute of Technology, Ph. D. dissertation in Linguistics.

BHATIA, Tej. 1975. The evolution of tones in Punjabi. Studies in the Linguistic Sciences. 5:2.12-24.

BUDDHADATTA, A. P. 1968. Concise Pali-English dictionary. Colombo: The Colombo Apothecaries.

BURZIO, Luigi. 1991. Principles in phonology. Paper presented at the annual meeting of the Linguistic Society of America, Chicago.

CHENG, C. C. 1973. A quantitative study of tone in Chinese. Journal of Chinese Linguistics. 1:93-110.

CLEMENTS, George. 1985. The geometry of phonological features. Phonology Yearbook 2:225-252.

DUANMU, San. 1990. Formal study of syllable, tone, stress and domain in Chinese languages. Massachusetts Institute of Technology Ph. D. dissertation in Linguistics.

GANDOUR, Jack. 1974. On the representation of tone in Siamese. UCLA Working Papers in Phonetics 27:118-146.

——. 1979. Tonal rules for English loanwords in Thai. Southeast Asian Linguistic Studies 4:131-144.

——. 1982. The relative frequency of tones in Thai. Pacific Linguistics, Series A-No. 62: Papers in South-East Asian Linguistics 8: Tonation. David Bradley (Ed.), 155-159.

HAAS, M. R. 1964. Thai-English student's dictionary. Stanford: Stanford University Press.

HALLE, Morris, & K. N. Stevens. A note on laryngeal features. R.L.E. Quarterly Progress Report. 101:198-213.

——, & Jean-Roger Vergnaud. 1987. An Essay on Stress. Cambridge: The MIT Press.

HAUDRICOURT, A. G. 1961. Bipartition et tripartition des systèmes de tons dans quelques langues d'Extrême Orient. Bulletin de la Societé Linguistique de Paris 56:163-180.

HENDERSON, E. J. A. 1949. Prosodies in Siamese. Asia Major (New Series) 1:189-215. (Reprinted in Prosodic Analysis , F. R. Palmer (Ed.). London: Oxford University Press, 1970).

——. 1967. Grammar and tone in south east Asian languages. Wissenschaftlische Zeitschrift der Karl-Marx-Universität. Leipzig 16:171-8.

HOCK, Hans H. 1986. Principles of Historical Linguistics. Amsterdam: Mouton de Gruyter.

HOMBERT, J.-M. 1978. Consonant types, vowel quality, and tone. Tone: A linguistic survey. Victoria Fromkin (Ed.). N.Y.: Academic Press, 77-111.

JEEL, V. 1975. An investigation of the fundamental frequency of vowels after various Danish consonants, in particular stop consonants. Annual Report of the Institute of Phonetics, University of Copenhagen. 9:191-211.

KIU, K. L. 1977. Tonal rules for English loan words in Cantonese. Journal of the International Phonetic Association. 7/1:17-22.

LEA, W. A. 1972. Intonational cues to the constituent structure and phonemics of spoken English . Purdue University Ph.D. dissertation in Linguistics.

LEBEN, W. 1971. On the segmental nature of tone in Thai. R.L.E. Quarterly Progress Report. 101:221-4.

——. 1973. Suprasegmental Phonology . Massachusetts Institute of Technology, Ph.D. dissertation in Linguistics.

OHALA, John J. 1973. The physiology of tone. In L. Hyman (Ed.), Consonant types and tone. Southern California Papers in Linguistics. 1:1-14.

——. 1978. Production of tone. Tone:A linguistic survey. Victoria Fromkin (Ed.). N.Y.: Academic Press, 5-39.

PEYASANTIWONG, Patcharin. 1986. Stress in Thai. Papers from a Conference on Thai Studies in Honor of William J. Gedney. Robert J. Bickner, et al. (Eds.), 211-230. (Originally presented at the 13th International Conference on Sino-Tibetan Languages and Linguistics, University of Virginia, October 1980.)

PHANTHUMETHA, Banchop. 1975. Bali Sanskrit nai phasaa Thai. Bangkok: Ramkhamhaeng University Press.

SAGEY, Elizabeth. 1986. The Representation of Features and Relations in Nonlinear Phonology. Massachusetts Institute of Technology, Ph.D. dissertation in Linguistics.

STRECKER, David. 1987. Tai languages. The World's Major Languages. Bernard Comrie (Ed.). Oxford: Oxford University Press, 747-775.

SURINTRAMONT, A. 1973. Some aspects of underlying syllable structure in Thai:evidence from KHAMPUAN-a Thai word game. Studies in the Linguistic Sciences. 3:1.121-142.

HRH. Prince Kitiyakara Krommaphra Chandaburinarunath. 1969. Pali-Thai-English-Sanskrit Dictionary. Bangkok: Mahaamakutaraajavidyaalaya King Mahaamakutaraajavidyaalaya.

TUMTAVITIKUL, A. 1991. Perhaps, the tones are in the consonants? Paper presented at the annual meeting of the Linguistic Society of America, Chicago.

WAROTAMASIKHADIT, U. 1967. Some phonological rules in Thai. Journal of the American Oriental Society. 87:4.541-574.

WONG-OPASI, U. 1987. Lexical Phonology and the Spanish Lexicon. Bloomington: Indiana University Linguistics Club (1990).

——. (Forthcoming). An autosegmental account of tone and stress in language contact. MS.

YIP, Moira. 1980. The tonal phonology of Chinese. Bloomington: Indiana University Linguistics Club.

——. 1982. Against a segmental analysis of Zahao and Thai: A laryngeal tier proposal. Linguistic Analysis 9:1.79-94.

Studies in the Linguistic Sciences
Volume 23, Number 2, Fall 1993 [publ. October 1996]

ADJECTIVAL AND DETERMINATIVE MEASURE PHRASES AND NP INTERPRETATIONS IN MANDARIN CHINESE*

Mary A. Wu

In Mandarin Chinese (MC), the order of prenominal elements such as measure phrases (MP) and modifying phrases (MOD-*de*) correlates with different readings of the NP relative to the definite and indefinite distinction (Annear 1965; J. Huang 1982, 1983). In this paper, I show first that definite and indefinite readings of NPs correlate not only with the order of MP and MOD-*de* but also with the syntactic category that combines with *de*. Secondly, I argue that the order-related definiteness facts fall out from Partee's 1987 assumption that numerals, hence MPs, may be adjectival or determinative, and from the assumption that NPs in MC may have a null determiner which, in principle, can be interpreted as definite or as indefinite.

1. Introduction: order-related definite interpretations

Mandarin Chinese (MC) lacks a word for the definite article. Definiteness of MC NPs may be signaled by the order of words (more precisely, of constituents) within the NPs. The relation between word order and definiteness in MC NPs is the subject of this paper.

In MC NPs, measure phrases such as *liang-tou* 'two head (of)' occur prenominally.[1] MC NPs may also contain prenominal modifying phrases marked by the particle *de*, like for example *hen xin de* 'very new *de*'. I will refer to modifiers of this form as MOD-*de* and Measure Phrases as MP. MP and MOD-*de* may be ordered as in (1) and (2) below.

(1) MP MOD-*de* N'

(2) MOD-*de* MP N'

The order of MP and MOD-*de* appears to correlate with different readings of the NP (Chao 1968; Annear 1965; J. Huang 1982:152-153, 1983).[2] For example, in (3), MP precedes MOD-*de* and the NP may be interpreted as indefinite or definite, but in (4), MOD-*de* precedes MP and the NP has a definite reading only. (*cl* = CLASSIFIER in the gloss.)

(3) [$_{NP}$ MP MOD-*de* N']
Liang-tang shangwu qu Badaling de lieche yijing fachuqule.
two-cl morning go Badaling de train already dispatched
'Two/The two trains going to Badaling in the morning have already been dispatched.'

(4) [NP MOD-*de* MP N']
Shangwu qu Badaling de liang-tang lieche yijing fachuqule.
morning go Badaling de two-cl train already dispatched
'The two trains going to Badaling in the morning have already been dispatched.'

While the correlation of the definite reading with the order MOD-*de* preceding MP has been observed before in the literature, it has been described as a tendency (J. Huang 1983:51-52) or associated with a demonstrative in the post-MOD-*de* MP (Chao 1968, Annear 1965). There is evidence, however, that the correlation may be more than just a tendency. It systematically holds for NPs containing certain types of MOD-*de* but not for certain others, and it is not limited to cases with demonstratives in the MP. For instance, the order MOD-*de* MP is associated with the definite interpretation in (4), in which a VP-*de* precedes the MP. But indefinite readings are quite common where MOD-*de* precedes MP if MOD-*de* contains a possessive NP or an NP_{loc}, i.e., an NP followed by a locational particle such as *li* 'in', *shang* 'on', etc. (Li & Thompson 1981:391).[3] As in (5a) and (5b), both indefinite and definite readings of the larger NP are possible even though MOD-*de* precedes MP.

(5) a. [NP MOD-*de* MP N']
dashan li de liang-ge xuesheng
mountain in de two-cl student
'two/the two students in the mountain(s)'

b. [NP MOD-*de* MP N']
Zhangsan de liang-zhi shouzhi
Zhangsan de two-cl finger
'two/the two fingers of Zhangsan's'

An adequate account of Mandarin NPs should thus explain not only why the order MOD-*de* MP correlates with the definite-only reading in (4) but also why the same order fails to produce a definite-only reading for the NPs in (5).

Example (6) with a bare Noun *shu* 'book(s)' shows that in the absence of an MP and MOD-*de*, both definite and indefinite readings are in principle possible. (*part* = SENTENTIAL PARTICLE in the gloss.)

(6) Lisi maile [NP shu] le
Lisi bought book part
'Lisi has bought a book/some books.' or
'Lisi has bought the book(s).'

Sentence (7) says that Lisi is informed of the fine quality of cake in the dining-hall by a student/some students/the student(s) who work there, which shows that an NP containing a MOD-*de* but no MP may have both definite and indefinite readings as well.

(7) [NP Zai canting dagong de xuesheng] gaosu Lisi nali de dangao bucuo.
at dining-hall work de student tell Lisi there de cake not-bad
'A/Some student(s) working at the dining-hall told Lisi that cake there is pretty good.'
'The student(s) working at the dining-hall told Lisi that cake there is pretty good.'

MOD-*de* may also precede or follow quantificational phrases (QP) containing quantificational words such as *mei* 'every', as in (8), but the difference in the order of MOD-*de* and *mei-tang* 'every-cl', on a superficial examination, does not seem to correlate with a difference in truth-conditions.

(8) a. [NP QP MOD-*de* N']
mei-tang shangwu qu Badaling de lieche zai Shahezhen
every-cl morning go Badalingde train at Shahezhen
ting wu-fenzhong.
stop five-minute
'Every train going to Badaling in the morning stops at Shahezhen for five minutes.'

b. [NP MOD-*de* QP N']
Shangwu qu Badaling de mei-tang lieche zai Shahezhen
morning go Badaling de every-cl train at Shahezhen
ting wu-fenzhong.
stop five-minute
'Every train going to Badaling in the morning stops at Shahezhen for five minutes.'

While the focus of this paper is on the interpretation of NPs containing numerals and MOD-*de*, an adequate account needs to be able to account for cases like (6), (7) and (8) as well.

In the following I first show that when MOD-*de* precedes MP in the NP, definite-only readings of the NPs pattern with a number of syntactic categories in MOD-*de*, whereas when MP precedes MOD-*de*, both definite and indefinite readings are in principle possible with all categories in MOD-*de*. I then propose an analysis to account for the patterns observed. I need to point out that the generalizations argued for in this paper do not carry over to NPs in predicative positions or with emphatic stress. For example, an NP like *a student* in *Lisi is a student*, or NPs with emphatic stress, e.g., *shangwu qu Badaling de LIANG-TANG lieche* 'TWO trains going to Badaling in the morning', may not fit the patterns here.

2. The syntactic category combining with *de* and definite-only readings of the NPs

2.1. [NP MOD-de MP N'] with definite-only readings

The examples in (9) illustrate the kind of contrasts used in determining whether an NP has only a definite reading or not. By (9a), there are four trains going to Badaling in the morning. Adding (9b) to (9a) means two out of the four trains going to Badaling in the morning have already been dispatched. But adding

(9c) results in an infelicitous discourse. The only reading for (9c) is that there are just two trains going to Badaling in the morning, and this is at odds with (9a), which states that there exist four such trains. The infelicity resulting from continuing (9a) with (9c) makes it clear that the NP in (9c) with VP-*de* preceding MP has a definite reading only.

(9) a. Shangwu you si-tang lieche qu Badaling.
morning have four-cl train go Badaling
'In the morning there are four trains going to Badaling.'

b. [$_{NP}$ MP MOD-*de* N']
Liang-tang shangwu qu Badaling de lieche yijing fachuqule.
two-cl morning go Badaling de train already dispatched
'Two trains going to Badaling in the morning have already been dispatched.'

c. #[$_{NP}$ MOD-*de* MP N']
Shangwu qu Badaling de liang-tang lieche yijing fachuqule.
morning go Badaling de two-cl train already dispatched
'The two trains going to Badaling in the morning have already been dispatched.'

If we have (9a') instead of (9a), then both (9b) and (9c) may follow (9a') felicitously. The subject NPs in both (9b) and (9c) may be understood as anaphorically related to the two trains mentioned in (9a'), i.e., they may have the definite reading 'the two trains going to Badaling in the morning'. This again shows that NPs with the order [$_{NP}$ VP-*de* MP N'] have definite readings. It also shows that NPs with the order [$_{NP}$ MP VP-*de* N'] may have definite readings as well.

(9) a'. Shangwu you liang-tang lieche qu Badaling.
morning have two-cl train go Badaling
'In the morning there are two trains going to Badaling.'

b. [$_{NP}$ MP MOD-*de* N']
Liang-tang shangwu qu Badaling de lieche yijing fachuqule.
two-cl morning go Badaling de train already dispatched
'The two trains going to Badaling in the morning have already been dispatched.'

In (10), MOD-*de* is an NP-*de* which is not a possessive. According to (10a), five of the ten birds in the cage have black tails. (10b) may follow (10a) felicitously, but (10c) is infelicitous after (10a). (10c) may be used felicitously only if the number of black-tailed birds that are salient in the context is just two, i.e., it has a definite reading only.

(10) a. Longzi li you shi-zhi niao, qizhong wu-zhi weiba shi hei de.
cage in have ten-cl bird among-which five-cl tail be black de
'In the cage are ten birds, among which five have tails that are black.'

b. [$_{NP}$ Liang-zhi heiweiba de niao] hai hen xiao, bu hui fei.
two-cl black-tail de bird still very young not can fly
'Two birds with black tails are still very young and cannot fly.'

c. #[NP Heiweiba de liang-zhi niao] hai hen xiao, bu hui fei.
black-tail de two-cl bird still very young not can fly
'The two birds with black tails are still very young and cannot fly.'

If we replace (10a) with (10a'), then both (10b) and (10c) follow (10a') felicitously. In particular, it is possible to understand the NPs as referring back to the two black-tailed birds mentioned in (10a'). Therefore, [NP NP-*de* MP N'] has only a definite reading while both definite and indefinite readings are possible with [NP MP NP-*de* N'].

(10) a'. Longzi li you shi-zhi niao, qizhong liang-zhi weiba shi hei de.
cage in have ten-cl bird among-which two-cl tail be black de
'In the cage are ten birds, among which two have tails that are black.'

b. [NP Liang-zhi heiweiba de niao] hai hen xiao, bu hui fei.
two-cl black-tail de bird still very young not can fly
'The two birds with black tails are still very young and cannot fly.'

2.2. [NP MOD-*de* MP N'] with definite and indefinite readings

2.2.1. NP-*de* MP N' (with possessive NP-*de*)

Example (11) is similar to (10) in the sense that MOD-*de* is NP-*de*. Unlike in (10), however, NP-*de* is now possessive. Sentence (11a) says that five of the ten birds in the cage are Zhangsan's. Both (11b) and (11c) may follow (11a) felicitously with the interpretation that there are two birds that are Zhangsan's that are still very young and cannot fly. In either case, Zhangsan is assumed to have more than two birds, given (11a). Therefore, unlike the cases of [NP NP-*de* MP N'] observed in (10) with a non-possessive NP-*de*, [NP NP-*de* MP N'] with a possessive NP-de has a reading that is not definite here.

(11) a. Longzi li you shi-zhi niao, qizhong wu-zhi shi Zhangsan de.
cage in have ten-cl bird among-which five-cl be Zhangsan de
'In the cage are ten birds, among which five are Zhangsan's.'

b. [NP Liang-zhi Zhangsan de niao] hai hen xiao, bu hui fei.
two-cl Zhangsan de bird still very young not can fly
'Two birds that are Zhangsan's are still very young and cannot fly.'

c. [NP Zhangsan de liang-zhi niao] hai hen xiao, bu hui fei.
Zhangsan de two-cl bird still very young not can fly
'Two birds that are Zhangsan's are still very young and cannot fly.'

If we replace (11a) with (11a'), then both (11b) and (11c) may follow (11a') felicitously with a definite reading referring back to the two birds that are Zhangsan's in (11a').

(11) a'. Longzi li you shi-zhi niao, qizhong liang-zhi shi Zhangsan de.
cage in have ten-cl bird among-which two-cl be Zhangsan de
'In the cage are ten birds, among which two are Zhangsan's.'

b. [NP Liang-zhi Zhangsan de niao] hai hen xiao, bu hui fei.
two-cl Zhangsan de bird still very young not can fly
'The two birds that are Zhangsan's are still very young and cannot fly.'

c. [NP Zhangsan de liang-zhi niao] hai hen xiao, bu hui fei.
Zhangsan de two-cl bird still very young not can fly
'The two birds that are Zhangsan's are still very young and cannot fly.'

Cases like (12) provide further evidence that [NP NP-*de* MP N'] with the possessive NP-*de* may have indefinite interpretations. (12a) with [NP MP NP-*de* N'] seems to convey a sense of contrast, e.g., Zhangsan's fingers compared to someone else's; (12b) [NP NP-*de* MP N'] lacks this contrastive implication and sounds more natural than (12a) in non-contrastive contexts.

(12) a. #[NP San-zhi Zhangsan de shouzhitou] youdian zhong.
three-cl Zhangsan de finger somewhat swollen
'Three fingers that are Zhangsan's are somewhat swollen.'
b. [NP Zhangsan de san-zhi shouzhitou] youdian zhong.
Zhangsan de three-cl finger somewhat swollen
'Three fingers of Zhangsan's are somewhat swollen.'

More needs to be said about the contrastive implications of (12a). For the purpose of the present discussion, the point is that (12b) may be used felicitously if Zhangsan has ten fingers as people normally do, whether any of them are salient in the context or not. This should not be possible if [NP NP-*de* MP N'] had a definite reading only. Such a reading would require Zhangsan to have exactly three fingers or to have three that are salient in the context, as is the case with the English possessive in (13).

(13) John's three fingers were injured in an accident.

In (13), *John's three fingers* is definite and may be used felicitously only if John has exactly three fingers or three of his fingers are salient in the context. Therefore, [NP NP-*de* MP N'] with a possessive NP-*de* has a reading that is not definite in (12b).

2.2.2. NP_{loc}-*de* MP N'

In (14), MOD-*de* is NP_{loc}-*de*. According to (14a), there are six pots of flowers in the yard and six in the house. Both (14b) and (14c) may follow (14a) felicitously. This shows that both [NP MP NP_{loc}-*de* N'] and [NP NP_{loc}-*de* MP N'] may have an indefinite reading.

(14) a. Yuanzi li he wuzi li ge you liu-pen hua.
yard in and house in each have six-cl flower
'There are six pots of flowers in the yard and in the house respectively.'
b. [NP Wu-pen yuanzi li de hua] hen haokan.
five-cl yard in de flower very pretty
'Five pots of flowers in the yard are very pretty.'
c. [NP Yuanzi li de wu-pen hua] hen haokan.
yard in de five-cl flower very pretty
'Five pots of flowers in the yard are very pretty.'

If we replace (14a) with (14a'), then (14b) and (14c) may both follow (14a') felicitously with a definite reading referring back to the five pots of flowers men-

tioned in (14a'). Therefore, both definite and indefinite readings are possible for [NP NP_{loc}-*de* MP N'] and [NP MP NP_{loc}-*de* N'].

(14) a'. Yuanzi li he wuzi li ge you wu-pen hua.
yard in and house in each have five-cl flower
'There are five pots of flowers in the yard and in the house respectively.'

b. [NP Wu-pen yuanzi li de hua] hen haokan.
five-cl yard in de flower very pretty
'The five pots of flowers in the yard are very pretty.'

c. [NP Yuanzi li de wu-pen hua] hen haokan.
yard in de five-cl flower very pretty
'The five pots of flowers in the yard are very pretty.'

Applying the same kind of tests to NPs containing MP and various MOD-*de*'s in subject and object positions, we get the results in Table 1, which we will get to shortly.

3. Readings for NPs with MOD-*de* and *mei* 'every'

As already mentioned, an NP may contain a MOD-*de* and a QP headed by a quantificational word like *mei* 'every'. MOD-*de* may precede or follow QP, as in (8) repeated below.[4] NPs with QP preceding MOD-*de* and those with QP following MOD-*de* appear to have similar interpretations. However, a careful examination reveals that there are some meaning differences between the two orders, which I illustrate in the subsections below.

(8) a. [NP QP MOD-*de* N']
mei-tang shangwu qu Badaling de lieche zai Shahezhen
every-cl morning go Badalingde train at Shahezhen
ting wu-fenzhong.
stop five-minute

b. [NP MOD-*de* QP N']
Shangwu qu Badaling de mei-tang lieche zai Shahezhen
morning go Badaling de every-cl train at Shahezhen
ting wu-fenzhong.
stop five-minute

3.1. Partitive and non-partitive readings

Sentences like (15a) and (15b) bring out a difference with respect to partitive and non-partitive readings of the NPs containing MOD-*de* and QP. Both (15a) and (15b) may mean Lisi sympathizes with every student wearing glasses within a particular group, i.e., the NPs may have the partitive reading 'every one of the students wearing glasses'. However, while (15a) may also mean Lisi sympathizes with any student who wears glasses, (15b) lacks such a reading.

(15) a. [NP QP VP-*de* N']
Lisi tongqing mei-ge dai yanjing de xuesheng.
Lisi sympathize every-cl wear glasses de student
'Lisi sympathizes with any student wearing glasses.'
'Lisi sympathizes with every one of the students wearing glasses.'

b. [NP VP-*de* QP N']
Lisi tongqing dai yanjing de mei-ge xuesheng.
Lisi sympathize wear glasses de every-cl student
'Lisi sympathizes with every one of the students wearing glasses.'

3.2. Distributive and non-distributive readings

Sentences like (16a) and (16b) show a difference in the availability of distributive readings and collective readings. In (16a), where QP precedes MOD-*de*, there is only a distributive reading, i.e., each puppy performed two numbers. But (16b) is ambiguous. While a distributive reading is possible, the sentence may also mean the puppies jointly performed two numbers.

(16) a. [NP QP AP-*de* N']
Mei-zhi tiaopi de xiaogou biaoyanle liang-ge jiemu.
every-cl naughty de puppy performed two-cl number
'Every naughty puppy performed two numbers.'

b. [NP AP-*de* QP N']
Tiaopi de mei-zhi xiaogou biaoyanle liang-ge jiemu.
naughty de every-cl puppy performed two-cl number
'Every one of the naughty puppies (jointly) performed two numbers.' or
'Every one of the naughty puppies performed two numbers.'

4. Summary

The patterns of definite and indefinite readings of NPs vis-à-vis the order of MOD-*de*, MP and QP are summarized in Table 1.

Table 1. Patterns of Definite and Indefinite Readings of the NP vis-à-vis the Order of MOD-*de*, MP, and QP (every)

bare Nouns: def./indef. MOD-*de* N': def./indef.

MOD-*de* MP N'	MOD-*de* N'	MOD-*de* MP N'	QP MOD-*de* N'	MOD-*de* QP N'
VP-*de*	def./indef.	def. only	-pt./pt., distr.	pt., distr./coll.
AP-*de*	def./indef.	def. only	-pt./pt., distr.	pt., distr./coll.
PP-*de*	def./indef.	def. only	-pt./Pt., distr.	pt., distr./coll.
SP-*de*	def./indef.	def. only	-pt./pt., distr.	pt., distr./coll.
S/NP_{obj}-*de*	def./indef.	def. only	-pt./pt., distr.	pt., distr./coll.
MP-*de*	def./indef.	def. only	-pt./pt., distr.	?? pt.
NP-*de* [-poss]	def./indef.	def. only	-pt./pt., distr.	pt., distr./coll.
NP-*de* [+poss]	def./indef.	def./indef.	-pt./pt., distr.	pt., distr./coll.
NP_{loc}-*de*	def./indef.	def./indef.	-pt./pt., distr.	pt., distr./coll.

pt.=partitive; -pt.=non-partitive; distr.=distributive; coll.=collective;
SP=sentential predicate phrase, e.g., *yanjing hao* 'eye(s) good';
S/NP_{obj}=clause missing an object.

Both definite and indefinite readings are possible for bare Nouns and NPs containing MOD-*de* but no MP or QP. NPs with the order MP-MOD-*de*-N' have both definite and indefinite readings, whereas those with the order MOD-*de*-MP-

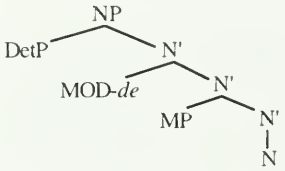
N' have only definite readings except when MOD-*de* is a possessive NP or NP_{loc}, in which case both definite and indefinite readings are possible. There are meaning differences between NPs containing QPs with quantificational words like *mei* 'every' as well. While both QP-MOD-*de*-N' and MOD-*de*-QP-N' have partitive readings, MOD-*de*-QP-N' lacks the more general reading of 'every' available for QP-MOD-*de*-N'. Also, both orders have a distributive reading, but QP-MOD-*de*-N' is only distributive, whereas MOD-*de*-QP-N' allows also a reading that is collective.

5. Towards an analysis of the patterns

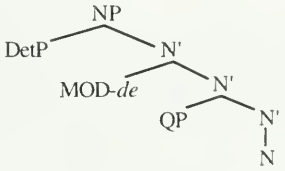
5.1. Structural realizations of the different word orders in Mandarin NPs

We already showed that MOD-*de* may precede or follow MP or QP. Since MOD-*de* is plausibly regarded as an N' modifier, it is also plausible to assume that, if MOD-*de* precedes MP or QP, then MP and QP are dominated by N' and the determiner position is phonologically not realized.[5] This gives us the structures in (17) corresponding to the order MOD-*de* MP/QP, where MP and QP occur under N' and the determiner phrase (DetP) is phonologically null.

(17) a.

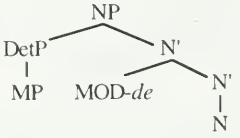


b.

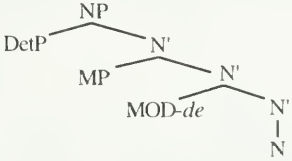


I assume, however, that, if MP or QP occur NP-initially, then MP and QP may occur in the DetP. Since MP and QP may occur under N', and the determiner need not be phonologically realized, it follows that the order MP/QP MOD-*de* does not univocally determine the structural position of MP and QP, since this order is also compatible with these phrases occurring under N'. Structures corresponding to the order MP/QP MOD-*de* are therefore represented as in (18) and (19). In (18a) and (19a), MP and QP are DetPs. In (18b) and (19b), they occur under N', and DetP is phonologically null.

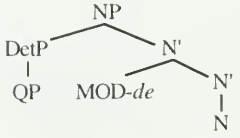
(18) a.



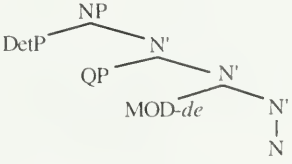
b.



(19) a.



b.



5.2. The semantic import of MP in different positions

The view that from a syntactic standpoint numerals can either be determiners or be dominated by N' is not new (Partee 1988). As suggested in Partee 1987, numeral words may have a double life. For instance, *three* may be an adjective or a determiner, and the two *three*'s have different semantic types (1987:130). Partee's analysis is motivated to explain the diversity of NP interpretations on the bases of general syntactic and semantic principles (1987:115). Following Partee, I assume that numerals can be either adjectives or determiners. Accordingly, MPs can either be adjectival MPs, hereafter $MP_{[+Adj]}$, which combine with an N' to form an N', or be determiner MPs, hereafter $MP_{[+D]}$, which are determiner phrases (DetPs) that combine with an N' to form an NP. Semantically, $MP_{[+Adj]}$ marks the cardinality of the N' it combines with. For example, the function of the adjectival MP 'three' in

$[_{N'}\ [_{MP[+Adj]}$ three] $[_{N'}$ men]]

is to restrict the set denoted by $[_{N'}$ men] to a set whose members are groups of three. On the other hand, $MP_{[+D]}$ denotes a relation between two sets and is in this sense quantificational. For example, the function of the determiner 'three' in

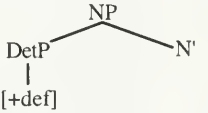
$[_{S}\ [_{NP}\ [_{DetP}$ three] $[_{N'}$ men]] $[_{VP}$ left]]

is to indicate that the intersection of the set denoted by the N' 'men' and the set denoted by the VP 'left' is not empty and has three members. The distinction between cardinal and quantificational MPs by itself does not give us the distinction between definite and indefinite readings of the NP. However, given this distinction and the assumption that when numerals occur under N', DetP is still present, though phonologically null, the definite and indefinite readings of the NP may then follow from the realizations of the DetP.

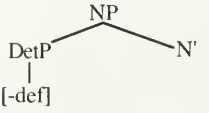
5.3. The realizations of the null determiner

I want to suggest that when the determiner is phonologically null, it may still be semantically active. In particular, the phonologically null det can in principle be realized as [+def] or [-def]. This means that, in the absence of grammatical factors constraining the realization of the null determiner, an NP whose determiner position is phonologically null may have the structures in (20):

(20)a.



b.



I am assuming, however, that the availability of the [+def] or [-def] interpretation of the null determiner is subject to the same constraints that limit the distribution of the determiners *some* and *the*. It may be noticed that while the definite determiner *the* may co-occur with (adjectival) numerals in English, the indefinite determiner *some* cannot (I am ignoring the 'about' reading of (21b)). The same may be said of the Chinese examples in (22): (22a) is fine, but (22b) is not.

(21) a. the three men
b.?? some three men

(22) a. Nei san-ge xuesheng zoule.
that three-cl student left
'Those three students left.'
b. *Yixie san-ge xuesheng zoule.
some three-cl student left

The exact nature of the constraint responsible for the unacceptability of (21b) and (22b) deserves further investigation. For the purpose of this paper, I will simply assume the existence of a grammatical constraint which rules out structures like (23) (whether or not DetP is phonologically realized):

(23) *[$_{NP}$ [$_{DetP}$ -def] [$_{N'}$ [$_{MP[+Adj]}$ num] N']]

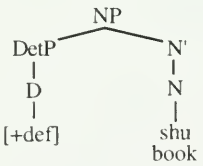
(23) says cardinal adjectives may not co-occur with an indefinite determiner in an NP. Given (23), the contrasts in (21) and (22) are expected. I now turn to some predictions this analysis makes for MC NPs.

6. Some predictions

6.1. The interpretation of NPs without MPs or QPs

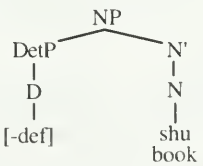
With the analysis proposed here, when no numeral or overt quantificational element is present in DetP, DetP can in principle be realized as [+def] 'the' or [-def] 'some', as in (24) and (25). This accounts for the indefinite and definite readings found with bare Nouns and with NPs containing MOD-*de* but no MP or QP, as in (6) and (7), respectively.

(24)a.



'the book(s)'

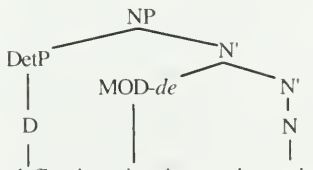
b.



'a/some book(s)'

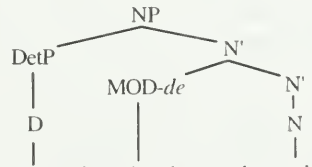
(6) Lisi maile [$_{NP}$ shu] le
Lisi bought book part
'Lisi has bought a book/some books.' or
'Lisi has bought the book(s).'

(25) a.



'the student(s) working at the dining-hall'

b.



'a/some student(s) working at the dining-hall'

(7) [NP Zai canting dagong de xuesheng] gaosu Lisi nali de dangao bucuo.
at dining-hall work de student tell Lisi there de cake not-bad
'A/Some student(s) working at the dining-hall told Lisi that cake there is pretty good.'
'The student(s) working at the dining-hall told Lisi that cake there is pretty good.'

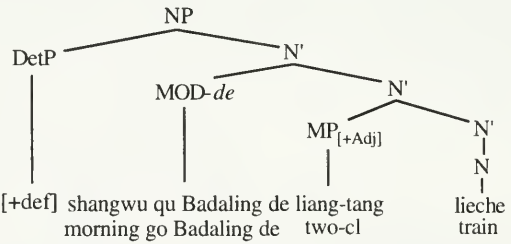
6.2. The interpretation of NPs with numerals

By the analysis proposed here, the null determiner can in principle be realized either as [+def] or as [-def], which are responsible for definite and indefinite readings of the NP. We observed however that in the case of (4), where MOD-*de* precedes the MP, only a definite reading is possible.

(4) [NP MOD-*de* MP N']
Shangwu qu Badaling de liang-tang lieche yijing fachuqule.
morning go Badaling de two-cl train already dispatched
'The two trains going to Badaling in the morning have already been dispatched.'

Why? Given our assumptions about the possible realizations of the null determiner, the subject NP in (4) has the structure in (26).

(26)



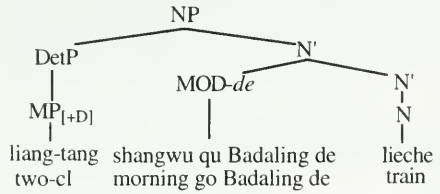
'the two trains going to Badaling in the morning'

The [-def] reading is ruled out by the same co-occurrence restriction which is responsible for ruling out expressions like **Some three boys left* in English. Due to the co-occurrence restriction on determiners and cardinal adjectives in the NP, when MP follows MOD-*de* and therefore is adjectival, DetP can only be realized as [+def], and the NP has a definite interpretation only. On the other hand, when MP precedes MOD-*de*, the MP may be under N' hence be adjectival, in which case DetP is [+def] again and the NP is definite. The other possibility is that the MP may be in DetP, in which case the NP is indefinite on the standard assumption that in the determiner interpretation numerals are existential quantifiers. This explains why cases like (4) with MOD-*de* preceding MP have a definite reading only and accounts for cases like (3), where MP precedes MOD-*de* and both definite and indefinite readings are possible.

(3) [$_{NP}$ MP MOD-*de* N']
Liang-tang shangwu qu Badaling de lieche yijing fachuqule.
two-cl morning go Badaling de train already dispatched
'Two/The two trains going to Badaling in the morning have already been dispatched.'

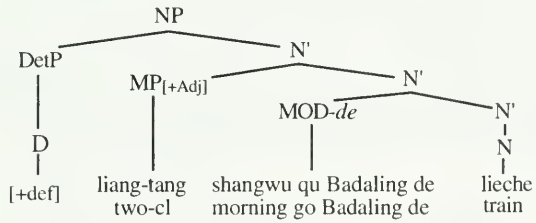
In (3), where MP precedes MOD-*de*, the MP may be in DetP or under N', as in (27) and (28) respectively. Since the determiner 'two' is indefinite, the NP in (27) has an indefinite reading. But the NP in (28) has no overt determiner, and its DetP is [+def] due to the co-occurrence restriction on determiners and cardinal adjectives in the NP. Therefore, the NP in (28) has a definite reading. And between (27) and (28), we have the two possible readings for (3).

(27)



'two trains going to Badaling in the morning'

(28)



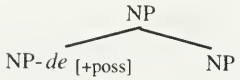
'the two trains going to Badaling in the morning'

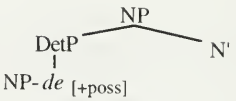
The analysis proposed here accounts for the patterns of definite and indefinite readings in Table 1 except for the cases where MP follows a possessive NP-*de* or an NP$_{loc}$-*de*. In these cases, both definite and indefinite readings are possible. We turn to such cases now.

6.3. Possessive NP-*de* and NP$_{loc}$-*de*

A plausible analysis for the readings found with NPs containing possessive NP-*de* and NP$_{loc}$-*de* would be that possessive NP-*de* and NP$_{loc}$-*de* have syntactic and semantic properties that set them apart from other types of MOD-*de*'s. But what may these properties be?

In L & T, NP-*de* is treated as an associative phrase (ASSOCP), which combines with an NP to form another NP (1981:113) or with a noun to form an NP (1981:126). In our terms this means that possessive NPs can occupy the positions in the trees in (29):

(29) a. 

b. 

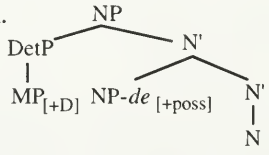
As Huang points out, however, possessive NPs in MC can receive both definite and indefinite interpretations. For example, *wo de shu* (Lit. 'I de book') stands for 'my book(s)' as well as 'book(s) of mine', which in English are definite and indefinite respectively. Also, possessive NP-*de* may co-occur with a demonstrative prenominally, as in *wo de neiben shu* 'my that book' (= that book of mine). Huang suggests that NP-*de* is therefore not a determiner and has no definitizing function as determiners do (1987:252). If this is correct, then the structure (29b) above is ruled out for possessive NPs, and we are left with structure (29a). L & T also claim that NP-*de* only occurs NP-initially (1981:124). Sentence (30a), however, is a counterexample to this claim which incorrectly rules out the reading in (30c).

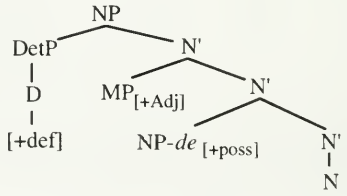
(30) a. liang-ge Beida de xuesheng
two-cl BU de student

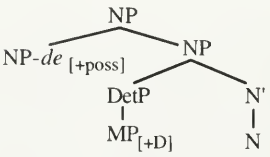
b. [$_{NP\text{-}de}$ liang-ge Beida de] xuesheng
two-cl BU de student
'students of two BU's'

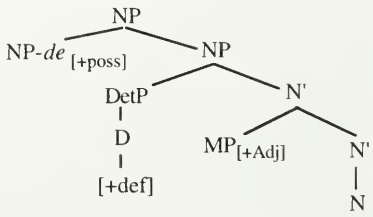
c. liang-ge [$_{NP\text{-}de}$ Beida de] xuesheng
two-cl BU de student
'two students of BU'

Following our assumption that MPs can either be DetPs or be cardinal adjectives, we are thus led to recognize the following possible structures as licensed. MP precedes the possessive NP in (31) and follows it in (32).

(31)a. 

b. 

(32)a. 

b. 

Given our analysis of the possible realizations of the null DetP, this means that while (33a) will allow the structures in (33b) and (33c), (34a) will allow both those in (34b) and (34c).

(33) a. liang-ben Zhangsan de shu
two-cl Zhangsan de book
b. [NP [DetP two-cl] [N' [NP-*de*[+poss] Zhangsan de] [N' book]]]
c. [NP [DetP +def] [N' [MP two-cl] [N' [NP-*de*[+poss] Zhangsan de]
[N' book]]]]

(34) a. Zhangsan de liang-ben shu
Zhangsan de two-cl book
b. [NP [NP-*de*[+poss] Zhangsan de] [NP [DetP two-cl] [N' book]]]
c. [NP [NP-*de*[+poss] Zhangsan de] [NP [DetP +def] [N' two-cl book]]]

This predicts correctly that both orders should allow both a definite and an indefinite reading. The analysis proposed here for possessive NP-*de* requires additional investigation, in particular independent evidence is needed for the structures I am assuming. This analysis provides, however, a plausible lead that could account for the exceptional behavior of possessive NPs. The analysis may work for NP_{loc}-*de* as well. Assuming that these NPs can occupy the same positions as possessive NPs, then again definite and indefinite readings are predicted for both orders of NP_{loc}-*de* and MP.

7. The interpretation of NPs with MOD-*de* and QP

As shown in (8), where a quantificational word *mei* 'every' is present, the different orders of the QP *mei-ge* 'every-cl' and MOD-*de* are still available.

(8) a. [NP QP MOD-*de* N']
mei-tang shangwu qu Badaling de lieche zai Shahezhen
every-cl morning go Badalingde train at Shahezhen
ting wu-fenzhong.
stop five-minute
'Every train going to Badaling in the morning stops at Shahezhen for five minutes.'

b. [NP MOD-*de* QP N']
Shangwu qu Badaling de mei-tang lieche zai Shahezhen
morning go Badaling de every-cl train at Shahezhen
ting wu-fenzhong.
stop five-minute
'Every train going to Badaling in the morning stops at Shahezhen for five minutes.'

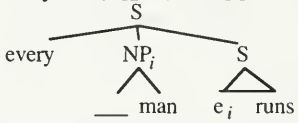
Two issues arise in the analysis of NPs containing MOD-*de* and QP here. First of all, although syntactically QP in (8b) does not have scope over MOD-*de*, semantically QP takes scope over the entire NP whether it precedes or follows MOD-*de* in the NP. Namely, (8b), like (8a), is true only on condition that every train going to Badaling in the morning stops at Shahezhen for five minutes. So what is the relation between the syntax and the semantics of NPs containing MOD-*de* and QP? Secondly, we observed that the difference in order correlates with a difference in partitive and non-partitive readings of the NPs. I will not at-

tempt a full treatment of these two issues here but will sketch some possible solutions in the sections below.[6]

7.1. Scope relations

Quantificational words like *every* are inherently relational, i.e., they always denote a relation between two sets. For example, in *every man runs*, *every* expresses a relation between the set of individuals with the property MAN and the set of individuals with the property RUN. In particular, *every* says the first set is included in the second. In comparison, numerals, as we saw, need not be relational in this sense, but may be cardinality predicates. For example, in *three men run*, *three* may simply mark the cardinality of sets of individuals that are men. In order to capture the relational nature of quantificational words like *every*, we may follow Heim (1982, 1990) and assume that in addition to Quantifier Raising (QR), there is a Quantifier Construal Rule (QC) which maps SS onto LF. QR is an operation that raises every non-pronominal NP out of S and adjoins them to S. QC is an operation that then attaches every quantifier as a leftmost immediate constituent of S (1982:132-136). For example, (35a) has the structure (35c) after QR and QC have applied.

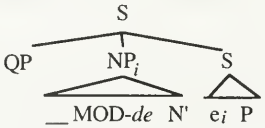
(35) a. every man runs
b. after QR: [S every man$_i$ [S e$_i$ runs]]
c. after QC:



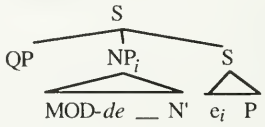
The tripartite structure in (35c) resulting from QR and QC consists of the quantifier *every*, its restrictor [$_{NP}$___man], and its nuclear scope [S e$_i$ runs]. Both the restrictor and the nuclear scope can be seen as denoting sets, namely as being of type <e, t>, and both are in the scope of *every*.

The analysis may be applied to NPs containing MOD-*de* and QP as well. As illustrated in (36a) and (36b), whether QP precedes or follows MOD-*de* syntactically, after QR and QC have applied, QP semantically takes scope over both the restrictor NP and the nuclear scope S and expresses a relation between two sets. It follows that since QP has scope over the NP at LF, it also has scope over MOD-de within the NP.

(36)a.



b.



7.2. Partitive and non-partitive readings

My proposal is that the partitive readings of NPs containing the quantifier *mei-ge* 'every' arise as a result of *mei-ge* co-occurring with the null [+def] deter-

miner. In other words, I am suggesting that the partitive readings for MC NPs containing the quantifier *mei-ge* correspond to LFs of form (37):

(37) [$_{QP}$ mei-ge] [$_{NP}$ [+def] N'] S
every the

Since *mei-ge* need not occur in determiner position and can occur instead under the N', the determiner position is in principle available for the null determiner. This makes LF(37) possible. There is independent evidence from English, moreover, that, while universal and definite DetPs can co-occur to form partitive NPs, universal and indefinite DetPs cannot. The contrast in (38) makes the point:

(38) a. Every one of the students read the book.
b. *Every one of some students read the book.

Again, the exact nature of the constraint responsible for the ill-formedness of (38b) needs to be investigated further. Notice, nonetheless, that if a similar constraint constrains the co-occurrence of *mei-ge* 'every', [+def] and [-def], we should expect that LFs of form (39), unlike LFs of form (37), should be impossible:

(39) *[$_{QP}$ mei-ge] [$_{NP}$ [-def] N'] S
every a

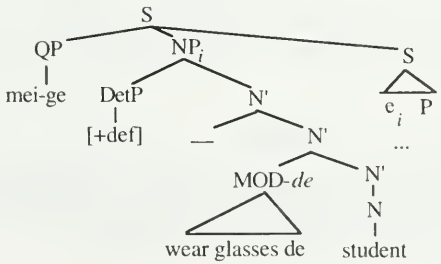
Let's now go back to the correlation between word order and partitive readings. The observation was that with the order in (15a) both partitive and non-partitive readings of the NP are possible, but with the order in (15b) only the partitive reading is possible:

(15) a. [$_{NP}$ QP VP-*de* N']
Lisi tongqing mei-ge dai yanjing de xuesheng.
Lisi sympathize every-cl wear glasses de student
'Lisi sympathizes with any student wearing glasses.'
'Lisi sympathizes with every one of the students wearing glasses.'

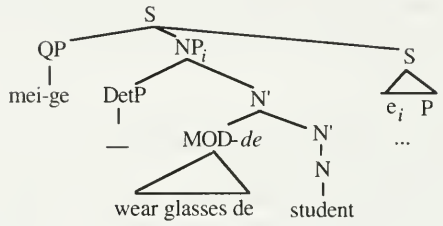
b. [$_{NP}$ VP-*de* QP N']
Lisi tongqing dai yanjing de mei-ge xuesheng.
Lisi sympathize wear glasses de every-cl student
'Lisi sympathizes with every one of the students wearing glasses.'

Given that co-occurrences of type (39) are barred, my analysis predicts that sentence (15a) should be ambiguous between LF (40a) and LF (40b), corresponding respectively to the case in which QP is base-generated under N' and to the case in which QP is generated in DetP position.

(40)a.



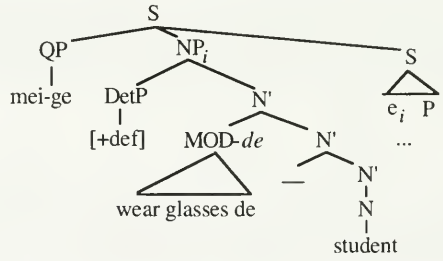
b.



(40a) in which *mei-ge* co-occurs with the definite determiner corresponds to the partitive reading for the NP in (15a), while (40b) corresponds to the standard (non-partitive) reading of universally quantified NPs assumed in Heim. Therefore, the theory predicts correctly that the NP in (15a) should have both a partitive and a non-partitive reading.

On the other hand, in (15b) the quantifier *mei-ge* is base generated under N' (since it is preceded by MOD-*de*). Thus, the determiner position will be occupied by a phonologically null determiner. Given the co-occurrence restriction in (39), however, only [+def] will be able to occur in DetP. Thus, only LF (41), which corresponds to the partitive reading of the NP, is possible for (15b).

(41)



Namely, the NP in (15b) is correctly predicted to have the partitive reading only.

8. Conclusion

I have shown that different interpretations of NPs containing numerals and prenominal modifier phrases correlate with the order of MP and MOD-*de* as well as with the syntactic categories in the MOD-*de*. The definiteness facts may fall out from the syntactic and semantic properties of the NP and its elements if we assume with Partee 1987 that numerals, hence MPs, may be adjectival or determinative, and from the assumption that NPs in MC may have a null determiner which, in principle, can be interpreted as definite or indefinite. The analysis also accounts for partitive and non-partitive readings of NPs containing MOD-*de* and QPs such as *mei-ge* 'every'.

NOTES

* I thank Profs. Alessandro Zucchi, Louise McNally, Anna Szabolcsi, and Chris Barker for comments on earlier versions of this paper, part of which was presented at the NACCL5 at the University of Delaware in May, 1993.

[1] Except for numerals and measure words, which I mark as a single unit with a hyphen, the pinyin orthography here is based on principles and rules by the Committees on Education and Languages (1988).

[2] Annear's observations also concern restrictive vs. non-restrictive readings of MOD-*de* where a demonstrative is present in the MP. These readings are also discussed in Chao 1968. The discussion in this paper concerns MPs without demonstratives.

[3] For the purpose of this paper it is not essential that I make a decision on whether these phrases are NPs or PPs, provided that we recognize that there is a syntactic distinction between prepositional and postpositional phrases. I shall follow, however, L & T's view that these locative phrases are NPs.

[4] An exception would be when MOD-*de* is MP-*de*, in which case MP-*de*-QP-N' sounds odd. Cases with demonstratives are not considered in this paper. It should be mentioned, however, that QP-NP-*de*-N' also sounds odd when a demonstrative is present in NP-*de*.

[5] The assumption that if MOD-*de* precedes MP or QP, then MP and QP are dominated by N' is shared by Huang (1982:67). Both numerals and quantificational words like 'every' are assigned to the syntactic category QP in Huang. I use MP and QP to reflect the different semantic properties of the two categories here.

[6] I will not try to address the contrast with respect to the collective/distributive distinction mentioned in section 3.2.

REFERENCES

ANNEAR, S.S. 1965. English and Mandarin Chinese: definite and indefinite determiners and modifying clause structures. Mimeograph.

CHAO, Y.R. 1968. A grammar of spoken Chinese. Berkeley, CA: University of California Press.

HEIM, Irene Roswitha. 1982. The semantics of definite and indefinite noun phrases. University of Massachusetts Ph.D. dissertation in linguistics.

——. 1990. E-type pronouns and donkey anaphora. Linguistics & Philosophy 13:2.137-177.

HUANG, C.-T. James. 1982. Logical relations in Chinese and the theory of grammar. MIT Ph.D. dissertation in linguistics.

——. 1983 On the representation of scope in Chinese. Journal of Chinese Linguistics 11:1.36-91.

——. 1987. Existential sentences in Chinese and (in)definiteness. The representation of (in)definiteness, ed. by Reuland & ter Meulen, 226-253. MA: The MIT Press.

LI, Charles N., and Sandra A. Thompson. 1981. Mandarin Chinese: a functional reference grammar. Berkeley, CA: University of California Press.

PARTEE, Barbara H. 1987. Noun phrase interpretation and type-shifting principles. Studies in discourse representation theory and the theory of

generalized quantifiers, ed. by J. Groenendijk, D. de Jongh, M. Stokhof, 115-143. Foris.
——. 1988. Many quantifiers. ESCOL '88 Proceedings.

STUDIES IN THE LINGUISTIC SCIENCES

The following issues are available

Vol. 19:1	Spring 1989	Papers in General Linguistics	$7.50
Vol. 19:2	Fall 1989	The Contribution of African Linguistics to Linguistic Theory, Vol. 1, ed. by Eyamba G. Bokamba.	$7.50
Vol. 20:1	Spring 1990	The Contribution of African Linguistics to Linguistic Theory, Vol. 2, ed. by Eyamba G. Bokamba.	$7.50
Vol. 20:2	Fall 1990	Linguistics for the Nineties: Papers from a Lecture Series in Celebration of the Department's Twenty-Fifth Anniversary, ed. by Hans Henrich Hock.	$7.50
Vol. 20:3	Spring 1991	Meeting Handbook: Thirteenth South Asian Languages Analysis Roundtable, 25 - 27 May 1991, University of Illinois at Urbana-Champaign, ed. by Hans Henrich Hock.	$5.00
Vol. 21:1	Spring 1991	Papers in General Linguistics	$7.50
Vol. 21:2	Fall 1991	Illinois Studies in Korean Linguistics 2, ed. by Chin-W. Kim.	$7.50
Vol. 22:1	Spring 1992	Papers in General Linguistics	$7.50
Vol. 22:2	Fall 1992	Twenty-Five Years of Linguistic Research and Teaching at the University of Illinois at Urbana-Champaign, ed. by Braj B. Kachru	$10.00
Vol. 23:1	Spring 1993	Papers in General Linguistics	$8.00
Vol. 23:2	Fall 1993	Papers in General Linguistics	$8.00
Vol. 24:1/2	Spr./Fall 1994	Proceedings of the 5th Annual Conference of the Formal Linguistics Society of Mid-America, ed. by James H. Yoon.	$16.00
Vol 25:2	Fall 1995	Language, Gender and Power (Editor: Yamuna Kachru)	$8.00

IN PREPARATION:

Vol. 25:1	Spring 1995	Papers in General Linguistics	$8.00
Vol. 26:1	Spring 1996	Illinois Studies in Korean Linguistics 3, ed. by James H. Yoon	$8.00
Vol. 26:2	Fall 1996		$8.00

Orders should be sent to:
SLS Subscriptions, Department of Linguistics
University of Illinois
4088 Foreign Languages Building
707 S. Mathews
Urbana, Illinois 61801

All checks and money orders must be in U.S. Dollars, drawn on a U.S. Bank, and made payable to University of Illinois. Sales Tax must be included as follows (except tax exempt organizations: **IL** $7\frac{1}{4}$%, **IN** 5%, **MI** 4%, **MN** 6%, **OH** 5%, **WI** 5%.

While quantities last, Volumes 8:1 through 17:2 are being offered at half-price.

Vol. 8:1	Spring 1978	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 8:2	Fall 1978	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 9:1	Spring 1979	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 9:2	Fall 1979	Relational Grammar and Semantics, ed. by Jerry L. Morgan.	$2.50	~~$5.00~~
Vol. 10:1	Spring 1980	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 10:2	Fall 1980	Studies in Arabic Linguistics, ed. by Michael J. Kenstowicz.	$2.50	~~$5.00~~
Vol. 11:1	Spring 1981	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 11:2	Fall 1981	Dimensions of South Asian Linguistics, ed. by Yamuna Kachru.	$2.50	~~$5.00~~
Vol. 12:1	Spring 1982	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 12:2	Fall 1982	Papers on Diachronic Syntax: Six Case Studies, ed. by Hans Henrich Hock.	$2.50	~~$5.00~~
Vol. 13:1	Spring 1983	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 13:2	Fall 1983	Studies in Language Variation: Mpmwestern Case Studies, ed. by Braj B. Kachru.	$2.50	~~$5.00~~
Vol. 14:1	Spring 1984	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 14:2	Fall 1984	Language in African Culture and Society, ed. by Eyamba G. Bokamba.	$2.50	~~$5.00~~
Vol. 15:1	Spring 1985	Papers in General Linguistics	$2.50	~~$5.00~~
Vol. 15:2	Fall 1985	Linguistic Studies in Memory of Theodore M. Lightner, ed. by Michael J. Kenstowicz.	$2.50	~~$5.00~~
Vol. 16:1	Spring 1986	Papers in General Linguistics	$3.00	~~$6.00~~
Vol. 16:2	Fall 1986	Illinois Studies in Korean Linguistics, ed. by Chin-W. Kim)	$3.00	~~$6.00~~
Vol. 17:1	Spring 1987	Papers from the 1986 South Asian Languages Analysis Roundtable, ed. by Hans Henrich Hock.	$3.00	~~$6.00~~
Vol. 17:2	Fall 1987	Papers in General Linguistics	$3.00	~~$6.00~~
Vol. 18:1	Spring 1988	Papers in General Linguistics	$3.00	~~$6.00~~
Vol. 18:2	Fall 1988	Papers in General Linguistics	$3.00	~~$6.00~~

STUDIES IN THE LINGUISTIC SCIENCES
VOLUME 23, NO. 2 FALL 1993
[PUBLISHED OCTOBER 1996]

PAPERS IN GENERAL LINGUISTICS

CONTENTS

Editor's Note and Correction iv

SAE-YOUN CHO: Auxiliary verb constructions in Korean 1

CASSANDRE CRESWELL: Criticizing with a question 25

ABDUL AZIZ DIOP: Vowel deletion in Pulaar: Rime and nuclear mergers and the issue of the syntax-phonology interface 33

HANS HENRICH HOCK: Subversion or convergence? The issue of pre-Vedic retroflexion reexamined 73

JONI KAY HURLEY: Request formation in Ecuadorian Quichua 117

NKONKO M. KAMWANGAMALU: Advancement in some Asian and African languages 137

EMMANUEL KWEKU OSAM: Animacy distinctions in Akan grammar 153

UTHAIWAN WONG-OPASI: The interplay between tone, stress, and syllabification in Thai 165

MARY A. WU: Adjectival and determinate measure phrases and NP interpretations in Mandarin Chinese 193

THE LIBRARY OF THE
NOV 07 1996
UNIVERSITY OF ILLINOIS
URBANA-CHAMPAIGN

UNIVERSITY OF ILLINOIS-URBANA

3 0112 033490373