SICOL

Proceedings of the Second International Conference on Oceanic Linguistics: vo1.1, Language Contact

Tent, J. and Mugler, F. editors. *SICOL- Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.*
C-141, ix + 146 pages. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.cover*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.

PACIFIC LINGUISTICS

FOUNDING EDITOR: Stephen A. Wurm

EDITORIAL BOARD: Malcolm D. Ross and Darrell T. Tryon (Managing Editors), Thomas E. Dutton, Nikolaus P. Himmelmann, Andrew K. Pawley

Pacific Linguistics is a publisher specialising in linguistic descriptions, dictionaries, atlases and other material on languages of the Pacific, the Philippines, Indonesia and southeast Asia. The authors and editors of Pacific Linguistics publications are drawn from a wide range of institutions around the world.

Pacific Linguistics is associated with the Research School of Pacific and Asian Studies at the Australian National University. Pacific Linguistics was established in 1963 through an initial grant from the Hunter Douglas Fund. It is a non-profit-making body financed largely from the sales of its books to libraries and individuals throughout the world, with some assistance from the School.

The Editorial Board of Pacific Linguistics is made up of the academic staff of the School's Department of Linguistics. The Board also appoints a body of editorial advisors drawn from the international community of linguists. Publications in Series A, B and C and textbooks in Series D are refereed by scholars with relevant expertise who are normally not members of the editorial board.

To date Pacific Linguistics has published over 400 volumes in four series:

- **Series A:** **Occasional Papers**; collections of shorter papers, usually on a single topic or area.
- **Series B:** **Monographs** of intermediate length.
- **Series C:** **Books**; publications of greater length, especially reference books such as dictionaries and grammars, and conference proceedings.
- **Series D:** **Special Publications**; including archival materials, pedagogical works, maps, audiovisual productions, and materials that do not fit into the other series.

Series C, Volume 141

SICOL

PROCEEDINGS OF THE SECOND INTERNATIONAL CONFERENCE ON OCEANIC LINGUISTICS: VOL.1, LANGUAGE CONTACT

edited by
Jan Tent and France Mugler

Pacific Linguistics

Research School of Pacific and Asian Studies
The Australian National University
Canberra

Published by Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University
PO Box 1428
Canberra ACT 2601
Australia

Copyright © The authors

First published 1998

National Library of Australia Cataloguing-in-Publication entry:

International Conference on Oceanic Linguistics (2nd: 1995: Suva, Fiji).
SICOL: proceedings of the Second International Conference on Oceanic Linguistics.

Bibliography.
ISSN 0078-7558
ISBN 0 85883 488 X (v. 1).

1. Oceanic languages. I. Tent, Jan. II. Mugler, France.
III. Australian National University.
Research School of Pacific and Asian Studies.
Pacific Linguistics. IV. Title. (Series: Pacific Linguistics.
Series C–141).

499

Copyedited by Joan Birnie
Typeset by Jeanette Coombes
Printed by ANU Printing Service, Canberra
Bound by F & M Perfect Bookbinding, Canberra

TABLE OF CONTENTS

CONTRIBUTORS TO THIS VOLUME

DAVID G. ARMS
Columban Fathers
Suva, Fiji

CHRIS CORNE
Department of Romance Languages
University of Auckland
New Zealand

GENEVIEVE ESCURE
Department of English
University of Minnesota
Minneapolis, USA

ANTHEA FRASER GUPTA
School of English
University of Leeds
Leeds, England
(Formerly: Department of English
Language and Literature
National University of Singapore)

WILLIAM JENNINGS
Department of Romance Languages
University of Waikato
Hamilton, New Zealand

ERNEST W. LEE
Summer Institute of Linguistics
Honiara, Solomon Islands

MIRIAM MEYERHOFF
Department of Linguistics
University of Pennsylvania
Philadelphia, USA

FRANCE MUGLER
Department of Literature and Language
The University of the South Pacific
Suva, Fiji

GEOFF P. SMITH
English Centre
The University of Hong Kong
Hong Kong

JAN TENT
School of English, Linguistics and Media
Macquarie University
Marsfield, Australia

PREFACE

The idea to hold the First International Conference on Oceanic Linguistics (FICOL) was first mooted in early 1992 by John Lynch and Paul Geraghty. FICOL was held in Port Vila at the Vanuatu complex of the University of the South Pacific from July 4 to 9, 1993. It attracted some 60 participants from all over the world and it was decided to hold subsequent conferences every two years. Proceedings were published by Pacific Linguistics in 1997.

The Second International Conference on Oceanic Linguistics (SICOL) was held in Suva, Fiji, on the Laucala Bay campus of the University of the South Pacific from July 3 to 7, 1995. More than 50 participants came from Australia, Canada, Fiji, France, Germany, Hawai'i, Hong Kong, Japan, the Netherlands, New Caledonia, New Zealand, Niue, Norway, Papua New Guinea, Singapore, Solomon Islands and the United States of America.

A special session on contact varieties was included in the program at the suggestion of Chris Corne and Jeff Siegel. The idea for this session came out of the 1993 Amsterdam meeting of the Society for Pidgin and Creole Linguistics, where Jeff Siegel and Bill Jennings noted that it would be good to have regular meetings of pidginists and creolists closer to the Pacific region. At the end of SICOL the Pacific Area Contact Linguistics Association (PACLA) was formed. While its members would continue to support both the Society of Pidgin and Creole Linguistics and the Society for Pidgins and Creoles in Melanesia, PACLA would provide a specific forum not only for Pacific pidgins and creoles, but also for other language contact phenomena in the region. A periodic newsletter would be edited by Chris Corne and future meetings would take place as part of the ICOL series.

This volume contains most of the papers presented at SICOL's session on language contact. While the collection is relatively small, it nevertheless reflects the enormous diversity of the field. Not surprisingly, most of the papers are about the Pacific and in particular Melanesia, the part of the region with the greatest linguistic diversity and a rich history of language contact. Other papers range farther afield, to Southeast Asia (Singapore English), the Americas (Belizean Creole), and the Caribbean (French Antillean creoles)—an area of the world of particular interest to creolists. Language varieties include of course pidgins, creoles and languages that are at once pidgins for some speakers and creoles for others, like Melanesian Pidgin. The volume also contains discussions of other languages which are the result of contact phenomena (e.g. the Fiji Hindi koinê) or have been greatly affected by contact (e.g. the fast disappearing Dravidian languages of Fiji). Some of the other issues these papers touch on include: structural features and their relationship with sociological factors (e.g. gender), classification and typology, historical origins, language change and language loss, regional varieties and controversies about how these relate to the standard, discourse features, language use and attitudes in a multilingual society, and computer-aided analysis of features of language use.

We thoroughly enjoyed organising and staging SICOL, as well as editing the papers in this volume. We sincerely hope that they will interest scholars and inspire further research, leading to worthwhile and entertaining discussions at future conferences.

Finally, we express our gratitude to the University of the South Pacific and its School of Humanities, without whose generous financial support SICOL would not have been possible.

Tent, J. and Mugler, F. editors. *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.*
C-141, ix + 146 pages. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.cover*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.

TENDENCIES IN FIJI HINDI

DAVID G. ARMS

1. INTRODUCTION

Fiji Hindi is a koinê (see Siegel 1987:188) and thus does not technically belong in a discussion dealing with pidgins and creoles. Nevertheless its particular history makes it of special interest to this area of study, and it would be a pity if such an important language of the Pacific were not considered at this time.

Pidgin Fiji Hindi is a different entity to the Fiji Hindi being discussed here, although the two have certainly influenced each other. Standard Hindi is different again, being the form of Hindi used in Fiji on formal occasions (to the degree the speaker has any knowledge of it). I will not discuss these varieties of Hindi, nor the sociolinguistics or history of Hindi in Fiji since they have been excellently covered by Siegel (1987).

My intention here is to expand a little—perhaps even correct—what is known about the phonology and grammar of Fiji Hindi. The main sources on this are Moag (1977), Siegel (1977, 1987) and Hobbs (1985).

It is often alleged that Fiji Hindi is amazingly uniform throughout Fiji considering the different origins of its speakers. I tend to agree with this opinion, but this uniformity should not be exaggerated. There are a few significant differences in vocabulary and verb inflections. Some rural areas and the northern coast of Vanua Levu (the second largest island) stand out as being different from the rest of Fiji. These, however, are very general observations. A good dialect study is needed to pinpoint where differences exist and their extent. The bulk of my own study was done in Naleba (on Vanua Levu, 16km north-east of Labasa), Suva, and, for short periods, Raviravi, between Ba and Lautoka, on the western side of Viti Levu.

2. PHONOLOGY

My first reaction in studying the phonology of Fiji Hindi was one of disappointment. With people from so many different linguistic backgrounds thrown together, I had thought that Fiji Hindi might have developed some exotic phonology of its own. In this expectation I had overlooked the fact that most of the indentured labourers came from the north-east of India and spoke different varieties of Hindi. It is basically the phonology of these varieties that Fiji Hindi has inherited. The presence of speakers of other Indian languages did not override this,

Jan Tent and France Mugler, eds *SICOL: proceedings of the Second International Conference on Oceanic Linguistics:* vol. 1, *Language Contact*, 1–10.
Pacific Linguistics, C–141, 1998.
© David G. Arms

Arms, D.G. "Tendencies in Fiji Hindi". In Tent, J. and Mugler, F. editors, *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.*
C-141:1-10. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.1*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.

although they may have contributed to some of the changes which have taken place or are taking place.

The Standard Hindi consonants are listed below in the order of traditional Hindi grammars:

Stops:						
velar		*k*	*k*^{h}	*g*	*g*^{h}	
palatal		*c*	*c*^{h}	*j*	*j*^{h}	
retroflex		*ʈ*	*ʈ*^{h}	*ɖ*	*ɖ*^{h}	*(ɳ)*
dental		*t*	*t*^{h}	*d*	*d*^{h}	*n*
labial		*p*	*p*^{h}	*b*	*b*^{h}	*m*
Fricatives:	*h*	*(x)*	*(ɣ)*	*(f)*		
Sibilants:	*ʃ*	*s*	*(z)*			
Liquids:	*ɾ*	*ɾ*^{h}	*r*	*l*		
Semivowels:	*y*	*ʋ*				
Extra stop:	*(q)*					

FIGURE 1: FIJI HINDI CONSONANT INVENTORY

The symbols in parentheses are what we might call secondary consonants—used only in careful speech, by learned speakers, or by speakers of Urdu. In Fiji Hindi, only the primary consonants are used, with the following exceptions:

(a) The secondary consonant [f] has completely replaced the primary one [p^h]. This is not surprising, and has also occurred in some dialects in India, while other dialects have the two sounds in free variation. They are certainly not in free variation in Fiji, but [f] has in some cases given way to unaspirated [p]:

(1)a. [həpta] 'week'

b. [fʊppa] 'father's sister's husband'

Note that in the second example above, [f] has been retained initially, but has changed to [p] medially. For some speakers the change of [f] to [p] takes place optionally in many vocabulary items.

(b) The secondary consonant [z] does occur occasionally, although it is usually replaced by [j]. It would probably have disappeared altogether if it were not for a few Urdu words and a few English borrowings. Even these latter are sometimes realized as [j], or occasionally [s].

(2)a. [bɪjji] 'busy'

b. [jesi] 'jersey, pullover'

(c) The primary consonant [ʃ] has merged with [s] for many speakers, especially in rural areas. Loss of [ʃ] could well be complete if it were not for the influence of English.

(3)a. [sadi] 'wedding'

b. [kosɪs] 'try, effort'

Fiji Hindi has basically the same vowel system as Standard Hindi. The ten vowels can be divided into two sets of five, which are related to each other orthographically in the Devanagari script and phonetically:

First:	*a*	*i*	*u*	*e*	*o*
Second:	*ə*	*ɪ*	*ʊ*	*æ*	*ɔ*
				(əɪ)	*(əʊ)*

FIGURE 2: FIJI HINDI VOWEL INVENTORY

In India, the pronunciation of the last two vowels differs according to dialect, being either monophthongal ([æ] and [ɔ]) or diphthongal ([əɪ] and [əʊ]). It is the latter diphthongal pronunciation which is current in Fiji. The first three vowels of the second set ([ə], [ɪ] and [ʊ]) are termed short vowels and do not occur word-finally in Fiji Hindi. The diphthongs [əɪ] and [əʊ] also do not occur word-finally except in a few monosyllabic words. In fact, they are infrequent, constituting only about 1 per cent of total vowel occurrences.

Stress seems largely predictable in Fiji Hindi, although the rule is complex and the matter requires further research. Stress most often occurs on the penultimate syllable:

(4)a. [pa'sina] 'sweat'

b. ['gob^{h}i] 'cabbage'

c. ['marɪs] 'he hit'

However, it occurs on the final syllable if it is a closed syllable with a long vowel:

(5)a. [bɪs'was] 'belief'

b. [məj'but] 'strong'

Stress also occurs on the final syllable (in disyllabic words), or on the antepenultimate syllable (in polysyllabic words) if the penultimate syllable is an open syllable with a short vowel and if the final syllable is also open:

(6)a. [sə'fa] 'clean'

b. ['kəɾʊwa] 'bitter'

Further study, hopefully, will clarify the issue of stress, and also the phonemic status of the short vowels [ɪ] and [ʊ] (vis-à-vis [i] and [u] respectively). If stress turns out to be phonemic, it is possible that these two short vowels are *not* phonemic, as their occurrence is very closely related to stress placement. Minimal pair candidates seem limited to final closed syllables, including monosyllables. Even in these cases, however, there may be a solution, at least from the point of view of an economic writing system: one could argue for writing the final consonant as doubled to indicate a final 'short' vowel. An alternative in that case to the transcription given in example (7) would be *pull* versus *pul.*

(7)a. [pʊl] 'bridge'

b. [pul] 'pool'

Fiji Hindi has, of course, been influenced by English. It has not taken over any sounds from English, but the high percentage of words borrowed does have its effect. English borrowings into Fiji Hindi are treated as in Standard Hindi except for the following :

(a) English [ə] and [ɔ] are not realised as Hindi [ə] and [ɔ] (which are pronounced as diphthongs in Fiji Hindi) but as [e] and [o]. This is a little unfortunate (even though quite natural, of course), as it creates some homonyms while leaving [ə] and [ɔ] grossly under-utilised:

(8) [plen] 'plan' (or) 'plane'

One does sometimes hear [pleɪn] for the latter meaning.

(b) English words ending in *-er* have that syllable realised as [-a] in Fiji Hindi, not (-ər]. The few words with the latter pronunciation in Fiji Hindi are English words that have come via India.

(9) [bəʈʈa] 'butter' but [moʈər] '(motor-)car'

English alveolar stops are realised as Fiji Hindi retroflex stops. In Standard Hindi there are considerably fewer words in the vocabulary employing the retroflex series than there are the dental. The high incidence of English borrowings in Fiji Hindi changes the situation and puts the two on a more even footing—giving Fiji Hindi a distinctive sound in the process, perhaps.

Interestingly, the Fijian phonemic system has a set of dental stops, not alveolar like English. Fiji Hindi faithfully borrows Fijian words with its dental series:

(10)a. [d̪akua] '*dakua*, kauri type'

b. [t̪anoa] '*tanoa*, *yaqona* bowl'

Occasionally, a Fijian word is heard with a retroflex stop. While hard to prove, it seems plausible that such a word may have come in through English:

(11) [ɖuruka], [ɖəruka] or [ɖuruka] '*duruka*, cane type'

Both Fijian and English [ð] are borrowed as dental [d̪]. English [θ] is borrowed as Hindi [tʰ] in the following example, and perhaps elsewhere:

(12) [tʰaʊsən] 'thousand'

It is interesting that, although linguists often illustrate in their phonetics classes the difference between aspirated and non-aspirated stops from English (e.g. [tʰɒp] versus [st=ɒp]) and although this difference is phonemic in Hindi, neither Standard Hindi nor Fiji Hindi speakers hear English initial stops as aspirated. They are always borrowed as non-aspirated stops.

While the phonology of Fiji Hindi does not deviate sharply from Standard Hindi, there are nevertheless some interesting tendencies. Some have been mentioned already, and others have been well covered by Moag (1977:281–285). In addition, note the following:

(a) Most content words beginning with [ʋ] have an alternative pronunciation with [b], which is preferred:

(13) [ʋɪcar] or [bɪcar] 'opinion'

(b) A dental articulation is sometimes replaced in the lexicon by a retroflex one. It is difficult to see why; the change seems unpredictable:

(14) [pəpiʈa] 'pawpaw' (< [pəpita])

(15) [bəkʈi] 'female goat' (< [bəkri])

3. GRAMMAR

Standard Hindi nouns are declined for number, gender and case. Adjectives and verbs in large part agree in number and gender with the nouns they modify. Much of this complex morphology has been simplified in Fiji Hindi and some has disappeared.

Whereas in Standard Hindi all nouns are treated grammatically as masculine or feminine, in Fiji Hindi gender distinction has been all but abolished. Human or animal females *may* take special feminine agreement, but even that is purely optional. If Standard Hindi were to lose its influence, one could predict, I think, that the feminine forms would disappear altogether. In the following example, the optional feminine form would be (əcc^{h}i].

(16) *I əcc^{h}a ləṛki hæ.*
this good girl is
She is a good girl.

Number has also been much modified. There is a special plural marker [log] or [logən] which is used to indicate plurality with words referring to humans and sometimes other animates. Otherwise plurality is left unmarked, or is expressed by quantifiers.[1]

(17) *ɔrət log awe hæ.*
woman plural come.PRES is
The women are coming.

Case has also disappeared. Nouns retain the same form whatever their function in the sentence. Their particular relationship to the action or event is conveyed chiefly by postpositions. However, a rather peculiar development has taken place. In Standard Hindi a masculine noun like, say, [rasta] 'road' would change to [raste] (the 'oblique' case, as it is called) before a postposition. In Fiji Hindi, the noun does not change in this way, but all the postpositions have alternative forms which happen to end in [-e]. Has some sort of case function ('oblique') invaded these postpositions?

TABLE 1: POSTPOSITION CHANGES

Standard Hindi	Fiji Hindi	
ka, ki, ke	*ke*	of the road
ko	*ke*	to the road
raste mẽ	*rasta me*	in the road
pər	*pe*	on the road
se	*se*	from the road
tək	*le*	until the road

One does not feel that the final [-e] occurring as it can on practically all the Fiji Hindi postpositions can be terribly important. Apparently, native speakers feel the same way, for after a noun ending in a vowel (and occasionally even after one ending in a consonant), the *-e* may be dropped (except in the case of [le]). This would leave [rasta m], for example, which may in fact become [rastəm]. The wheel has gone full circle, and another case system seems to be developing.

---

1 Abbreviations used in this paper are as follows:

EXP	expectational	INF	infinitive
HAB	habitual	PART	participle
IMP	imperative	PRES	present

I will now briefly look at verbs by considering a verb paradigm and then commenting on just a few parts of it. The forms on the left of the paradigm are the more common forms for Fiji Hindi, those on the right, labelled 'North', are substituted for the usual forms on the northern coast of Vanua Levu and probably also in some rural areas elsewhere. The numbers 1, 2, 3 are entered where forms differ for first, second and third person. Forms in square brackets are optional plurals. The verb [hæ] 'is', put in parentheses for the present tense, may be omitted. Alternative forms are separated by commas.

TABLE 2: PARADIGM OF VERB ENDINGS

		Usual	North
Independent Infinitive		*-o*	
Dependent Infinitive		*-e*	
Permissive Infinitive		*-ən, -ɪn*	
Future Imperative		*-na*	
Present Imperative		*-o*	
Subjunctive	1	*-i*	
	2	*-o*	
	3	*-e*	
Conditional		*-ta*	*-ta, ət, -te*
Present Participle		*-ət, te*	
Passive Participle		*-an*	
Future	1	*-ega*	*-ab, ɪb*
	2	*-ega*	*-ɪyo*
	3	*-i*	*-i (-ɪẽ)*
Present	1	*-ta (hæ)*	*-ɪt (hæ), -e (hæ)*
	2	*-ta (hæ)*	*-e (hæ)*
	3	*-e (hæ)*	*-e (hæ)*
(Extra Present)			*-ət hæ*
Past Continuous		*-ət rəha, -ta rəha*	*-ət reha*
Past Simple	1	*-a*	
	2	*-a*	
(transitive)	3	*-ɪs*, [-ɪn]	
(intransitive)	3	*-a, -ɪ*, [-ɪn]	
Past Habitual	1	*-i*	
	2	*-a*	
	3	*-i*	
Expectational		*-bə kəro*	

3.1 INFINITIVES

The [-o] infinitive form in Fiji Hindi, probably derived from the imperative form of Pidgin Hindi, as Siegel (1987:197) indicates, has become the stand-alone infinitive and dictionary entry for words in Fiji Hindi, replacing the Standard Hindi infinitive in [-na] for these

functions. The [-e] infinitive is termed dependent because it may occur only before a postposition or in construction with another verb (18). The [-ən] infinitive occurs, it would seem, only with the verb [deo] 'give', to indicate permission (19). Moag (1977:267) declares it is an option to the [-e] infinitive in this usage, but it seems to have taken over entirely for some speakers, at least in the north.

(18) *Həm i kam kare nəi sakɪt.*
I this work do.INF not can.PRES
I can't do this work.

(19) *Piən do.*
drink.INF give.IMP
Let him drink.

3.2 IMPERATIVES

It has been common to refer to the [-na] imperative as a 'polite' imperative. This is true, but I believe that its primary function is as a future imperative. Its politeness derives from the fact that to ask someone to do something in the future is less forceful than to ask them to do it now. Sometimes it is used more like a subordinate verb than a genuine imperative. Thus in the context of (20) below, the person was *not* being instructed to go to town by the first imperative; the first clause is more like a temporal clause, the second being a true imperative:

(20) *Taʊn jana təb lana.*
town go.IMP then bring.IMP
Bring it when you go to town.

3.3 PARTICIPLES

A use of the present participle in [-te] that is, I think, undocumented so far, is its occurrence with the negative, to give a rather elliptical negative present indicative (21); it provides a good example of how a new tense could develop. The passive participle ending [-an] is a form not found in Standard Hindi; it is sometimes used almost impersonally (22).

(21) *Ca pite nəi.*
tea drink.PART not
I don't drink tea.

(22) *Kʊc^{h} nəi dek^{h}an həme.*
something (is).not seen.PART to.me
I didn't see anything.

3.4 FUTURE

Siegel (1987:206–207) goes into much detail about the origin of the future tense forms of Fiji Hindi. He seems to imply there are no *-əb* or *-ɪb* first person future forms in Fiji, even though Moag (1977:219) mentions the latter. The forms do exist, however, at least in the north. These northern first person forms are, nevertheless, much more restricted in use (considered very 'rustic') than are the second and third person forms (including the plural—see §3.7), which are heard quite frequently.

3.5 PRESENT

The distinction of Standard Hindi between the simple present and the present continuous has been lost in Fiji Hindi. The extra set of present forms in the north, however, suggests that a separate present continuous tense could reconstitute itself. The inclusion or omission of the optional [hæ] could also take on meaning, but as far as I can tell, that has not yet happened.

3.6 PAST CONTINUOUS

This tense also does service as the past habitual (but see §3.8). The distinction between the two was lost in Fiji Hindi as just reported above for the present tense. The actual form used for the past continuous is reported by Moag as always being [-ət reha]. In the Suva area at least, the [-ta reha] form is the normal one for many speakers. This may reflect the influence of Standard Hindi on Fiji Hindi in the city. However, it makes sense to use the same participial form for the present and past continuous tenses, so Standard Hindi may not be entirely to blame.

3.7 PAST SIMPLE

The main point to note about this tense is that Moag, Siegel and Hobbs all treat its plural form as if it were *compulsory*. It is not. What's more, in some areas it is not commonly encountered and in fact occurs less than the optional future plural (see §3.4). Is this a change in usage, or was the original observation erroneous, which was then repeated by successive linguists? The latter seems unlikely, but ten years or so is a very short time for such a change to take hold. It can be noted, by the way, that the third person form for intransitive verbs is quite commonly [-is]. Whether this frequency represents an increase of its use, and thus a gradual shift towards regularisation, is difficult to say.

3.8 PAST HABITUAL

It was said above that the past continuous did service also for the past habitual (see Moag 1977:208, 221), and indeed the tense can be used in both senses. However, there is another set of forms, not recorded in the literature (to my knowledge), that has this specific function. What may have led to overlooking them is that the identical verb endings show up elsewhere in the verb paradigm, with different meanings. This is particularly problematic for the second person form, for it is identical with the simple past. I have had informants turn first or third person examples into second person, however, and they have regularly provided the [-a] form (rather than, say, the past continuous form). I conclude that it should be regarded as the third member of the set. The first and third person forms are the same as the subjunctive and future forms respectively, and thus have a future orientation that reminds one of the English habitual using 'would' (as in an alternative translation for example (23) 'We would work till two in the morning').

(23) *Həm log dʊi bəje rat le kam kəri.*
I plural two o'clock night till work do.HAB
We used to work until two in the morning.

(24) *Səb roj həmar gʰəre kʰana kʰai.*
every day my home food eat.HAB
He used to eat at my home every day.

3.9 EXPECTATIONAL

This unusual mood, which employs the suffix [-be] and the verb [kəro] 'do', is a formation not occurring in Standard Hindi. Moag (1977:211) presents it in a confusing and, it seems to me, mistaken way. Contrary to what he says, it can occur in a number of different tenses, not just one. A restriction I have observed, however, is that it seems to occur only in the negative.

(25) *Bolbe nəi kəro.*
speak.EXP not do.IMP
Don't answer him.

(The context here is a mother advising her daughter who is being teased by her brother; the advice is not to retort 'as he expects'.)

(26) *Abe nəi kərɪs.*
come.EXP not do.PAST
He did not come although (OR as) expected.

4. CONCLUDING REMARKS

I have looked selectively at a few grammatical features which seem of particular interest. Although Fiji Hindi has 'simplified' Hindi grammar, these changes have mainly involved simplification of the form of expression and the elimination of certain suppletive morphemes or constructions. There has been very little loss of semantic precision or flexibility, and indeed, as in §3.9, some convenient new turns of phrase have come into being. The most noticeable changes seem to be the disappearance of the distinction between the present simple and continuous forms and between the past habitual and continuous forms, and the collapse of the distinction between 'of' and 'to' for the Fiji Hindi postpositions (both being realised by [ke], see Table 3). There seems to be no trend to re-establish this latter contrast, but for the tense forms, the past habitual, as we saw in §3.8, has been reconstituted (perhaps recently) as a distinct tense, and it is possible that a similar reconstitution of a present tense could take place, or even *is* taking place (see §3.5).

The rate of change in Fiji Hindi appears to be quite fast. It is regrettable that more systematic study of Fiji Hindi did not begin at an earlier date; we might then have clearer evidence for certain changes and how they occur. Even so, from what we know already, it will be interesting to observe how current tendencies and other developments further evolve.

REFERENCES

Hobbs, Susan, 1985, *Fiji Hindi-English English-Fiji Hindi dictionary*. Suva: Ministry of Education.

Moag, Rodney F., 1977, *Fiji Hindi:a basic course and reference grammar*. Canberra: Australian National University Press.

Siegel, Jeff, 1977, *Say it in Fiji Hindi*. Sydney: Pacific Publications.

1987, *Language contact in a plantation environment: a sociolinguistic history of Fiji*. London: Cambridge University Press.

Arms, D.G. "Tendencies in Fiji Hindi". In Tent, J. and Mugler, F. editors, *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.*
C-141:1-10. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.1*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.

THE TYPOLOGY OF THE TAYO LANGUAGE OF ST LOUIS, NEW CALEDONIA

CHRIS CORNE

1. INTRODUCTION[1]

The aim of this paper is to identify and to describe briefly a number of structures in Tayo whose Melanesian origin can scarcely be doubted. Tayo did not exist prior to 1860, which was when the Marist mission was established at St Louis, some 15 kilometres from Noumea in the Far South of New Caledonia. In the village which grew alongside the Mission, there settled numbers of Melanesians, from different linguistic regions of New Caledonia, speaking related but mutually unintelligible languages. By around 1910, Tayo had become the sole first language (L_1) of children born as the third generation of the St Louis people (Ehrhart 1993:59–72). From a number of studies (Corne 1989, 1990a, 1990b; Ehrhart-Kneher & Corne 1996; and especially Ehrhart 1993; see also Chaudenson 1994; Ehrhart 1994, n.d.), it has become clear that we are in the fortunate position of being able to observe this language as it is currently spoken by the surviving members of the first generation for which it was the sole L_1 (as well as, of course, by subsequent generations). We have descriptions of the two main Melanesian languages spoken at St Louis in the first decades of the village's existence: Cèmuhî, the language of Touho (Rivierre 1980) and Drubéa (Shintani & Païta 1990), representative of the dialect group of the New Caledonian 'Far South' (see Rivierre 1973). There is as well an abundant literature concerning the New Caledonian Melanesian languages.

Tayo is a new language whose lexicon is mainly from the French spoken in New Caledonia, but whose semantic and syntactic organisation is fundamentally Melanesian in character.

Tayo continues to evolve, and there is still a range of variation which reflects in part the different tribal origins of the villagers (Ehrhart 1993:26–32). Within this variation, there is nonetheless a common core, and it is this core which is the baseline for the present investigation, which is more heuristic than definitive.[2]

---

1 This paper is a summary of my 1995 article entitled 'A contact-induced and vernacularised language: how Melanesian is Tayo?'. The research on which it is based was supported by the former Department of Romance Languages of the University of Auckland and by two research grants from the University of Auckland Research Committee (grants F3444075 and F3444079).

2 This study is based principally on the data collected by Ehrhart (1992a, 1992b, 1993), complemented by a tape recording made in 1978 by A.-G. Haudricourt and a few hours of fieldwork carried out by the author in 1988. My thanks go to Robert Carles, Sabine Ehrhart, Jim Hollyman, Vincent Holopopo,

Jan Tent and France Mugler, eds *SICOL: proceedings of the Second International Conference on Oceanic Linguistics:* vol.1, *Language Contact*, 11–26.
Pacific Linguistics, C–141, 1998.
© Chris Corne

Corne, C. "The typology of the Tayo language of St. Louis, New Caledonia". In Tent, J. and Mugler, F. editors, *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.*
C-141:11-26. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.11*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.

2. PERSONAL PRONOUNS AND SUBJECT INDICES

2.1 PRONOUN MORPHOLOGY

The forms of the pronouns are given below. Some variants are omitted from this table, as are the possessives, these being dealt with in the text.

TABLE: TAYO PERSONAL PRONOUNS

	Subject and subject index	Independent pronoun	Unmarked dependent and subject index
Singular			
1	*ma*	*mwa*	
2	*ta*	*twa*	
3	*la*	*lia, sa*	
Dual			
1	*nude*[3]	*nude*	*le*
2	*ude*	*ude*	
3	*lede*	*lede*	
Plural			
1	*nu*	*nu*	
2	*uso*	*uso*	
3	*sa, sola, lesot*	*sola, lesot*	

These forms have a number of variants, a partial listing of which follows. For the subject forms:[4]

(a) 1SG *ma* may become *m* before the vowel *a*; *mwa* is a relatively frequent form of the subject among older speakers, while *mã* or *mõ* (with nasal vowels) occurs only rarely and in the usage of the oldest speakers.

(b) 2SG *ta* may become *t* before *a*; *tã*, *tõ*—for the oldest speakers.

(c) 3SG *la* may become *l* before *a*; *lia-la* is a demonstrative pronoun (singular), and an emphatic form used in thematisation.

The dual is used only by older (40+) speakers:

(d) 1DU *nude* (inclusive and exclusive) has a form *nudø* which is felt to be "more modern"; *nude tu sel* 'exclusive dual'.

Joseph Katé; to all those in Noumea and at St Louis who helped in various ways; and to Peter Mühlhäusler who organised some financial assistance.

3 Voiced stops are usually prenasalised in Tayo.

4 Abbreviations used in this paper are as follows:

CONJ	conjunctive	PREP	preposition
DEF	definite	REL	relativiser
DU	dual	SG	singular
PL	plural	SI	subject-indexing pronoun

(e) 2DU *ude* (with or without prenasalisation of the *d*) has variants *udø*, *usot de*, *bude*; formerly, the 2DU was used as a form of respect when addressing an older person or a person holding tribal authority.

(f) 3DU *lede*, also *ledø*, *de*, *de-la*.

(g) 1PL *nu* may become *n* before the reduced form *a* of the future marker *va*.

(h) 2PL *(v)usot*, *uso*, *su*, according to speed of speech.

(i) 3PL *sot-la* is an archaic form remembered only by the oldest speakers; younger speakers prefer *lesot* to *sola*; *sa* has the variants *swa*, and *s* before *a*.

For the independent forms, note that *lia* (disyllabic) is usually *lya* (monosyllabic); *lia* is [3SG, +animate], *sa* is [3SG, -animate]; and that *usot* is in free variation with *uso*.

These observations show clearly that some variation is a function of evolution over time. Before examining this aspect, let us see how the system works.

2.2 THE SUBJECT AND SUBJECT-INDEXING PRONOUN

The subject pronoun is used either alone, preposed to the predicate, or as an index within VP when the subject is an independent pronoun or an NP (a lexical agent)—see §2.6:

(1) *Lesot sa vya war mwa e sola sa reste.*
3PL 3PL.SI come see 1SG and 3PL 3PL.SI stay
They came to see me and they (another group) stayed behind.[5]

(2) *Tule per sola arive de laba.*
PL priest 3PL.SI arrive from there
The priests came from there.

In (1), *lesot* and *sola* are independent pronoun subjects, while *sa* is a subject index; in (2), it is *sola* which is acting as the subject index. Note that *sa* '3PL, +animate' is never used as an independent pronoun, whereas *sa* '3SG, -animate' is so used.

2.3 THE INDEPENDENT PRONOUN

The independent pronoun is often used as the subject, generally but not obligatorily with a subject index. It may be in initial position, as in (1), or final, as in (3) below, according to the class of verb (minimally, ± stative), and it may be thematised in various ways, to be discussed below.

The independent pronoun is also used in non-subject functions as a direct object (*mwa* in (1) above) or as the object of a preposition (as in the possessive, see §2.5).

5 Tayo has few markers of tense/aspect, and they are always optional. For details, see Ehrhart (1992b, 1993:159–165).

2.4 THE DEPENDENT PRONOUN *le*

This pronoun is unmarked for number and person, and is always a subject index. Impersonal predicates have no subject: there are very few of these, but they are relatively frequent, and include:

(i) the existential predicates/verbs *na* and *napa* (see §6);

(ii) terms having to do with natural and meteorological phenomena such as *bota* 'be fine', *cho* 'be hot', *frwa* 'be cold', *fenwar* 'be dark/night'.

With (and only with) the second of these two categories, *le* appears to be usual (perhaps obligatory), and can be seen either as indexing the implied subject, or as an impersonal pronoun. In all other contexts it has the indexing function.

The seven contexts involved are:

(a) with stative predicates; the subject (independent pronoun or NP) is postposed in this context:

(3) *Le fu lia.*
SI crazy 3SG
He is crazy.

(b) when the subject precedes the predicate; a noun subject is usually specified by a demonstrative *-la* or a possessive; *le* has a thematising (emphasis, focus, demonstrative,...) function:

(4) *Lia le fu.*
3SG SI crazy
HE is crazy.

(c) when an independent pronoun is (redundantly(?)) focused by *se*:

(5) *Se twa le fe sa.*
it.is 2SG SI do that
You're the one who did it.

(d) after the relativiser *sa* (relativisation of the subject):

(6) *Ma bwar dolo-la sa le sal.*
1SG drink water-DEF REL SI dirty
I drank the dirty water.

(e) after a noun phrase (but not a personal pronoun) which is the subject within a relative clause introduced by *sa* (relativisation of a non-subject):

(7) *tule guyav sa wawa le pla:te*
Pl guava REL grandmother SI plant
the guava trees that Grandma planted (*wawa* < Dru)

Le is also used after certain non-personal and/or indefinite pronouns, and in some interrogative structures (see Ehrhart 1993:175–176; see also §4 below).

2.5 POSSESSIVES

The possessive is constructed with the preposition *pu(r)* + the independent pronoun. Sandhi phenomena produce reduced variants: *pur twa* > *pu twa* > *puta* > *pta*.

(8) *mater pu bude*
mother PREP 2DU
your (2DU) mother

2.6 DISCUSSION

The morphology of this system clearly comes from French forms, viz. the stressed (tonic, disjunctive) pronouns and similar forms which occur in spoken French, sometimes with the tag *là* added: *moi*, *toi*, *lui-là*, *nous*, *vous autres*, *(le)s autres(-là)*, *ça*. The dual, an obligatory semantic category in both Drubéa (Shintani & Païta 1990:57–58) and Cèmuhî (Rivierre 1980:61, 119–122), is added, but again the forms are derived from French: *nous deux*, *nous deux tous seuls*, *vous (autres) deux*, *(les) deux(-là)*. The indexing pronoun *le* is probably the result of a reanalysis of French *(i)l est* and 3SG *la* may well have a similar origin in French *(i)l a*. To be noted are the loss of the front rounded vowel [ø] and the presence of the (Melanesian) prenasalised voiced stops [ᵐb, ⁿd]. Phonetically, there has been some evolution, with the erosion of some forms and the (re)introduction of [ø]. Apart from the prenasalised stops, perhaps, there is nothing particularly noteworthy about the morphology of this system.

The way the system works, however, is strikingly different from either French or any other French-lexicon creole. In the dual, there is a near total absence of any distinction between inclusive and exclusive, although this distinction is obligatory in Drubéa and Cèmuhî, and no evidence that such a distinction has ever been a part of the Tayo pronominal system. However, the dual itself is now being lost. There is thus an on-going evolution towards the bipolar SG/PL system of French.

This evolution has little effect on the status of the pronouns, divided as they are into the two classes of dependent and independent elements. The independent pronouns are, variously, preposed agents, postposed objects, postposed possessives, and postposed statives (as in (3) above). These functions are identical to those of Cèmuhî, although the latter uses different classes of (usually affixed) personal 'modalities' to handle them (Rivierre 1980:61, 121). As for the apparently double role of the subject/subject-indexing pronouns, it is a direct reflection of Cèmuhî; there, the preposed personal modalities have the following features (among others): they are not clitics, since (as in Tayo) other particles (of tense, aspect, etc.) usually occur between them and the verb; they are within VP and normally precede the verb when it is followed by a subject; they are obligatory if the lexical subject is not expressed (Rivierre 1980:119). The multiple role of the Tayo independents is the same as in Drubéa, which has only two modalities, subject and non-subject (Shintani & Païta 1990:58, 69). (The honorific role of 2DU *ude* has been mentioned in §2.1 above: in Drubéa, 2PL has this role.) For possessives, Cèmuhî and Drubéa each has a (different) system of junctors divided into various semantic categories (Rivierre 1980:68, 151–156; Shintani & Païta 1990:55): in Tayo, all these categories are collapsed into the single paradigm using *pu*. The Tayo system is clearly the result of the transfer of Melanesian conceptions, those conceptions which Drubéa and Cèmuhî hold in common: conceptions, not forms, the forms in each having only a rather distant etymological connection.

To this transfer there has been added an innovation, the dependent *le*. This indexing pronoun is unmarked for person or number and is confined to certain contexts.

The pronominal system as sketched here offers a preliminary view of the nature of this language which arose in a plurilingual contact situation. The morphology is entirely derived from French forms, either directly or through reanalysis of French strings, while the semantic organisation of the system is essentially Melanesian. This organisation is reflected in an equally Melanesian syntax based on what Cèmuhî and Drubéa have in common, but there are innovations. The one concerned here (*le*) is one such which, while it conforms syntactically and semantically to the general system, seems not to have any direct, single model; and there is an on-going evolution which, in the pronominal system, affects:

(a) the dual (loss of a distinction of marginal import in French, or 'decreolisation', i.e. movement towards a bipolar system closer to French);

(b) the morphology (phonetic erosion: *vuzot* > *(v)usot* > *uso* > (*?*so*) > *su*, plus decreolisation: loss of prenasalisation of voiced stops, (re)appearance of rounded front vowels);

(c) the syntax (*se* for *le*, see §7).

It does not seem possible to speak here of any sort of *modification* of French, be it by simplification or by the accelerated evolution of tendencies inherent to spoken French. Nor is it a matter of the straightforward relexification of a Melanesian language, since the Tayo pronominal system shows a categorisation of the pronouns which is less complex than in Cèmuhî (which has different forms for the preposed agent, the postposed object, the possessive, and the stative) but more complex than in Drubéa (by virtue of the indexing role of *le*). The Tayo pronominal system is thus a new creation whose general inspiration is clearly Melanesian but the detail of which is unique to Tayo. Moreover, the usage of the different age groups at St Louis today shows a system which is changing, at least partly because of French/Tayo bilingualism.

These conclusions are based on a small subset of Tayo's grammar. The examination of other areas of grammar will allow an assessment of the accuracy or otherwise of the picture outlined thus far.

3. RELATIVISATION AND THEMATISATION

These two procedures are to some extent linked in Tayo because thematisation may be handled by clefting using a relative clause, as happens in French (*c'est à Laetitia qu'il s'adressera*, for example).[6]

3.1 RELATIVE CLAUSES

The relative construction has always, except in the one case to be discussed in §3.2, the subordinator/relativiser *sa*: antecedent + *sa* + clause. Subject and non-subject relativisation

6 The presentation here is a brief summary of Corne (1994), which is itself based on data culled from Ehrhart (1993).

are slightly different. In subject relativisation, there is obligatorily an indexing pronoun, either *le* or 3SG/DU?/PL,[7] preposed to the verb in the subordinate clause:

(9) *Ma war loto sa le vya.*
1SG see car REL SI come
I see the car coming.

(10) *fiy-la sa la okipe de nu*
girl.DEF REL 3SG look.after PREP 1PL
the girl who looked after us

With relativisation of an object, there are two cases: if the subject of the relative clause is a pronoun (11), or if it is a noun (12):

(11) *Le bon gato sa sa fe.*
SI good cake REL 3PL do
The cake they made is good.

(12) *tule guyav sa wawa le pla:te*
the guava trees that Grandma planted (= (7))

Here, *le* is obligatory as the subject index.

Locatives and temporals may follow the non-subject pattern, although the use of a presentative element *se* is more usual (see §3.2):

(13) *(se) la sa ta ete war a:ba*
there REL 2SG been see down.there
over there where you went to look

In summary, relative clauses are subordinated using *sa*. In all cases, a pronoun is required, either as subject or as subject index. Relativisation appears to concern only subjects and direct objects, and occasionally complements of place and time.

3.2 THEMATISATION

There are a number of thematising procedures which front a subject (an agent). One such has an optional presentative *se* and an obligatory NP + *le*:

(14) *(Se) twa le fe sa.*
You're the one who did it. (= (5))

Locatives are usually thematised by *se* and specified by a relative clause. Another thematisation uses *wala* + NP + *le*:

(15) *Wala per le desan.*
there.is priest SI descend
Here comes the priest.

---

7 The dual is not attested in the data currently available. The nature of the difference (if any) between *le* on the one hand and 3SG/DU/PL on the other is not as yet clear: *le* seems preferred when the embedded predicate is stative, the third person pronoun elsewhere.

These procedures have a competitor in a relativisation which follows the French clefting model, using *se* + NP + *ki*:[8]

(16) *Se lya ki fe sa.*
it.is 3SG REL do that
It is he who did it.

To sum up, Tayo thematisation concerns essentially an NP subject + *le*, and uses a relativisation strategy only with locatives. But there is also a competing relativisation strategy, with *ki* and without *le*, which occurs only in the context of thematisation and which is clearly French in inspiration.

3.3 DISTRIBUTION

According to the data, relativisation conforms to a near-universal hierarchy. Tayo relativises easily on subjects, almost as easily on direct objects, and occasionally on locatives and temporals: SU > DO > LOC/TEMP. Thematisation, in contrast, be it with (*se/wala*) + NP + *le* or with *se* + NP + *ki*, concerns only subjects, the first pattern being the more frequent.

3.4 DISCUSSION

If the origin of *le* can be ascribed with some certainty to a reanalysis of French *il est*, its use in relativisation and thematisation illustrates perfectly its present status, quite unconnected with any French model. In examples (9)–(15), where no other pronoun fills the role of subject or subject index, *le* is required, the relativiser *sa* being a conjunction, not a pronoun. But if (13) is contrasted with (16):

(13) *(se) twa le fe sa*

(16) *se lya ki fe sa*

it is clear that *ki*—a French influence—is indeed a pronoun in Tayo. There is therefore a change in the direction of French: currently well established in Tayo, relativisation with *ki* will quite possibly in the future introduce further perturbations in the original relativisation and thematisation patterns (as is suggested also by the form *ka*, see fn.7).

These original systems show, even more clearly than the pronominal system does, the extent to which Tayo derives from Melanesian grammatical and semantic conceptions.

In Cèmuhî, there are two relativisation strategies which use as the relativiser either one of three semantically differentiated demonstratives (deixis or localisation of the antecedent), or one of three semantically differentiated relative pronouns derived from articles (reference to an antecedent which is [± definite]). In both cases, there are agreement rules; relative clauses

---

8 There is also a form *ka* (*se lya ka met* 'it is he who puts...'), and some speakers allow non-subject thematisation with *ke* (*se mwa ke la ule war* 'it is I whom s/he wants to see') see Ehrhart (1993:153). The existential impersonal verbs *na*, *napa* may also occur in what appear to be thematising and/or focusing roles; the various constructions observed in the data include *na* NP *le* VP, *na* NP VP, *na* NP *ke/ki/ka/sa le* VP; while some of these appear to be prompted by French patterns, there are at present insufficient data available to allow a clear view of the exact semantic (if any) or sociolinguistic correlates of these constructions.

follow the antecedent; subject relativisation allows a following subject-indexing pronoun. Subjects and direct objects may be relativised (Rivierre 1980:172–176).

In Drubéa, subject relative pronouns are identical in form to the 3SG/DU/PL personal pronouns, and there is a fourth form unmarked for person/number; direct object relative pronouns have different forms, and are optional; there are agreement rules; relative clauses follow the antecedent; there are no subject-indexing rules (Shintani & Païta 1990:83–86).

Emphasis on a constituent in Cèmuhî may be handled by clefting and relativisation (Rivierre 1980:207, n.71), but this is 'much less frequent' than 'thematisation' (Rivierre 1980:207–212). Thematisation, in essence, fronts subjects or direct objects, the remainder of the sentence (the 'comment') being coordinated (not subordinated) and obligatorily containing a resumptive personal modality. Sentential adverbs of time and place, when fronted, do not have such a resumptive, but locative complements of the verb do. Any fronted element may be preceded by a presentative.

Thematisation in Drubéa (Shintani & Païta 1990:107) is less complex: a presentative is used, and while relativisation may occur, resumptive pronouns are generally absent. Only subjects are thematised.

These systems are considerably more complex than that of Tayo. Nevertheless, for Tayo relative clauses the parallels are obvious. Subject indexing is obligatory. Subject and direct object relativisation are allowed. Given that non-relative *sa* is, variously, 3PL animate subject, 3SG inanimate subject and non-subject ('this, that, it'), and, with postposed *-la*, the inanimate demonstrative pronoun (usually singular), its selection as the relative pronoun is unsurprising.

Thematisation in Tayo, with the exception of the *se* + NP + *ki* clefting pattern, displays some features reminiscent of Cèmuhî thematisation: subject indexing, an optional presentative, and sometimes (for locatives) emphasis through relativisation. Cèmuhî can front direct objects, while Drubéa cannot: this may account for Tayo's resistance to the fronting of anything other than the subject.

It is clear that Tayo's relativisation and thematisation strategies have a Melanesian architecture: they reflect Melanesian constructions, adapting the common ground and smoothing over a lot of the complex and conflicting detail.

In contrast, the clefting of the subject + relativisation with *ki* (and *ka*) is a French structure, but one which is contextually constrained in Tayo in that it applies only to subjects. (In French, clefting can occur with most constituents of a sentence.) This French-inspired construction is competing with the Melanesian-inspired thematisation using the subject index *le* and the presentative *se* (this last is of course directly derived from French *c'est*, but its optional status in Tayo parallels the similarly optional status of the Cèmuhî presentative (Rivierre 1980:211). Ehrhart (1993:152) informs us that *ki* is "more modern" than *sa* (a matter we shall return to in due course), but it must be pointed out here that the model is that of formal French (*c'est moi qui* VP) rather than of popular French (*c'est moi que je* VP).

Although the relative clause constructions characteristic of popular French (Frei 1929:183–191; Guiraud 1966) are not especially a feature of New Caledonian French (NCF), one could be tempted to try to establish some link between popular, spoken French and Tayo. In fact, it is clear that the relative clauses of Tayo do not follow any French model, formal or popular. Popular French allows, for example, *l'homme qu'il arrive*—invariable *que* + subject pronoun; it also allows *l'homme que je l'ai vu*—*que* + object pronoun. Thus,

Tayo agrees superficially and accidentally with popular French as far as subject relativisation is concerned, but refuses the object pronoun (**mek-la sa ma war lya* 'the man that I saw (him)'). Furthermore, neither formal nor popular French offers any explanation for the gaps that exist in the distribution of Tayo relative clauses.

In summary, then, Tayo relativisation has nothing to do at all with French, except for the *se* NP *ki* VP clefting pattern; even this last is constrained with respect to its domain (subject only) and is in competition with the Melanesian type of thematisation.

The situation here is thus identical to the one described earlier for the pronominal system. The relativisations and thematisations of Tayo are not modifications of French patterns, nor straightforward relexifications of Melanesian constructions; they are new patterns with a general Melanesian configuration but whose detail is peculiar to Tayo. The modern data, in addition, show ongoing changes which derive, at least in part, from Tayo/French bilingualism.

Thus far, we have seen two areas of Tayo grammar, each of which shows the language in the same light. The next three areas are covered a little more superficially, pending further investigation.

4. INTERROGATION

Only interrogation bearing on NP, and the interrogations 'where' and 'when' are considered here. The interrogatives are: *(se) ki* 'who(m)', *(se) kwa* 'what', *u* 'where', *ka* 'when'.

(a) [+ animate], subject: *se ki le* V; object: *se ki* NP V and NP V *ki*; statives: *se ki* NP and NP *se ki*; object of a preposition: Prep *ki*:

(17) *Se ki le vya laba?*
it.is who SI come over.there
Who's (that) coming?

(18)a. *Se ki ta war?*
it.is who 2SG see
Whom did you see?

b. *Ta war ki?*
2SG see who
Whom did you see?

(19) *Frer pu ta se ki?*
brother PREP 2SG it.is who
Who is your brother?

(20) *Sa pa kone se peti pu ki.*
3PL NEG know it.is child PREP who
They did not know to whom the child belonged.

There are variants, including relative clauses, sometimes with *sa* but also with the relativiser *ki* or with *(se) ki ki*, *ki se ki*, and other more-or-less French-inspired formulae.

(b) [- animate], subject: *se kwa sa le* V; object: V *kwa* and *se kwa sa* NP V; object of a preposition: Prep *kwa*:

(21) *Se kwa sa le to:be?*
it.is what REL SI fall
What fell?

(22)a. *Ta fe kwa?*
2SG do what
What are you doing?

b. *Se kwa sa la di?*
it.is what REL 3SG say
What did he say?

(23) *Ta kone ma rigole de kwa?*
2SG know 1SG laugh PREP what
Do you know at what (why) I am laughing?

(c) the locative *u* always follows the verb:

(24) *N a bare u?*
1PL FUT go where
Where will we go? (*bare* < NCF *barrer* 'go, leave')

but can also be the predicate head:

(25) *Le u lia?*
SI where 3SG
Where is he?

de u is always final:

(26) *Sola kuver ave lapay de u?*
3PL cover PREP straw from where
They cover the roof with straw from where?

(d) the temporal *ka* may be initial or final; *se ka sa* + relative clause also occurs:

(27)a. *La vya ka?*
3SG come when
When is she coming?

b. *Ka la vya?*
when 3SG come
When is she coming?

(28) *Se ka sa sola arive isi?*
it.is when REL 3SG arrive here
When did they arrive (here)?

With respect to the position of the interrogative word(s), most of these structures reflect either Cèmuhî or Drubéa or both:

(i) [+ animate], subject: initial in Dru (Shintani & Païta 1990:96) and in Cèm (Rivierre 1980:218); object: not known, but probably final; stative: final in Cèm (Rivierre 1980:218); Prep + Q: final in Dru (Shintani & Païta 1990:96) and in Cèm (Rivierre 1980:218).

(ii) [- animate], subject: initial in Dru (Shintani & Païta 1990:95) and in Cèm (Rivierre 1980:219); object: final in Dru (Shintani & Païta 1990:95) and in Cèm (Rivierre 1980:219); Prep + Q: final in Cèm (Rivierre 1980:219).

(iii) locative: always final in Dru (Shintani & Païta 1990:96), usually final in Cèm (Rivierre 1980:220).

(iv) temporal: initial and final in Dru (Shintani & Païta 1990:96) and in Cèm (Rivierre 1980:220; for the equivalent of *se ka sa* in Cèm, see Rivierre 1980:221).

As can be seen, the parallels are striking. Nonetheless, the conclusion that yet again Melanesian patterns are the direct motivations for the Tayo constructions must be qualified by the observation that in spoken New Caledonian French the same word orders occur. Comparative studies of interrogative word order in New Caledonian and popular French might perhaps allow an estimation of the extent of Melanesian influence on the local New Caledonian French, but for the moment all that can be done is to note the fact that there are parallels between Tayo and Cèm/Dru, and some differences (e.g. example (22b)) between Tayo and its Melanesian predecessors at St Louis.

5. IMPERATIVES

A declarative sentence may serve as an imperative with the appropriate intonation contour and in an appropriate (real-life) context; a particle *ko* may support such an imperative:

(29)a. *Ta fe bya!*
2SG do well
Be good!

b. *Ta ko fe bya!*
2SG *ko* do well
Be very good!

A subjectless verb may be used for the 2SG/PL imperative:

(30) *Pitchare lya!*
pinch 3SG
Pinch him/her! (*pitchare* < Dru)

and the independent pronoun may be pre- or postposed:

(31)a. *Muche, twa!*
blow.nose 2SG
Blow/Wipe your nose (, you)!

b. *Usot, desa:de de la!*
2PL descend from there
Get down from there, you lot!

In the first person plural, the use of the pronoun *nu* is obligatory, always preposed; the Future marker *va* is often used in this context (*nu va* > *nu a* > *na*):

(32) *Nu (va) asi!*
1PL (FUT) sit
Let's sit down!

There are other procedures (see Ehrhart 1993:177–179 for details), among which the use of *fo* 'be necessary' may be noted:

(33) *Fo sach!*
must well.behaved
Be good!

In the negative, *pa* is preposed to the verb, while *fo* > *fopa (ke)*:

(34)a. *Fopa parle kom sa.*
must.NEG speak like that
Don't speak like that.

b. *Fopa ke ta fe sa.*
must.NEG CONJ 2SG do that
Don't do that.[9]

(35) *Pa tuche!*
NEG.touch
Don't touch!

Another frequently used construction has *napa beswa de* 'there is no need to':

(36) *Napa beswa de di no pu lya.*
have.NEG need PREP say name PREP 3SG
Don't say his/her name.

These data may be briefly compared with imperatives in Cèmuhî and Drubéa. In the latter (Shintani & Païta 1990:97–98), a subject pronoun is optional; a verb equivalent to *fopa* is used sentence-initially, while another, equivalent to *napa beswa de*, is used predicate-initially; an element postposed to the verb can reinforce an imperative, cf. preposed *ko* in (29b). In Cèmuhî, a preposed personal modality (a pronoun) is optional, and there are two different verbs which express prohibition (negative imperative) (Rivierre 1980:222–223). Once again, a number of similarities can be seen between Tayo and the two Melanesian languages. All of the Tayo structures have an equivalent in French, to be sure, but a frequently used Tayo procedure, illustrated by (29) and (32), corresponds to a rather formal register in French (*tu seras sage, mon enfant!; nous nous assiérons, Messieurs, s'il vous plaît!*), while other Tayo imperatives correspond to more relaxed styles. The significant thing here is that Tayo has selected, from the range of French models putatively available, precisely those which correspond to Melanesian notions of ordering and forbidding.

6. THE IMPERSONAL INDEPENDENT VERB

The verb *na*, and its negative counterpart *napa*, are existential verbs which do not require a preposed personal modality. They have a number of variants which Ehrhart (1993:173) considers "decreolized": *ya*, *yana* in the present, *nave* and *yanave* in the past, *(ya)navepa* for the negative of the past.

The verb *na* is used for the notion 'to have, possess':

---

9 In other contexts, the same sentence could have other readings ('you mustn't do that', 'you shouldn't have done that', ...).

(37) *Na boku kusa pu mwa ke twa.*
have many cousin PREP 1SG than 2SG
I have more cousins than you do.

as well as that of 'be present':

(38) *Napa Childa? napa Childa.*
have.NEG Gilda
Is Gilda there? No.

and 'there is/are/etc.':

(39) *Na dolo partu.*
have water everywhere
There is/was water everywhere.

The Melanesian architecture seems very clear here. In Cèmuhî (Rivierre 1980:62, 70, 213, 215) there is an impersonal verb (three forms, 'there is', 'there isn't' and 'there is no longer'), but the meaning 'to have, possess' is usually expressed by a postposed personal modality (*two/the-2/child/of-him* 'he has two children'), to which may be optionally preposed a particle (*(there-is)/fear/of-him* 'he is afraid' [he has fear]). In Drubéa (Shintani & Païta 1990:68, 93), the match with Tayo is even closer: the same impersonal verb, in the affirmative or the negative, expresses 'there is' (39) and 'have, possess' (37), and in the negative 'there is not' and 'not be present' (38) (*there-is-not/of-you* 'you were not there'). The semantics of (37)–(39) do not come from French.[10]

7. DISCUSSION AND CONCLUSIONS

A large number of potentially revealing phenomena are not dealt with in this paper. For example, there are few adjectives in Tayo which are postposed to the noun: Ehrhart (1993:145–146) says that postposed adjectives seem to be an open class, but the usual strategy is relativisation (*dolo sa le sal* 'the water which is dirty/the dirty water'); as it happens, in Cèmuhî postposed stative verbs—the local equivalent of adjectives—are an abbreviated form of relative clause through deletion of an invariably recoverable relativiser (Rivierre 1980:143, 160–161); in Drubéa, the adjective follows the noun (Shintani & Païta 1990:54–55); both languages have however a short inventory (different in each) of preposed nominal determinants. Nothing has been said here about comparatives, little about the word order, nothing on the aspect markers (but see Corne 1993; Ehrhart 1992a, 1993:159–165), nothing on negatives. Yet these are all areas of interest: it is not, for example, impossible that the contrastive study of the tense/aspect systems of Drubéa (Shintani & Païta 1990:25–41, 44) and of Cèmuhî (Rivierre 1980:97–118) will provide the beginnings of an explanation for the existence in Tayo of some distinctions (Completive, Future, Continuative) and the absence of others which occur in other creole languages (Anterior, Irrealis,...).

While much descriptive and comparative work remains to be done, we have now a fairly clear picture, variation included, of Tayo some 130 years after the permanent settlement at St Louis of the Marists and their Cèmuhî-speaking associates. The grammatical areas examined

10 Nor from Réunionnais (Chaudenson 1994:135); although there are certainly some parallels in the use of the impersonal *na* (and associated forms) in Tayo and in Réunionnais, as in (39) for example, the semantics of (37)–(38) are no more Réunionnais than they are French.

here encapsulate the development of Tayo from its beginnings to the present: traces of its early stages are still visible; its emerged, conventionalised (or vernacularised), and essentially Melanesian form is well attested; and we can see the effects of recent French influences on this new language.

The relatively detailed study of the pronominal system and of the relativisation and thematisation strategies have shown that Tayo is in no way a modification of French, nor a relexification of a Melanesian language, but that it is a new creation which is typologically Melanesian. And yet, as soon as certain variants are considered, it is far from clear just what the syntactic rules actually are. For example, *le sek latam* 'the table is dry' is instantly recognisable as a Stative Predicate + Subject Melanesian pattern, but what about the variant form *se sek latam* 'it is dry, the table'?[11] It seems that replacing *le* by *se* causes a profound change in the grammar of the sentence, a change which is in the direction of French. Most variations of this kind noted here and especially by Ehrhart (1993), are part and parcel of the normal usage of those members of the tribe who have a certain level of knowledge of French. We are left then with a grammar of a Melanesian type onto which are grafted structures belonging to another grammar, that of French. This Melanesian-type grammar is, however, less complex than that of any given ancestral language. If sometimes it includes what is common to Cèmuhî and Drubéa,[12] by the regularisation or elimination of specific features, it is still more than a simplified calque, as is shown by such innovations as *le*.

In the light of these remarks, let us now consider an element of the St Louis tribal oral tradition (Ehrhart 1993:27):

> Before, all we did was express the Melanesian language in French because there were people from all over. It was a translation into French of the (Melanesian) language. We translate(d) the language into French. [My translation.]

This view captures a part of the truth. It claims that Tayo is merely relexified Melanesian. It is a correct view in so far as all the Melanesian languages involved in the formation of Tayo agree on how any self-respecting grammar should organise things, and it is this convergence which explains the Melanesian 'feel' of Tayo. But it is only a partial view, since no account is taken either of innovations or of the French input in the formation of the new language. For example, the adjectives which are preposed to the noun include: *gra* 'big, long, tall', *gro* 'big, fat', *peti* 'little, young', *vye* 'old', *move* 'bad'. Their preposing in Tayo can only have come from French, since neither Cèmuhî nor Drubéa allow this order for these items. The word order of Tayo must not be overlooked in this context, either: if the basic word order puts a lexical subject after the verb, Tayo also allows (beyond thematisation) an initial subject which reflects French grammar (as do present-day versions of some Melanesian languages of southern New Caledonia). If one were to take literally the view expressed by the oral tradition of the St Louis villagers, Tayo would indeed be "une langue mélanésienne à vocabulaire français", to rephrase Sylvain's famous formula. But, as has been seen, this is a partial view, since French too provided some input to the formation of Tayo.

Tayo, a language formed and relatively stabilised towards the beginning of the century, illustrates aptly the various procedures implicated in the formation of a new language. It may be noted in passing that the settlement and sociodemographic history of St Louis and the

---

11 cf. French *la table est sèche* and *c'est sec, la table*, the latter being an unremarkable case of right-dislocation.

12 And probably Xârâcùù as well; this language is not discussed here for practical reasons.

linguistic data advanced here provide mutual support: the social history prefigures the results of the linguistic analysis, just as these last reflect that history.

REFERENCES

Chaudenson, Robert, 1994, A propos de Sabine Ehrhart, Le créole français de Saint-Louis (le tayo) en Nouvelle-Calédonie. *Etudes créoles* 17/1:128–142.

Corne, Chris, 1989, Un créole à base lexicale française en Nouvelle-Calédonie: le tayo ou le patois de Saint-Louis. *Etudes créoles* 12/2:29–42.

1990a, L'agencement temporel des événements démographiques dans la création d'une langue créole et le tayo de Saint-Louis en Nouvelle-Calédonie. In Francine Tolron, ed. *L'Homme et le temps*, 11–27. Nouméa: C.O.R.A.I.L.

1990b, Tayo pronouns: a sketch of the pronominal system of a French-lexicon Creole language of the South Pacific. *Te Reo* 33:3–24.

1993, Creole French: of continuity, change, and creation. *Prudentia* 25/2:47–71.

1994, Relativization and thematization in Tayo and the implications for Creole genesis. *Journal of Pidgin and Creole Languages* 9/2:283–304.

1995, A contact-induced and vernacularised language: how Melanesian is Tayo? In Philip Baker, ed. *From contact to creole and beyond*, 121–148. London: University of Westminster Press.

Ehrhart, Sabine, 1992a, The development of the preverbal markers in St-Louis Creole: the formation of a TMA system? *Language Sciences* 14/3:233–247.

1992b, La fête dans les traditions de la tribu de Saint-Louis, Nouvelle-Calédonie: brève étude linguistique d'un récit en créole français (tayo). In B. Gasser, ed. *La fête*, 147–164. Nouméa: C.O.R.A.I.L.

1993, *Le créole français de St-Louis (le tayo) en Nouvelle-Calédonie.* (Kreolische Bibliothek, 10) Hamburg: Buske.

1994, Der Einfluß des Réunionnais auf das Tayo in Neukaledonien—Mythos oder Realität? Paper presented to the 16th Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft, Münster.

n.d., Quelques réflexions concernant la genèse du tayo en Nouvelle-Calédonie et discussion des éventuels apports venant de l'extérieur, notamment du créole réunionnais. MS.

Ehrhart-Kneher, Sabine and Chris Corne, 1996, The Creole language Tayo and language contact in the 'Far South' region of New Caledonia. In Peter Mühlhäusler, Darrell Tryon and Stephen Wurm, eds *Atlas of languages of intercultural communication in the Pacific, Asia, and the Americas*, 265–270. Berlin: Mouton de Gruyter.

Frei, Henri, 1929, *La grammaire des fautes*. Paris: Geuthner, Leipzig: Harrassowitz and Geneva: Kundig.

Guiraud, P., 1966, Le système du relatif en français populaire. *Langages* 3:40–48.

Rivierre, Jean-Claude, 1973, *Phonologie comparée des dialectes de l'extrême-sud de la Nouvelle-Calédonie*. Paris: Société d'Etudes Linguistiques et Anthropologiques de France.

1980, *La langue de Touho: phonologie et grammaire du cèmuhî (Nouvelle-Calédonie).* (Langues et civilisations à tradition orale, 38.) Paris: Société d'Etudes Linguistiques et Anthropologiques de France.

Shintani, T.L.A. and Yvonne Païta, 1990, *Grammaire de la langue de Païta*. Nouméa: Sociéte d'Etudes Historiques de Nouvelle-Calédonie.

Corne, C. "The typology of the Tayo language of St Louis, New Caledonia". In Tent, J. and Mugler, F. editors, *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.*
C-141:11-26. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.11*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.

GENDER AND TENSE/ASPECT IN BELIZEAN CREOLE

GENEVIEVE ESCURE

1. INTRODUCTION[1]

Claims relating to the role of gender in language change have been varied and often contradictory. For example, it has been argued that women favour prestige varieties as a way to compensate for their lack of power, whereas men prefer the feeling of camaraderie which results from the use of vernacular variants (Trudgill 1978). A related claim is that women initiate change, but only in the direction of the standard. On the other hand, linguistic change is assumed to go in the direction of vernacular forms (Labov 1972), which would imply that indeed men lead in the development of linguistic change. Does this also mean that men and women increasingly diverge in their language patterns? The claim has in fact been made that men and women use distinct communicative mechanisms, or even speak different 'languages'. The U.S. mass media regularly report simplified abstracts of (largely unpublished) neurological research suggesting that men and women use different parts of the brain. For example, it is stated that "male brains are more asymmetrical than female brains", and that "men use [when reading] a minute area in the left side of the brain while women use areas in both sides of the brain" (*New York Times*, 16 February 1995); that "women are better at verbal memory...while men consistently do better at spatial reasoning" (*Minneapolis Star-Tribune*, 11 February 1996); that "men as a group excel at tasks that involve orienting objects in space;...women, on the other hand, seem to be more adept at communication, both verbal and nonverbal" (*Time*, 17 July 1995).

One problematic aspect of most gender studies is their limitation to Western societies, and primarily the middle classes, hardly a valid platform leading to gender generalisations. The few studies investigating more ethnically and socioeconomically diverse groups in Western-type societies or in developing countries have found that sex-based differentiation is much more complex than the brief description above. The putative effect of gender on linguistic change is particularly relevant in the context of a colonial past, which is likely to lead to a complex multidimensional space (Le Page & Tabouret-Keller 1985). In a context forcing men and women into equally subjugated positions, both sexes learn similar survival mechanisms, which no doubt entail extensive lectal shifting for necessary social adjustments. It may thus

1 My sincere thanks to Chris Corne and France Mugler for their helpful comments.

Jan Tent and France Mugler, eds *SICOL: proceedings of the Second International Conference on Oceanic Linguistics:* vol.1, *Language Contact*, 27–41.
Pacific Linguistics, C–141, 1998.
© Genevieve Escure

Escure, G. "Gender and tense/aspect in Belizean creole". In Tent, J. and Mugler, F. editors, *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.*
C-141:27-41. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.27*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.

be hypothesised that postcolonial situations, and others, in which men and women share powerlessness, may involve fewer instances of sex-based language differentiation.[2]

The problem in many gender studies is one of methodological scope: explanations are directly dependent on the kind and range of speech data collected, and analyses based on socially restricted cross-sections of men and women can hardly provide universal explanations of gender behaviour. This study is an investigation of the putative role of gender as a social variable in the diffusion of linguistic change in a creole continuum—specifically, the English-based continuum in Belize, Central America. A previous study of copular variation in a Belizean rural community concludes that women extensively use the vernacular (creole basilects) in community activities and thus cooperate with men to preserve local identity and the traditional values rooted in the creole vernacular (Escure 1991). In addition, age appears to be a relevant factor which intersects with gender: middle-aged women exhibit especially extensive style-shifting, commonly using acrolects for certain official functions. Some age groups evidence no sex-based language differentiation, whereas others do, and differences appear to be related to local social status more than to sex, or gender. In fact, all Belizeans, men as well as women, evidence to a certain extent their ability to style-shift in various contexts, yet women exhibit a greater overall ability or willingness to shift across lects. I call this cross-lectal shifting tendency 'linguistic bipolarity'. This finding is based solely on the analysis of copular variants and must of course be further tested in the context of other linguistic features.

With this goal in mind, I investigated the functioning of certain tense/aspect categories (especially as they involve past tense reference) in the community of Placencia. If the distribution of tense-marking variants conforms to that of copular forms, then the hypothesis of women's linguistic bipolarity will be strengthened. Generally, the broad range of linguistic choices available to the members of a Creole community aptly reflects the conflicting identities common to postcolonial societies. Extensive lectal shifts may be part of the decreolisation process which is often assumed to be happening in the West Indies, as well as in other developing or postcolonial societies. If decreolisation is viewed as internal change away from an earlier grammar, systematic formal differences observed between basilects and mesolects could be viewed as indicators of ongoing historical change, and it would then be particularly interesting to observe whether one sex or the other is more actively involved in ongoing linguistic change.

Lectal shifting can be viewed as an extension of individual repertoires, that is, the addition of acquired second dialects to native vernacular basilects, rather than the substitution of more standardised varieties for non-standard lects. This interpretation is supported by the fact that most Creole speakers control a broad repertoire which keeps expanding during their lifetime.

2. LECT-SHIFTING AND METHODOLOGY

A prerequisite in the investigation of the role of gender in linguistic change is the accurate identification of the features which characterise basilects and mesolects in a typically fluctuating continuum of overlapping varieties. In addition, it is essential to ascertain the

---

2 See Escure (1991) for a brief overview of gender and language research. A typical problem in gender research has been the confusion of such notions as 'sex' (the biological differences between men and women), and 'gender' (the construction or interpretation of the social roles assigned to men and women).

extent of individual repertoires for both sexes, since a single speech sample for each individual may lead to the fallacy that this individual always performs at a single stylistic level. Thus a variety of natural contexts should preferably be included in the survey of speech varieties.

The speech data that make up this corpus result from years of participant-observation in the Belizean village of Placencia, located on an isolated peninsula of the Caribbean coast of Central America (Escure 1983b).[3] Spontaneous conversations were recorded either by a local fieldworker or by myself, and were conducted in natural settings, such as private homes, the beach, or local shops, systematically avoiding formal interviews. This data base is derived from several interactions with five men and six women, who range in age between 30 and 65.[4] Selection of those samples was based on three criteria: (a) spontaneity; (b) reference to past events; and (c) representation of distinct styles, in particular basilects and mesolects, since acrolects are relatively rare in a natural local context. Among the samples included in this analysis, the closest to an acrolect was produced by the village Anglican priest (Flo) during a church service, and indeed church may be the only truly appropriate context for acrolectal production in Placencia, and probably the highest variety serving as a counterpart to Standard Caribbean English. Speakers in formal contexts are much more likely to produce meso-acrolects, that is, highly volatile varieties sharing various lectal features.

There is no absolute definition of lectal boundaries. Given that all speakers are able to shift at least between vernacular, traditional styles (basilects), and less casual varieties approaching Standard English in variable ways (mesolects or meso-acrolects), speakers' choices are determined by the context as well as by their own interpretation of stylistic appropriateness, and this may vary from one individual to another. I consider that a reliable measure of lectal level may be derived from comparing competing variants for the same feature: for example, in Belize, copular variants include *de* (clearly basilectal); *zero-morpheme* (which occurs at all levels depending on following grammatical environments); and inflected forms of English *be* (mesolectal and acrolectal). Relative frequencies of those variants permit an assessment of lectal level, which can independently be confirmed by other linguistic or extra linguistic features, such as phonological variation and contextual knowledge.

In the case of *past* tense/aspect marking, three variants can also help assess lectal level: *me* (exclusively basilectal); *zero-marking* (distributed across basilects and mesolects); and English standard-like *preterite* forms, which occur with strong verbs, and mostly in mesolects and acrolects. An interesting change is noted in the Placencia community, namely the apparent loss of constraints on differential markings for the anterior aspect and the stative semantic class of verbs.[5]

3. IS THERE PAST REFERENCE IN BELIZE?

Bickerton (1975) has claimed that creole verbal systems, contrary to their Indo-European superstrates, do not have tense, but rely instead on aspectual and semantic distinctions. In the case of the reflexes of Standard English past tense marking in English-based Caribbean

---

3 In the last few years, a road and an airstrip have been constructed, opening up the area to a fledgling industry. Escure (forthcoming) discusses the impact of this change on the region and on the language.

4 The women are overall older than the men, which can admittedly constitute a potentially significant difference, but the age variable will not be included in this study.

5 The phonological constraints on unmarked verbal past forms are not investigated here.

creoles, relevant distinctions would include [punctual versus non-punctual] and [anterior versus non-anterior] aspects; and [stative versus non-stative] as semantic verbal subcategorisation. According to Bickerton, Guyanese Creole non-stative propositions use the bare verb stem form for 'unmarked simple past', whereas stative verbs use a preverbal morpheme to refer to simple past. Non-stative verbs are marked by preverbal morphemes for 'nonpast', that is, either for 'anterior aspect', which is often referred to as 'past before past', or 'continuative or iterative' aspect. Adjectives (functioning as stative verbs) are said to take a combination of markers when a process rather than a state is indicated (Bickerton 1975:30). The system suggested by Bickerton, and endorsed by many other creolists is summarised in Table 1.

TABLE 1: CREOLE VERBAL SYSTEMS (BICKERTON 1975)

1.	0	[+stative] = nonpast
2.	0	[-stative] = simple past
3.	Anterior	[+stative] = simple past
4.	Anterior	[-stative] = past before past (said to be rare)

However, this analysis has been variously challenged for Guyanese Creole (Gibson 1986), for Jamaican Creole (Pollard 1989), and generally for Caribbean English creoles (Winford 1993:66). It is demonstrated below that the stative/non-stative distinction for anterior/non-anterior marking is not, or is perhaps no longer, represented in Belizean. Thus, the claim that English, which has marked past tense, and unmarked nonpast (in fact habitual or iterative aspect), is entirely different from its related creole, is open to question. It is not obvious that creoles are indeed so different from their lexical base or any other language in terms of their tense/aspect system, as also stated in Corne (1995), Escure (1993a), Tagliamonte and Poplack (1993), not to mention possible counterexamples in non-standard forms and casual styles of standard varieties.

There are clearly similarities between English and English-based creoles: English, like creole, marks progressive aspect, which is nonpast (*I'm singing; I de sing*); some English statives, like their creole counterparts, cannot be used in the progressive aspect, as in **I'm knowing*; **I'm believing*.[6] Most importantly, Bickerton's claim that the use of the 'anterior' aspect marker (*me* in Belize) is rare before non-stative verbs, and if it occurs means 'past-before-past', or at least is 'related to a prior state', is not consistently supported by my Belizean data, as shown below in (2), (3) and (4). Bickerton's claim that the copula/zero copula does not co-occur with the anterior marker unless it refers to a process rather than a state is contradicted as well by sentences such as (4). In conclusion, it is not obvious that *me* is semantically differentiated from the past reference involved in the unmarked verb stem. What is clear, however, is that there *is* a rich system for past reference in creoles, often independent from aspect, which is represented throughout the varieties of the Belizean continuum. Whether or not the coexisting morphemes are related to significant semantic distinctions, they no doubt have strong lectal or stylistic associations, and provide clues to the linguistic choices effected by the speaker. It is from this perspective that I will evaluate

6 *I'm believing* is sometimes acceptable in certain English varieties (e.g. New Zealand English, according to Chris Corne (pers. comm.)). It is possible in American English as well, but in contexts usually marked for emphasis. These options may indicate an ongoing shift in the use of the progressive aspect with statives.

the function of gender in ongoing changes affecting past tense referential mechanisms. Each of the three lectal variants is discussed below.

3.1 LECTAL VARIANT 1

Basilectal *me* ('bin/ben' in other West-Indian creoles) is a preverbal morpheme, with a negative counterpart *neva* (= *me* + NEG). It is glossed as [ANT] for [anterior] in the following examples, in spite of the fact that the semantic value of *me* seems to be evolving, and that the putative 'past-before-past' or 'prior' semantic connotation does not always apply, as illustrated below. The +/- [ANT] value indicates whether prior reference applies to the verbal unit in the context of a given interaction. The semantic interpretation is determined by pragmatic considerations, and in particular knowledge of the discourse context, and background cultural assumptions requiring of course intimate awareness of the situation.

The following sentences illustrate the changing function of the morpheme *me*. They are all derived from the narrative of a single event: the shipwreck of a Belizean oil tanker off the coast of Guatemala, told by one of the crewmen (Bli) to his friend (Rol), both Placencia natives:

(1) *Who unu* **me** *go da trip fa?*[7]
who you.PL [+ANT] go that trip for
For whom had you been working (on that trip)? (Rol 20:15)

Me before the punctual (non-stative) verb 'go' refers to an event **prior** to the 'shipwreck' event, a combination which Bickerton claims is rare (combination 4 in Table 1).

(2) *A tink da* **me** *only six tousand galan* **me** *come out.*
I think TOP [-ANT] only six thousand gallons [-ANT] come out
I think that only 6,000 gallons leaked out. (Bli 20:17)[8]

This sentence refers to the oil slick following the shipwreck, and begins with a presentative structure which highlights a new topic, namely the amount of oil having escaped from the tanker. In such cases, the topic particle *da* (labelled TOP in the interlinear gloss) frequently occurs, as is evidenced in (3), (4) and (8). The first *me* which occurs between the topic particle *da* and the topic *six tousand galan* is obviously not related to any anterior value, nor is the second *me* preceding the punctual 'come out' (thus violating combination 4 in Table 1). In this sentence, both instances of the putative anterior morpheme now appear to have a topic-marking value, indicating a possible grammaticalistion of *me*, to be discussed in §6.

The uncertain status of *me* is shown when comparing its use as anterior in (1), versus its non-anterior status in (2), although this morpheme co-occurs in both cases with a non-stative verb. Sentence (3) is a paraphrase of (2), which it immediately follows in the conversation: *neva*, the negative form of *me*, occurs in the same context as above, after a topic marker, and indicating also a highlighting function; *me* does not occur before the punctual verb 'come',

---

7 The transliteration representing creole data adopts standard English orthography as much as possible, with a broad phonemic representation in the case of creole (non-English) morphemes such as *me* (past); *de* (progressive); or *unu* (you-plural). Regular phonological differences are marked (the absence of interdentals is represented by an orthographic [d] or [t]).

8 [TOP] in the literal translations refers to topic markers or highlighters, especially *da*—probably derived from the English deictic 'that', as argued in Escure (1983a).

contrary to (2), yet no semantic change is involved. Thus, in this case, combination 2 is upheld:

(3) *Da* **neva** *much oil* **come** *outa di barge.*
TOP not[-ANT] much oil come out of the barge
Not much oil leaked out of the tanker. (Bli 20:17)

(4) *Da* **me** *wan propaganda ting; an i* **come** *wan*
TOP [-ANT] a propaganda thing and it come a

time wen dis Guatemala question **me** *kinda hot.*
time when this Guatemalan question [+ ANT] kind.of hot
That was pure propaganda, and it came at the height of the problems with Guatemala. (Guatemala threatened to invade and annex Belize.) (Rol 20:17)

In (4), the speaker is saying that the crew of the shipwrecked tanker was kept hostage in a Guatemalan prison because of prior political problems with Belize (Guatemala had threatened to annex Belize). Thus the first *me* has non-anterior reference (a simple past reference before [stative] zero-copula, agreeing with combination 3 in Table 1), and this simple past value is also represented in the non-marking of the verb *come*, but it is also part of a presentative phrase and is, as above, likely to carry a focusing function. As to the second *me*, it has possible anterior reference (also before a [stative] zero-copula) 'kinda hot', now violating the same combination 3.

(5) *A* **me** *tink you me know de guy.*
I [-ANT] think you [-ANT] know the guy
I thought you knew the guy. (Co 21:5)

In this sentence, there are two stative verbs, with simple past reference, both agreeing with combination 3, but there is an additional emphatic or contrastive value evidenced in the context, as the sentence is uttered in response to the statement that 'the guy' was not known to the speaker's interlocutor.

In the above examples, only two out of the eight cases of the so-called anterior aspect morpheme *me/neva*can be associated with a truly prior event in (1) and (4), as indicated in the glosses. But most instances of *me* (7 out of 8), regardless of the stative/non-stative status of the following verb, consistently illustrate non-anterior topic-marking functions.

3.2 LECTAL VARIANT 2

The 'bare stem form' has typically the value of simple past, and may appear in basilects and mesolects, as well as in acrolects, as in (6). As can be seen in (3), (4) and (8), this stem freely combines with *me/neva*, and, as indicated above, there is no evidence that the marking/zero marking distinction is related to the stative or non-stative semantic verbal categorisation. The bare stem occurs as well with stative verbs referring to continuing events such as *live*, or existential events such as *baan* 'be born' and *grow* 'grow up' in (7). This is in agreement with combination 2:

(6) *One* **come** *da Wednesday.*
One (a baby's birth) came on Wednesday. (Til 12:10)

(7) *All my life I* ***live*** *here, I* ***baan*** *and* ***grow*** *here.*
All my life I have.lived here, I was.born and grew.up here.
(Dor 3:3)

(8) *When I da* ***me*** *wan group leader de, I* ***know*** *dem*
when I TOP [+ANT] a group leader there I knew those

girl, dey ***like****, well, steady* ***go run go tell, run go tell*** *pan*
girls they liked well always go run go tell run go tell on

dis girl; I ***try*** *cover up fu dem all de time.*
this girl I tried cover up for them all the time
When I was group leader, I knew that those girls (office workers) always liked to tell on other girls. I tried to cover for them all the time. (Rol 21:3)

In this sentence, unmarked verbs include statives ('know, like') as well as non-statives (the serial strings of punctual verbs ('go, run, tell'), but they refer to a past situation which implies an iterative context. This habitual reference is captured by the adverbial 'steady' which is commonly grammaticalised into an aspect morpheme.

3.3 LECTAL VARIANT 3

The marking of past tense (preterite) as in Standard English occurs commonly in acrolects, as in (9):

(9) *Dis* ***was*** *about two years ago when I first* ***met*** *him.*
I met him about two years ago. (Co 21:4)

What is striking in the case of marked preterites is that there is a high frequency at **all** lectal levels of the two verbal auxiliaries *had* and *was*. A close investigation of verbs marked for past tense in this sample, shows that most of the standard preterite forms produced consist of the two verbs *had* and *was*: they represent overall 42.7% of all standard past forms found among Placencia speakers, as represented in Table 2. Most specifically, men use *had/was* 41.8% of the time they produce English-type preterites, and women 43% of the time, that is, an insignificant difference.

TABLE 2: DISTRIBUTION OF *had* AND *was*

Gender	N	n	*had* (%)	*was* (%)	Other (%)
Men	536	194	32 (16.5%)	49 (25.3%)	113 (58.2%)
			both 41.8%		
Women	1,067	581	80 (13.8%)	170 (29.3%)	331 (57%)
			both 43%		
All	**1,603**	**775**	**112 (14.5%)**	**219 (28.3%)**	**444 (57.3%)**
			both 42.7%		

(N = Total number of verbs; n = Total number of standard type preterites.)

It is not unusual to find sentences in which a past-marked verb co-occurs with an unmarked verb as in (10). Furthermore, *had* and *was* frequently occur in a non-acrolectal

context, as in (11), which is part of a narrative produced in a mesolectal style by a 65-year old woman:

(10) *I* ***was*** *surpris' when dey* ***tell*** *me de price of cemen'.*
I was surprised when they told me the price of cement.

(11) *We* ***had*** *bad weda, Placencia* ***was*** *under water.*
We had bad weather, and Placencia was under the water.

It is also very clear that *had* and *was* do not always function as past tense markers: in (12), *had* co-occurs with *me*, with no anterior reference, thus apparently creating a redundant past context. In fact the possibility of psychiatric treatment would have to be **posterior** to the shipwreck and following imprisonment of the crew of the shipwrecked tanker, because of the stress they suffered in the Guatemalan prison (they were held as pawns in exchange for financial compensation by the U.S. company which owned the tanker). Here again *me* appears to be restructured as an emphatic marker (see §6):

(12) *I hear i* ***me*** *nearly* ***had*** *to get apsychiatric treatment*
I hear he [-ANT] nearly had to get psychiatric treatment
a de rass.
(expletive)
I heard that he was almost forced to undergo some damned psychiatric treatment.

This emphatic function evidenced by *me..had* in (12) is also represented in the context of (13): *'we* ***had*** *to catch de plane'* is clearly not a modal (since the midwife and mother should already have been on the plane, if the weather had allowed), but may be more appropriately translated as 'We finally **did** get on the plane'.

4. MIXED MARKING STRATEGIES

It is interesting to note that the three past variants of verbs are commonly intermixed in the same text, as shown in the following conversation produced by Til, Placencia's midwife and *obeah* woman:

(13) *I* ***had*** *one case, it* ***was*** *twins. Well, one* ***come*** *da Wednesday, and we* ***had*** *bad weda, Placencia* ***was*** *under wata, no plane* ***can*** *land at di airstrip, no boat* ***come,*** *the road* ***block*** *off, the bridge dey* ***flood.*** *Well data baby* ***neva*** *come. Well, di wata* ***broke*** *Friday morning, and we* ***had*** *to catch di plane and* ***gone*** *up. Dat baby* ***born*** *on di plane just when di plane* ***pitch*** *into Belize. But di ambulance* ***come*** *wid, it* ***had*** *oxygen, you know,* ***come*** *a help di baby. And di baby* ***died*** *on de way to hospital.*
I had a case of twins. One of them was born on a Wednesday. There was a storm that day and Placencia was flooded. Planes couldn't land at the airstrip, boats couldn't dock, the road was blocked, and the bridges flooded. Well, the other baby wouldn't come. When the water broke Friday morning, we finally got on the plane. The baby was born on the plane, just as we descended toward Belize-City. The ambulance met us with oxygen to help the baby. The baby died on the way to the hospital. (Til 12:10)

Out of 19 verbs, nine are marked for past, and among them, four are *had*, and two are *was* (amounting together to 66.6% of all preterite-marked verbs). Of the three remaining

verbs, which are apparently marked, three are in fact nonpast, that is, not truly marked for past, although the context **is** past: *broke*, *gone* and *born* are used in English-based creoles as the base verb forms (they are thus equivalent to 'break', 'go' and 'be born'). In other words, the English preterite has been lexicalised in creoles as the present form. There are other verbs which follow the same pattern (e.g. *lef* 'to leave', *got* 'to get', *ful* 'to fill', *las* 'to lose'). The following examples collected in spontaneous conversations highlight the nonpast value of these relexified forms:

(14) *Bra Anansi* ***got*** *up pan di tree and he* ***see*** *Bra Tiger.*

(15) *I* ***had*** *to* ***lef*** *dat one.*

(16) *De dance* ***had*** *to* ***brok*** *up.*

(17) *Dat must* ***got*** *cancer.*

(18) *Everybody use to* ***got*** *baby easy.*

Those false preterites are more likely to function in the system as unmarked variants, and it would follow from this observation that past marking is even less common in the samples investigated than is represented in the tables. Since the relative incidence of those four or five verbs is somewhat limited, their inclusion as 'marked preterites', although inaccurate, should not change notably the overall significance of the results.

5. DISTRIBUTION OF PAST VARIANTS IN BELIZE AND GENDER

When comparing men's and women's use of the three past variants the difference in their linguistic behaviour appears to confirm claims that men use more vernacular forms than women, at least assuming that the most basilectal forms represent the vernacular. As shown in Table 3, men appear to be overall more partial to the basilectal variant *me/neva*, which occurs over 30% of the time, almost on a par with the other two variants: *unmarked verb forms* (33%), and *marked verbs* according to the standard manner (36%). On the other hand, women display an almost complementary use of *zero* and standard *past* marking (*zero* = 41.9%; and *past* = 54.5%), whereas *me/neva* occurs less than 4% of the time. But when looking at individual behaviour, another interesting pattern emerges: the six women studied display relatively homogeneous usage of variants (Table 4), whereas men contrast much more sharply in their relative use of past variants (Table 5).

TABLE 3: WOMEN'S AND MEN'S USE OF PAST VARIANTS

Gender	N	*me/neva* n (%)	Zero n (%)	Preterites n (%)
Men	536	165 (30.8%)	177 (33%)	194 (36.2%)
Women	1,067	39 (3.7%)	447 (41.9%)	581 (54.5%)
All	**1,603**	**204 (12.7%)**	**624 (38.9%)**	**775 (48.3%)**

TABLE 4: WOMEN'S USAGE OF THE PAST VARIABLE

Name, age	N	*me/neva* % (n)	Zero % (n)	Preterites % (n)	Lect
Dor, 65	147	4.8% (5/1)	71.4% (105)	24.4% (36)	basilect
Til, 66	229	2.1% (2/3)	55.8% (128)	41.9% (96)	basi/mesolect
Ten, 37	332	4.8% (2/14)	28.6% (95)	66.5% (221)	mesolect
Fran, 68	81	2.4% (0/2)	37% (30)	60.4% (49)	mesolect
Ag, 55	209	4.7% (2/8)	39.7% (83)	55.5% (116)	mesolect
Son, 38	24	0	25% (6)	75% (63)	meso/acrolect
All	**1,603**	**3.6%**	**41.9%**	**54.5%**	

TABLE 5: MEN'S USAGE OF THE PAST VARIABLE

Name, age	N	*me/neva* % (n)	Zero % (n)	Preterites % (n)	Lect
Rol, 28	114	66.6% (70/6)	26.3% (30)	7% (8)	basilect
Bli, 27	241	36.9% (65/24)	41.9% (101)	21.1% (51)	basilect
Hab/R, 51	35	0	40% (14)	60% (21)	mesolect
Hab/G, 51	86	0	26.4% (23)	72.4% (63)	mesolect
Flo, 36	60	0	15% (9)	85% (51)	acrolect
All	**536**	**30.8%**	**33%**	**36.2%**	

(Speaker Hab is represented twice in two different contexts)

The limited overall use of *me/neva* by women is consistent across all individuals. Their lectal differences are marked through the differential usage of unmarked and preterite-marked variants. In contrast, men differ sharply in their use of *me/neva*, which is almost exclusively used by the two young men recounting the Guatemalan shipwreck. In fact the 30.8% *me* men's average depends solely on Rol and Bli's contributions, which also constitute the longest sample in the corpus. This is not tantamount to saying that Rol and Bli, or any other speaker, always perform at the same lectal level. The range displayed in the corpus represents actual possibilities for each individual, but naturally not the complete stylistic repertoire available to each. However, the data presented in Table 5 are considered to be representative of men's linguistic behaviour in the community. On the basis of this limited sample, and a restricted linguistic feature, it would appear that men and women are equally capable of shifting between lects. The linguistic bipolarity observed earlier in women (for copular variants) is not particularly evidenced in the case of the past variable: in this case, men show even more internal variability within their gender group. They switch between lects which show almost complementary percentages of the three past variants *me/zero/preterite* (respectively, 66.6% / 26.3% / 7% versus 0 / 15% / 85%). Women mostly vary in their use of *zero* and *preterite*, ranging for those two variants between a basilectal 71.4% / 24.4% combination, and an acrolectal 25% / 75% combination. The women's

systematically restricted usage of the basilectal morpheme (only 3.6%, in Table 4) indicates even in the most vernacular variants an overall loss of the traditional creole morpheme *me*, which coincides with the apparent loss of anterior versus non-anterior aspect distinctions.

However, this development does not entail a clear movement towards the standard since women use more unmarking (*zero past*) than men, namely 41.9% for women versus only 33% for men. Furthermore, marking in the standard fashion is prominently represented in two special auxiliaries *had* and *was*—a huge 42.7% of all marked past forms—which have taken new nonpast functions. Therefore, women replace a non-standard feature by another one, a feature, however, which is less marked, less stigmatised than the creole morpheme *me/neva.*

5.1 EFFECT OF SEX AND GENDER ON LINGUISTIC CHANGE

It is possible to interpret the facts described above as an indication that women are more sensitive than men to the progression of linguistic change in the direction of the standard, and are more active in implementing it, but only if the loss of creole features is viewed as a move towards the standard or acrolectal level. According to this interpretation, women would spearhead the disappearance of *me/neva* morphemes, while losing the distinctive anteriority function, but would simultaneously assign new functions to incoming standard forms. This conclusion fits with my previous analysis of copular variation (Escure 1993b), in which women were found to be more selectively active in removing the basilectal morpheme *de*, but only in the locative context, not in the progressive context, which still seems to be well established in the basilect.

However, the argument for gender- or sex-based linguistic differentiation is weak, at least in the context of the Belizean community's mechanisms for marking past. The findings presented above are too limited to validate the claim that there is sex-based differentiation in the village of Placencia. And the significance of gender as a variable representing the social construction of sex roles is even more difficult to assess. Although men and women have apparently distinct roles in a traditional sense, Placencia is not a community where there is a power imbalance (and women typically constitute the powerless sex). At the time of collection of the speech data used here (1982–87), men were fully involved in fishing activities that would take them away from home for two to four weeks at a time, whereas women were fully in charge of household chores and village administration, including important decisions regarding local economic and social issues (see Escure (1991) for a fuller discussion). The older women included in the tables (Dor, Til, Son) have particularly high social status, yet they produce distinct lects. Ten, who does not actively participate in village decisions (she has 14 children to raise, not to mention grandchildren) produces a variety similar to that of Son (who takes most decisions as the wife of a regularly absent village chairman—he is either working or 'relaxing'). On the other hand, men seem to be split according to age, although differentiation may be due to the context of the interactions. Thus, no final conclusion can be drawn here as to the effect of either age or gender as social variables, until further data are adduced.

In fact, the present data confirm my earlier hypothesis that in a creole (postcolonial) community with earlier social levelling due to the historical subjugation of both sexes, gender will not easily surface as a factor determining linguistic choices in the context of the extensive range of linguistic variability available to most creole speakers.

6. TENSE/ASPECT CHANGE THROUGH GRAMMATICALISTION

The phenomenon observed in Placencia provides an insight into the process of linguistic change in the creole continuum, which some would call decreolisation, or standardisation. The observation of past marking suggests that this change occurs in incremental steps, notably involving restructuring of lexical items, and reassignment of grammatical, semantic and pragmatic functions, a process commonly referred to as grammaticalistion.

There is some indication that both men and women implement changes in tense/aspect categories through a combination of grammaticalistion and extension. In the case of past marking, this dynamic process takes place in several quasi-synchronous or overlapping stages involving basilectal *me*, as well as the English items *had* and *was*, which show the following series of changes:

(a) the loss of the anterior/non-anterior distinction, and of its dependence on the stative/non-stative class of verbs (in (1) *me* has anterior value, whereas *me* in (2) has non-anterior value);

(b) the substitution of *had/was* for the anterior morpheme *me*, still preserving some anterior aspectual value, as in (12), (9), (11);

(c) the acquisition of new topic-marking pragmatic functions associated with *me/neva* (when it still occurs, as in (2), (3), (4), (5)), and with English auxiliaries, such as *had* and *was*.

6.1 GRAMMATICALISTION

The type of relexification known as 'grammaticalistion' has been defined as "the process whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions, and once grammaticalised, continue to develop new grammatical functions" (Hopper & Traugott 1993:xv). In other words, the lexicon of language A (here creole basilect) is replaced by that of language B (English), while the grammar remains A. Thus, *me* is first replaced by *was/were/had*, which may refer to an anterior aspect of a current event as in (19):

(19) *The cassaba* ***were*** *strain'; it comes into a powder.*
After the cassava **has been** strained, it becomes a powder.

It is also possible to interpret the use of *had/was*, as in (20), as referring to the setting of the story, here to events anterior to the death of the second twin (see context in (13) above):

(20) *We* ***had*** *bad weather; Placencia* ***was*** *under water.*
We had had bad weather, Placencia was under water. (= (11))

According to this scenario, there is no longer a clear anterior aspect represented by a simple morpheme, and the semantic stative/non-stative verbal classification is no longer relevant. As *had/was* replace *me*, they first assume the latter's anterior function, then replace it, also losing all past tense reference, at least in basilects and mesolects. Considering the bleaching of past marking, the distribution of standard past variants (as represented in Table 2) could be reinterpreted with instances of *had/was* counted as cases of non-marking. This reinterpretation involves adding the *had/was* cases (an overall 331 tokens, from Table 2) to the zero category in Table 3; and subtracting those same 331 tokens form the preterite

category in Table 3. The resulting analysis modifying Table 3 is shown as Table 6, indicating a predominance of non-marking (*me* =12.7%; *zero* = 59.6%; *preterite* = 27.7%).

TABLE 6: OVERALL USE OF PAST VARIANTS (WITH *had/was* AS UNMARKED)

Gender	N	*me/neva* n (%)	Zero n (%)	Preterites n (%)
All	**1,603**	**204 (12.7%)**	**957 (59.7%)**	**444 (27.7%)**

6.2 EXTENSION

The second process involves an innovative semantic extension. In basilects *me* can still be used, but with additional pragmatic functions, as a topic marker, or a general emphatic strategy. In such contexts, *me* no longer carries any essential tense/aspect function:

(21) *A tink da me only six tousand galan me come out.* (= (2); also (3), (4), (5) and (12))

This process appears to be widespread in current basilects and mesolects: there is an overwhelming presence of adverbials relexlified into topic highlighters, often concomitant with the complete loss of their initial semantic value. This is the case for the adverb *just* illustrated in (22). The speaker does not refer to a punctual event in the immediate past, but to a repeated event (the visit of a cruise ship) which no longer applies (the cruise ship ceased making a stop in the village):

(22) *We just* ***have*** *dat ship comin in here; dat* ***was*** *good because it only* ***spend*** *couple hours; de touris' it just* ***stay*** *couple hours.*
What happened, we had that ship visiting; it was good because it would only spend a couple hours here; the tourists just stayed a couple hours.[9]

The second *just* is more ambiguous, since it seems to combine emphasis and restrictive meaning. However, other contexts indicate that *only* (as well as other items, such as *still* in the following sentence) has taken a contrastive value not contained in the English gloss, as shown in (23) recently collected in Placencia. The same grammaticalisation appears to affect the adverbial *again* (24);

(23) *We* ***still*** *speak creole, a like my creole language* ***only*** *because G. say dat we take it and mess it up.*
We speak creole **regardless** (of the fact that tourism is developing), I like my creole language, **even though** G. says that we mess up English.

(24) *Dat not cool* ***again****, it's hot now, it isn't cool* ***again****, dat isn't cool* ***again****, it hot, it isn't nice* ***again****, don't do it, don't.*
That's no longer cool, don't do it. (woman admonishing a child not to drink what's left in a glass)

The above examples indicate that a complex process is taking place, simultaneously relexifying the creole anterior morpheme *me* with English past auxiliaries (*had/was*), and extending the use of *me* to pragmatic contexts by analogy with similar topic-marking functions assigned to a variety of preverbal adverbials.

9 This approximate translation attempts to capture the topicalisation (non-restrictive) of 'ship comin'.

7. CONCLUSION

This analysis of past reference in the Belizean Creole continuum indicates that there is clearly an ongoing restructuring which provides insights into the pattern of linguistic change in a creole continuum. The conservative *me* past variant is eliminated and replaced by *zero* (unmarked verb), or by an English preterite. This process is accompanied by complex adjustments, removing certain aspectual distinctions (anterior), semantic differentiation (stative), and adding pragmatic mechanisms: more specifically, the 'anterior' value of *me* is no longer clearly distinguished from simple past, assuming it ever was exclusively used with an anterior value; it is relexified through English auxiliaries, and becomes part of a general topic-marking process affecting preverbal adverbials. The English forms *had* and *was* spread to the creole end of the spectrum, and appear to hold an especially significant intermediary position in this overall grammaticalising of past.

As far as gender is concerned, women appear at first to be actively involved in linguistic diffusion in the sense that they accentuate innovations, especially with the intensive loss of the basilectal morpheme *me/neva*. Those innovations could be interpreted as a move towards more standardised speech, thus supporting the traditional claim that women favour prestige varieties. Yet, this loss is not necessarily indicative of standardisation, since it is concomitant with an increase of other non-standard (though perhaps less stigmatised) forms, such as the *zero-past* variant, and non-standard use of preterites.

There is another crucial factor in interpreting the Belizean results as indicators of the directionality of linguistic change: this involves defining what constitutes the prestige variety in Belize. In most creole (and other) societies, there is a conflict between official norms (usually assigned 'overt' prestige), and local norms (assigned 'covert' prestige). Although Standard English is the overt prestigious variety, as official and educational medium, it cannot be considered to be the only 'prestige' variety in Placencia: basilects and mesolects have strong values of identity, and are more commonly used than the external standard. In fact, vernacular values are not really covert, they are indeed overtly expressed and recognised, as the Placencia community is well aware of the diglossic value of language. Women may be more particularly conscious of the complementary functions of creole and the standard, and of the necessity of the function of local identity in maintaining community ties.[10]

Although there is overall little indication that sex, or gender, has a strong effect on linguistic choices in a postcolonial Caribbean/Central American society like Placencia, there is some support for the hypothesis that women initiate, or at least, are highly sensitive to linguistic change, as indicated in their restricted use of the basilectal past variant, relative to that of men. This is not tantamount to saying that women no longer control basilects, but it may mean that men and women somewhat differ in their definitions of basilects and mesolects.

Given the importance of topic marking and focus constructions in Belizean (and generally in all creoles), the innovative assignment of pragmatic functions to past morphemes appears to be a typical example of superstrate items moulded to fit concepts essential to the basilectal vernacular. This type of linguistic change may occur in various types of contact situations: for example, it is well illustrated in the development of Tayo, a 'vernacularised', relatively

10 Escure (1991) discusses the possible function of women as mediators in Placencia.

young creole spoken in New Caledonia, which "uses French lexical items to convey Melanesian conceptualisations" (Corne 1995:23).

REFERENCES

Bickerton, Derek, 1975, *Dynamics of a creole system*. Cambridge: Cambridge University Press.

Corne, Chris, 1995, A contact-induced and vernacularised language: how Melanesian is Tayo? In Philip Baker, ed. *From contact to creole and beyond*, 121–148. London: University of Westminster Press.

Escure, Geneviève, 1983a, The Belizean copula: a case of semantactic shift. In Laurence D. Carrington, ed. *Studies in Caribbean language*, 190–202. Trinidad: Society for Caribbean Linguistics.

1983b, Belizean Creole. In John Holm, ed. *Central American English*, 29–70. Amsterdam: Benjamins.

1991, Gender roles and linguistic variation in the Belize Creole community. In Jenny Cheshire, ed. *English around the world: sociolinguistic perspectives*, 595–608. Cambridge: Cambridge University Press.

1993a, Anaphoric reference in creoles and noncreoles. *Language Variation and Change* 5/3:233–250.

1993b, Gender and linguistic change in Belize. In Kira Hall et al., eds *Locating power*, vol.1, 118–131. Berkeley: University of California Press.

forthcoming, *Creole and dialect continua: standard acquisition processes in Belize and China (PRC)*. Amsterdam: Benjamins.

Gibson, Kean, 1986, The ordering of auxiliary notions in Guyanese Creole. *Language* 62: 571–586.

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

Labov, William,1972, *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Le Page, R.B. and Andrée Tabouret-Keller, 1985, *Acts of identity: creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.

Pollard, Velma, 1989, The particle *en* in Jamaican Creole: a discourse-related account. *English World-Wide* 10/1:55–68.

Tagliamonte, Sali and Shana Poplack, 1993, The zero-marked verb: testing the creole hypothesis. *Journal of Pidgin and Creole Languages* 8/2:171–206.

Trudgill, Peter, 1978, *Sociolinguistics patterns in British English*. London: Edward Arnold.

Winford, Don, 1993, *Predication in Caribbean English Creoles*. Amsterdam: Benjamins.

SINGAPORE COLLOQUIAL ENGLISH OR DEVIANT STANDARD ENGLISH?

ANTHEA FRASER GUPTA

1. INTRODUCTION

Like Fiji (Siegel 1987), Singapore offers a complex multilingual setting where contact varieties appear and disappear, but where the languages that have created them remain. The approaches made to Singapore English raise issues in two concepts that Le Page (1988:30) has said are fundamental in sociolinguistics:

(a) the definition of 'a language';

(b) the definition of 'native language' (/mother tongue).

I will not go into (a) at this point. As for (b), which in a way depends on it, I think that all normal humans have at least one native language and that many people have more than one, and that a native language is a language you learn in infancy, before learning any other. I have defended this view elsewhere (Gupta 1994:18). The concept 'native language' can surely not be defined (as it often is, even by linguists) on the basis of genetic ancestry or location of birth, and your personal first experiences of language learning are formative. Also, for contact varieties, the process of nativisation is rather crucial.

Although this paper is about Singapore English, much of what I have to say also applies to Malaysian English. Despite demographic and political differences in the role and distribution of English, there remain varieties of Malaysian English that are virtually identical to their parallels in Singapore (although there is greater variation in Malaysia than in tiny Singapore).

Singapore English is a range of Englishes. The contact variety which I call Singapore Colloquial English (SCE) is strikingly different from Standard English (StdE), especially in syntax. No English pidgin has ever been historically attested for Singapore or Malaysia, and present-day Singapore English has no relationship whatever to Chinese Pidgin English or Melanesian Pidgin English. There was a Malay-based pidgin (Bazaar Malay), which is still used a little today, and a related creolised Malay used by a Chinese community (Baba Malay) spoken by the Straits Chinese, and still used a little. These have been of importance in the formation of SCE. It is generally agreed that SCE had its origins in the English-medium schools of the Straits Settlements in the early twentieth century (Platt & Weber 1980; Bloom 1986; Gupta 1994). Historiographic data from nineteenth and early twentieth century Singapore (Gupta 1994, forthcoming) shows how the roots of SCE were based in the particular ethnic mix of these schools, which in the nineteenth century were composed

Jan Tent and France Mugler, eds *SICOL: proceedings of the Second International Conference on Oceanic Linguistics:* vol.1, *Language Contact*, 43–57.
Pacific Linguistics, C–141, 1998.
© Anthea Fraser Gupta

Gupta, A.F. "Singapore Colloquial English or deviant Standard English?". In Tent, J. and Mugler, F. editors, *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.*
C-141:43-57. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.43*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.

predominantly of English and (pidgin/creole) Malay-speaking pupils (Europeans, Eurasians, Jews, Armenians and Straits Chinese) and where the teachers consisted roughly equally of Eurasians, Europeans and Indians. In the early twentieth century there was an influx of non-Chinese pupils who spoke varieties of Chinese. Female education in English is crucial for English to become a native language. Until well into the twentieth century female English-medium education was virtually restricted to the Europeans, Eurasians, Jews and Armenians. The first other girls to receive education in English were Straits Chinese. Again, the importance of the pidginised and creolised varieties of Malay emerges.

I would place myself in the old tradition (Weinreich 1953; Thomason & Kaufman 1988) that sees there as being a range of contact experiences, and that it is the processes and results of contact which are of especial interest. I don't want to get into arguments about whether SCE is or is not a creole. That it is a contact variety is undisputed. Rather than discussing what we might wish to mean by the terms 'creole' or 'creoloid', I prefer to see Singapore English in the context of languages or varieties which as a result of having been in a contact situation have "a form palpably different from either stock language" (Weinreich 1953:69). A natively spoken variety that carries the mark of any significant degree of discontinuity in development which results from contact can be seen as something like a creole in this sense. These varieties are referred to by Thomason and Kaufman (1988) as 'mixed languages': note that Singapore Standard English is not a contact variety.

Like many contact varieties, Singapore Colloquial English, usually referred to as 'Singlish', is the subject of ambivalent attention in Singapore. Singapore's former Prime Minister, Lee Kuan Yew, remarked in a speech to National University of Singapore students on 29 July 1994 (*The Sunday Times*, 31 July 1994) that Singlish is only for losers:

> I think it's important that you know the English language because it is the international language, and you speak it in the standard form.
>
> Do not speak Singlish! If you do, you are the loser. Only foreign academics like to write about it. You have to live with it. And your interlocutors, when they hear you, their ears go askew. You detract from the message that you're sending.
>
> I don't have to speak with an English upper-class accent. But I speak in a way which makes it easy for them to understand me and, therefore, they are not distracted by my background.

This foreign academic will continue to demonstrate her interest in writing about it....

2. MODES OF ANALYSIS

There are three main modes of analysis for Singapore English which have different theoretical implications. These are similar to the 'stages' of Bloom (1986) and of Foley (1988) but they are not chronological—the earliest studies, those of Killingley (1968, 1972), were in Mode (c).

(a) *Singapore English is imperfectly learnt Standard English*

Its features are errors. This approach attempts to identify interference features and describes Singapore English as a non-native variety. It can be found even in very recent books (e.g. Ko & Ho 1992). In this mode of analysis the contact variety is not seen as

having its own grammars: Singapore English is an accretion of mistakes as a result of imperfect learning. I live in the hope that this is no longer seen as a respectable approach.

(b) *The lectal continuum approach*

This is the approach associated with the late John Platt and his students. Singapore English is not to be seen as deficient, but is most usefully analysed in terms of its difference, or deviance, from better known varieties, such as British Standard English (all the work by Platt and his students; also Tay & Gupta 1983; Gupta 1986). This approach lends itself to a deviationist approach (near approach (a)) which treats features of dialects other than StdE/RP as deviations. For example, faced with Singapore English [o] in *know*, a follower of this approach would say that the RP diphthong had been monophthongised.

This approach treats the full range of proficiency. The usual research methodology is a modified Labovian one, in which large numbers of respondents, stratified by educational attainment, are recorded in an interview setting (Ho & Platt (1993) used 100 respondents). The Platt school approach tends to describe Singapore English as a non-native variety: respondents are not generally queried about their language backgrounds, as native speakers of Singapore English are given no special place in this analysis—the native speakers of English are all in countries like England and the USA. The Platt school use post-creole continuum terms—acrolect, mesolect, basilect. The acrolect is taken to be an idealised Standard British English—what Ho and Platt (1993:12) call "a prescriptive native speaker standard". 'Speakers' are classified as acrolectal, mesolectal or basilectal, according to their educational level. They are not recorded at different levels of formality.

Singapore English is described as a non-native variety, regardless of its status for the individual speaker. Its features are developmentally ordered, and are regarded as the result of 'mother tongue interference' and 'over-generalisation'.

(c) *The diglossia approach*

Singapore English is best analysed in relation to Singapore only. Standard English is locally established, and is not significantly different from other standard Englishes. The most informal variety of Singapore English, however, is a variety of English so different from StdE that it must be analysed entirely on its own terms (e.g. Gupta 1989, 1991, 1992a, 1992b, 1994; Kwan-Terry 1978, 1989, 1992; Pakir 1991; Alsagoff 1995; Gil 1994). Its grammar must be established by an examination of the usage and/or judgements of its native speakers. This approach artificially establishes a polarisation, identifies 'native speakers' and tends to ignore speakers of low proficiency. Native speakers are those who grew up with English as a primary language before attending school. Many researchers extend study to all proficient users whether native speakers or not, regarding as proficient someone who shows control of StdE in formal situations. I pay especial attention to native-speaker children over the age of five, on the grounds that they have usually not acquired StdE yet, and may thus be monovarietal speakers of L. Tay (1991) has criticised this on the grounds that I thus equate the L variety with baby talk.

Methodology is usually small scale and non-quantitative, often using data from a variety of sources including recordings in informal or domestic settings, overheard speech, elicited speech and judgements.

The diglossia approach uses terms H (StdE) and L (variously known as SCE/Colloquial Singapore English/Singlish).[1] Varieties are defined by both use and usage. Functionally, L is the variety used by either native speakers above the age of four or five, or used informally by speakers who also have the option of speaking StdE. Formally it is a variety sharply different in specified ways (especially in syntax) from StdE. This is the only approach possible if an aim is the analysis of an autonomous variety different from StdE.

3. COMPARISON OF STYLE OF ANALYSIS

My comparison is between the way approaches (b) and (c) have actually been used. It would be perfectly possible to have a lectal continuum approach which had some differences from the one actually used by the Platt school. The differences are summarised in the Appendix.

3.1 LECTAL CONTINUUM APPROACH

In Ho and Platt (1993) structures are analysed by their relationship to their StdE paraphrases, although in many cases there may be alternative paraphrases. Ho and Platt classify respondents according to their level of education. The examples below are taken from the section on "*Clause* + *be* and *be* + *clause* environments". Ho and Platt indicate that "[c]lauses used by informants with tertiary level education are generally similar to those used in SBrE [Standard British English]" (p.69), while among those with A level qualifications,[2] clauses are on the whole "fairly stable" (p.70) except where "gerunds" are concerned (e.g. *Give dem the basic foundation is important*). The examples below are, therefore, all from those with a maximum of O-Level qualifications.

WH clauses

(1) *dey offer you is four plus*
what they offer you in salary

Noun Phrase + *be* + Clause

(2) *Dey employ most of they are Malaysians.*
Most of the staff employed are Malaysians.

(3) *I like very much is de Domestic teacher.*
The teacher I liked best was the Domestic Science teacher.

---

1 Some scholars using the diglossia approach use the term *Singlish* while others avoid it. This is because *Singlish* is commonly used to mean both the informal variety of those proficient in English and also the low-proficiency learner variety. As the establishment of an autonomous variety precludes the analysis of the usage of non-native or low-proficiency speakers, it is important to those of us using the diglossia approach to separate these two meanings. The term *Singlish* is officially defined by SBC (*Straits Times*, 10 July 1993) as "ungrammatical English spoken by those with a poor command of English".

2 In the Singapore education system, A-levels are the external examinations taken at age 18, which are required for admission to university courses. O-Levels are the external examinations taken at age 16—a certain standard must be attained to allow students to progress to A-Level.

To-infinitive clauses

(4) *The purpose is have connections.*[3]

(5) *Earn dere easy, spend also de same.*
To earn a living there is easy, to spend all your money there is just as easy.

Gerunds

(6) *Work shift is very tired.*
Working on the night shift is very tiring.

Some diglossia approach proponents (like me) object to the use of 'eye dialect'—as this seems to be part of the deviance approach—here used for phonology. I would use either normal orthography or a broad phonetic transcription. Many of these sentences would not be regarded as bona fide L by proponents of the diglossia approach, as they would not be used in the L of native or proficient speakers. Let us put that aside for the moment: if they were L, the autonomy approach would discuss these utterances in terms of their structure in SCE. The basic description *Clause + be* would not be used, as, given the possibility of zero copula, there may well be no BE in the utterance (as in (5)). In (1), (3), (4), (5) and (6) there are noun clause structures, where finite clauses are functioning as the subject of an equational sentence. The terms *gerund*, *wh-clause* etc. will not be used, as they are artefacts of translation. There has probably been PRO-drop in (6); lack of context does not allow retrieval here, and the sentence may not be well formed in SCE. The noun clause in (4) may be an imperative; again context would provide the full structure. In (2) we seem to have a totally different structure, with relative clause (*Dey employ*), of which the head may be *they*. I do not see that *Dey employ most of they* can be a clause, as *most of they* is the subject of *are*.

3.2 DIGLOSSIA APPROACH

SCE should not be seen as *really* autonomous and this undoubtedly causes problems. When I first introduced the treatment of SCE as an autonomous variety, in 1988, in the paper later published in 1992, I had a prolonged debate with Le Page, who has attacked the autonomous concept as unrealistic—which it is. I very much agree with the analysis of Carrington (1993) which makes all competence varilingual. But I still think it is useful to focus on the grammatical structures in a way which is possible only in an abstracted delimited variety. Those who (unlike me) use a syntax based on speaker judgements have an even greater need to establish autonomy. The pseudo-autonomy of the diglossia approach will present them with some real problems, because Singapore's diglossia is very leaky indeed. The main area of leakiness is the use of H where a strict Ferguson definition (1959) would predict L.

In actual usage, speakers move across their varietal range in a way that is socially meaningful. Although individual utterances or stretches of discourse may be focused (on formal grounds) towards either StdE or SCE, discourses seldom sustain SCE for long continuous periods.

---

3 No gloss provided in original. The heading suggests that the gloss should be "The purpose is to have connections".

Someone asks a woman whether her son is doing his National Service. Among other things, she can answer:

(7) *Finish.*

(8) *Finish already.*

(9) *Finish already lah.*

(10) *Finish lah.*

(11) *Finished.*

(12) *Finished already.*

(13) *Finished already lah.*

(14) *Finished lah.*

(15) *He's finished already lah.*

(16) *He's finished lah.*

(17) *He's finished already.*

(18) *He's finished.*

She can't say:

(19)a. **Lah finish already.*

b. **Finish all-gone.*

c. **All-gone finish.*

d. **He done finish.*

e. **Done finish.*

f. **Has finished.*

The StdE forms here are:

(i) HAVE perfective

(ii) past participle morphology

The SCE forms are:

(i) PRO-drop

(ii) use of pragmatic particle

(iii) *already* as aspectual marker

However, structures may combine the StdE perfective with the SCE particle. The pronunciation of the final stop of the past participle is more likely to be present if the auxiliary is there but it may appear without the auxiliary, and may be absent with it.

The speaker's choices will be related to all sorts of variables which I won't go into (such as how the question was asked, and her relationship with the speaker). But the crucial question for the autonomy issue is: how many of these do we want to be from the same pseudo-autonomous variety, SCE? (1) to (8) are definitely SCE, and (18) is definitely StdE.

Many Singaporeans regard virtually any use of *already* as SCE, so may not regard (17) as StdE. Unless we want to accept that anything which is grammatical in StdE is grammatical in SCE (which is how speakers really see the situation) we have to artificially exclude certain StdE systems from the grammar of SCE. The artificiality of doing this makes reliance on judgement difficult.

As has been observed by many scholars in a range of contexts (for example, Bickerton 1977), in making judgements people who know a standard variety and are aware of concepts of correctness are liable to make their judgements on the basis of a sense of correctness which is often dependent on the formal written language. If asked about grammaticality this speaker will probably only accept *He's finished* as grammatical. It may or may not be possible to get her to admit that she sometimes uses the others. Many speakers will deny that educated speakers ever use these structures, even when faced with tape recordings. What people actually say even in the colloquial forms of StdE is much freer. In spoken usage, extralinguistic contextual information is vital. Conversations are built up over a period of time based on shared information. In non-standard varieties, and in creoles, there is an even greater gulf between what people say and what they consciously regard as correct. Also, where there is not a full range of functions, and especially where the formal functions are missing, there is no expanded, non-contextually based variety to which speakers can have recourse. This is likely to lead to rejection of sentences that are common in spoken usage. For example, a Singaporean graduate student who read my paper on pragmatic particles (Gupta 1992b) claimed that the attested utterance *Why lah?* (a child responding to his mother's prohibition on his going out) was impossible. He rejected the use of the particle *lah* in an interrogative. There is nothing unusual about using *lah* with *why*, where the use of this assertive particle marks the interrogative as an assertive, non-question, some sort of rebuke (= 'Why-ever did you put that dish in the oven?'). A reliance on judgements which pays no attention to usage runs the risk of producing a grammar based on stereotypical behaviour which does not encompass the real richness with which speakers operate under natural conditions. So far, the only linguist to have used judgements systematically has been David Gil, who has approached this issue with great caution and achieved many useful insights by working with informants.

I will give some examples of the diglossia analysis using my own data, as I do in Gupta (1994)—although these examples are not from the book itself as I prefer to give additional material. The context does not strictly determine the choice of L or H—an example of the leakiness of diglossia in Singapore English. Speakers can switch between H and L in the same context to achieve highlighted effect.

In a conversation between a 4:3-year-old girl (EG) and her mother (MG), we can see how the child consistently uses the L variety, while her mother moves from H to L.[4]

(20) 1. MG: Who are these?

2. EG: David.

3. MG: Which one is David?

[EG points to a child in the photograph]

---

4 EG is looking at a photograph taken at her kindergarten. *Botak* is a loan word from Malay, meaning 'bald'. EG's father's name is also 'David'.

4. MG: Yes.
5. AG: David?
6. We know a David don't we?
7. EG: Mummy our friend ah, our friend also David, our friend.
8. MG: Who else is David?
9. EG: David ah, *botak* head already *lah*, last time didn't, now *botak* head.

[this David's hair has been cut very short]

10. MG: Is it?
11. Why?
12. EG: He-| he |
13. MG: | His |
14. EG: Mummy want to cut.
15. MG: His mummy want to cut his hair?
16. EG: M.
17. AG: Oh so | *botak*. |
18. EG: | Jia Ting- |
19. Jia Ting also *botak* head.
20. MG: Oh | dear. |
21. AG: | Oh dear |
22. EG: Mandarin.

[here EG is identifying the language of the name *Jia Ting*]

23. MG: Oh.
24. EG: Then David is English.
25. MG: M.
26. Who else is David?
27. EG: Our school class *lor*. [very high rising]
28. MG: Some more—your Daddy *leh?*

By line 24, EG cannot understand why her mother is pestering her so much about David. Her *lor* in line 27, with its very high rising tone, appears to be expressing surprise that MG should imagine there is another David in her class. Her mother switches to L for line 28 using the grammatical particle *leh*, which marks an x-interrogative and can be glossed as 'What about-'. A contingent answer would be *My Daddy David.* She does not wait for an answer, but turns to me to tell me an anecdote about EG.

The following are two birthday greetings taken from the personal room of the pre-Internet NUS bulletin board. The first uses StdE of the sort that can be found in electronic

communication anywhere (there are no omissions—dots are in the original). Both these texts were produced by users who would be described as 'acrolectal speakers' by the Platt school.

(21) Birthday wishes.....

Announcement!! Announcement!!

On the 12th March 1992, My best friend birthday....

She is.....Miss Lim Siew Yian....

Grasping this chance..... I want to wish her A Very Happy 20th....

But it will be hard on her 'cos there will be a last paper for her on the 13th march.... Anyway i hope that she can have some fun on her birthday though she has to study for the paper tomorrow....

I also wish her good luck for her coming examinations...

May her coming birthday brings her good luck for her examinations in a few days time.....

May those who know her..... Wish her a happy 20 birthday.....

May God bless her.....

Her best friend..... Someone who know her for about 8 years...

Bye.... Take Care and good luck.

(Tay Mui Kheng, 24 February 1992)

This is a student not very secure in the morphology of StdE (*friend birthday*, *may X brings her good luck*, *someone who know*, *a few days time*). The presence of inflections is more significant than their absence, so the hypercorrect 'brings' is one of several clear indications, like the complex verb groups, that this writer is aiming at StdE. This is what the Platt school would identify as mesolectal—as it includes variable morphology, a features of the L variety that here results from lack of control of StdE.

Other birthday greetings may focus, entirely or in part, on SCE:

(22) Thank you, sifu, [*master/teacher*, here jokingly implying a disciple-teacher relationship between fellow students] for the Bday greeting

but sifu, you very the economical hor

one greeting to me n you ask for 5 in return

think I better learn from u

if not send entries send until blur

by the way, dont worry lah

I am not going to claim angpow[5] from u

well.....not yet.....hehe

who knows, may be soon u'll have to give angpows.

(Tan Hwee Eng, 8 February 1992)

This greeting begins with a focus on SCE, and illustrates three of the four characteristic elements of SCE (subject deletion, pragmatic particles *hor* and *lah*, and zero copula) but there is use of StdE features too in the last three lines (complex verb groups and noun and verb morphology). The motivation of this kind of rapid switching is not always clear, as is the

5 A gift of money in red envelope, normally given only to juniors, and also at weddings.

case here. I have found switches from SCE to StdE associated with a switch to the pedagogical mode in speech to children, and switches to SCE from StdE associated with clarification of StdE utterance that were not responded to, or with an increased emotional involvement. Pakir (1991:173) also has examples of such switches, associated with emotional responses. This text would presumably be described as basilectal by the Platt school, as it includes many features of their basilect.

There are numerous uses of the L variety in literature, used as one might expect, for characterisation, humour, and to show shifts of attitude. A popular comedy tape in Singapore (*Why U so like dat?* by Siva Choy), released in late 1990, and a runaway best-seller, makes much play of use of English. In one sketch, a disk jockey, 'DJ de Souza' (this name would be either Eurasian or Indian), is talking on the air, with Standard English and an exaggerated American accent. When his wife rings, as listeners have been invited to, he changes accents, and syntax:

(23) [Song over. Phone rings]

Hi there. This is old smiley face de Souza himself. And to what do I owe the pleasure of this phone call?

Alamak Yvonne, how many times I've told you don't call when I'm working ah. You think the other deejays' wives call them when they're working?

This is not as extreme a variety of SCE as the one used by Tan Hwee Eng on the bulletin board, with the DJ's lower variety still having some standard markers, particularly morphological ones, but there is a clear indication of a switch to the low variety.

4. PROS AND CONS

In common to both approaches is an agreement that individuals move across Englishes, and that the main differences between speakers and between styles are in syntax.

A diagram often seen in the Platt school texts is something like the figure on the following page.

Diglossia approach scholars would have some reservations about this diagram, which equates the informal variety of proficient speakers with the more limited range of less proficient speakers. We are more likely to see the continua of proficiency and formality as to some extent separate. We think that it is possible for SCE to be ungrammatical if spoken by someone whose native language it is not.

Ho and Platt (1993:15) refer to "some academics in Singapore who feel that research should concentrate on the speech and writing of the more educated members of society", and they defend the importance of "basilectal SE" for the understanding of how second language acquisition functions. It does so happen that at present native speakers of English are disproportionately from higher social classes, although the native-speaking children are not (as yet) highly educated. I agree with Ho and Platt that "the whole range of SgE" needs to be studied. It is true that the restriction to native speakers means that we are ignoring a large proportion of the English users of Singapore. However, it is also legitimate to look at the syntactic structure varieties uncluttered by issues of second language acquisition in the individual speaker.

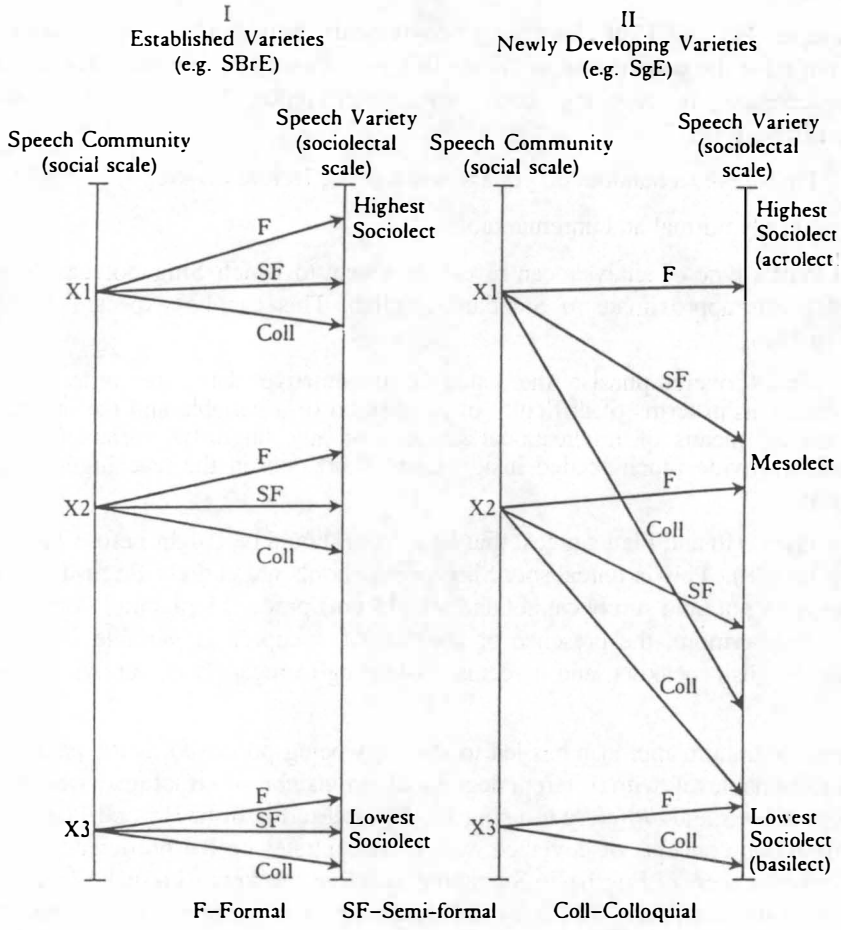


FIGURE: STYLISTIC AND SOCIAL VARIATION IN SINGAPORE ENGLISH
(HO AND PLATT 1993:5)

I don't see that these two approaches must be in opposition. One focuses on native speakers, and on their style shift, while the other focuses on the behaviour of all speakers in a formal context. The diglossia approach does not lend itself very well to large-scale quantitative research, to which the Platt school is committed. Surely both quantitative and non-quantitative approaches have a role.

There are intrinsic difficulties in a deviance approach. A lot of us have been caught out by this. I first came across the term *wash up* ('You look warm—go and wash up before we go out.') in Singapore. In British English *wash up* can only refer to washing dishes, not people. I therefore assumed that it was Singapore English (Shields 1977). Later I learnt that this use is widespread in American English. I once heard a talk on the English of the Philippines in which *overtake* ('The car overtook me at high speed.') was described as a Philippines English deviation—the speaker was not familiar with the usage in many varieties of English

and claimed that only *pass* was possible in American English. One is constantly stopped in one's tracks by these pseudo-deviances in the work of the lectal continuum school.

For example, Ho and Platt classify as non-standard Singaporean English uses of *take* which do not have the connotation of taking things without permission. They describe the following sentence as resulting both from interference from Chinese and from 'overgeneralisation' (p.170):

(24) Please *take* a handout on your way out of the lecture theatre.

I find this perfectly normal and unremarkable.

Ho and Platt's type of analysis can reveal the extent to which Singaporeans of different educational levels approximate to Standard English. This can be expected to feed into pedagogy (p.187):

> One cannot overemphasise the value of quantitative data: the ordering of environments in terms of difficulty of acquisition of a variable and the ordering of data by means of implicational scaling for any linguistic variable(s) can certainly provide much-needed insights and short cuts in the teaching/learning process.

As a corollary, Ho and Platt suggest that *be + Nom* should be taught before *be + Adj* and *be + V-ing* (p.188). This is unrealistic: Singapore schools are entirely English-medium and the syllabus does not (and surely cannot and should not) proceed structurally on an item-by-item basis. Furthermore, the presence or absence of a copula is variable in virtually all Singaporean English speakers, and it seems misleading to regard it in purely developmental terms.

The lectal continuum approach has led to speakers being portrayed as the passive victims of their educational level, with different degrees of *acquisition* of structures, who experience *confusion*, *problems* and *difficulty* in using English (all terms from Ho and Platt 1993). An approach based on concepts of deviance, which concentrates on the proficiency continuum, cannot portray speakers of English in Singapore as active speakers who make *acts of identity* (Le Page & Tabouret-Keller 1985) by manipulating their language repertoires in a rich sociolinguistic milieu. On the other hand, the diglossia approach, which does just this, ignores those people who really do experience *confusion*, *problems* and *difficulty* in using English.

At the moment I will stick with the diglossia approach, which continues to work for me. But in due course the whole spectrum of Singapore English will need to be looked at. The decision which to choose seems to be a matter of your own research priorities—for me it is to get a greater understanding of the syntax of the contact variety, SCE.

APPENDIX: SUMMARY OF DIFFERENCES

Issue	Lectal continuum	Diglossia
Image of varieties of SingE	Two dimensional continuum of acrolect – mesolects – basilect.	Proficiency continuum and formality continuum separated.
Who?	John Platt, Heidi Weber, Ho Mian Lian, etc.	Jack Richards, Anthea Fraser Gupta, David Gil, etc.
Syntax?	Deviations from StdBrE	Autonomy of L.
Phonology?	Deviations from RP. Use of eye dialect.	Attempt to establish phonemic system of SingE.
Informants?	All English educated Singaporeans.	Only Singaporeans who have English as a native language or are deemed to be of same proficiency as native speakers.
Source of data?	Large numbers of informants, stratified by education, but not necessarily selected by statistical sampling. Labovian interviews.	Small number of informants, unsystematically selected (except for native speaker status). Informal conversation, overheard data, speaker judgements.
Analysis of data?	Systematic Labovian analysis.	Often impressionistic. Seldom quantitative.
Nature of data?	Fairly formal speech.	Greatest interest in most informal varieties.
What is SingE?	A non-native variety. A creoloid.	A dialect of English. SCE is a contact variety/creole/ creoloid.
Central interest?	Effects of second language acquisition.	Syntactic structures of L.

REFERENCES

Alsagoff, Lubna, 1995, Colloquial Singapore English: the relative clause construction. In Teng Su Ching and Ho Mian Lian, eds *The English language in Singapore: implications for teaching*, 77–87. Singapore: Singapore Association for Applied Linguistics/Singapore Teachers Union.

Bickerton, Derek, 1977, Some problems of grammaticality and acceptability in pidgins and creoles. In Sidney Greenbaum, ed. *Acceptability in language*, 27–37. The Hague: Mouton.

Bloom, David, 1986, The English language and Singapore: a critical survey. In Basant K. Kapur, ed. *Singapore studies: critical surveys of the humanities and social sciences*, 337–458. Singapore: Singapore University Press.

Carrington, Lawrence D., 1993, Creole space—a rich sample of competence. *Journal of Pidgin and Creole Languages* 8/2:227–236.

Cheshire, Jenny, ed., 1991, *English around the world: sociolinguistic perspectives*. Cambridge: Cambridge University Press.

Crewe, W.J., ed., 1977, *The English language in Singapore*. Singapore: Eastern Universities Press.

Ferguson, Charles, 1959, Diglossia. *Word* 15:325–340.

Foley, Joseph, 1988, Studies in Singapore English: looking back and looking forward. In Joseph Foley, ed. *New Englishes: the case of Singapore*, 1–26. Singapore: Singapore University Press.

Gil, David, 1994, Genitives, number and (in)definiteness: some data from English, Singlish, Mandarin and Malay. MS.

Gupta, Anthea Fraser, 1986, A standard for written Singapore English? *English World-Wide* 7/1:75–99.

1989, Singapore Colloquial English and Standard English. *Singapore Journal of Education* 10/2:33–39.

1991, Acquisition of diglossia in Singapore English. In Anna Kwan-Terry, ed. *Child language development in Singapore and Malaysia*, 119–160. Singapore: Singapore University Press.

1992a, Contact features of Singapore Colloquial English. In Kingsley Bolton and Helen Kwok, eds *Sociolinguistics today: international perspectives*, 323–345. London: Routledge.

1992b, The pragmatic particles of Singapore Colloquial English. *Journal of Pragmatics* 17/3:39–65.

1994, *The step-tongue: children's English in Singapore*. Clevedon: Multilingual Matters.

forthcoming, English and Empire. In N.M. Mercer and J. Swann, eds *Learning English*. London: Routledge.

Ho Mian-Lian and John T. Platt, 1993, *Dynamics of a contact continuum: Singaporean English*. Oxford: Clarendon Press.

Killingley, Siew Yue, 1968, The phonology of Malayan English. *Orbis* 17/1:57–87.

1972, Clause and sentence types in Malayan English. *Orbis* 21/2:537–548.

Ko Peng Sim and Ho Wah Kam, eds, 1992, *Growing up in Singapore: the pre-school years*. Singapore: Longman.

Kwan-Terry, Anna, 1978, The meaning and the source of the 'la' and the 'what' particles in Singapore English. *RELC Journal* 9/2:22–36.

1989, The specification of stage by a child learning English and Cantonese simultaneously: a study of acquisition processes. In Hans W. Dechert and Manfred Raupach, eds *Interlingual processes*, 33–48. Tübingen: Gunter Narr.

1992, Code-switching and code-mixing: the case of a child learning English and Chinese simultaneously. *Journal of Multilingual and Multicultural Development* 13/3:243–259.

Le Page, Robert B., 1988, Some premises concerning the standardization of languages, with special reference to Caribbean Creole English. In John R. Rickford, ed. *Sociolinguistics and pidgin-creole studies. International Journal of the Sociology of Language* 71:25–36. The Hague: Mouton.

Le Page, R.B. and Andrée Tabouret-Keller, 1985, *Acts of identity: creole-based approaches to language and ethnicity.* Cambridge: Cambridge University Press.

Pakir, Anne, 1991, The range and depth of English-knowing bilinguals in Singapore. *World Englishes* 10/2:167–179.

Platt, John, 1987, Communicative functions of particles in Singapore English. In Ross Steele and Terry Threadgold, eds *Language topics: essays in honour of Michael Halliday*, vol. 1, 391–401. Amsterdam: Benjamins.

1991, Social and linguistic constraints on variation in the use of two grammatical variables in Singapore English. In Cheshire, ed. 1991:376–387.

Platt, John and Heidi Weber, 1980, *English in Singapore and Malaysia:status, features, functions.* Kuala Lumpur: Oxford University Press.

Richards, Jack C., 1977, Variation in Singapore English. In Crewe, ed. 1977:68–81.

1983, Singapore English: rhetorical and communicative styles. In Braj B. Kachru, ed. *The other tongue: English across cultures*, 154–167. Oxford: Pergamon Press.

Shields, Anthea Fraser, 1977, On the identification of Singapore English vocabulary. In Crewe, ed. 1977:120–140.

Siegel, Jeff, 1987, *Language contact in a plantation environment: a sociolinguistic history of Fiji.* Cambridge: Cambridge University Press.

Tay, Mary W.J., 1991, Southeast Asia and Hongkong. In Cheshire, ed. 1991:319–332.

Tay, Mary W.J. and Anthea Fraser Gupta, 1983, Towards a description of Standard Singapore English. In R. B. Noss, ed. *Varieties of English in Southeast Asia*, 173–189. SEAMEO Regional Language Centre Anthology Series 11. Singapore: Singapore University Press.

Thomason, Sarah Grey and Terrence Kaufman, 1988, *Language contact, creolization, and genetic linguistics.* Berkeley: University of California Press.

Weinreich, Uriel, 1953, *Languages in contact: findings and problems.* The Hague: Mouton.

SAINT-CHRISTOPHE: THE ORIGIN OF FRENCH ANTILLEAN CREOLES

WILLIAM JENNINGS

The small Lesser Antilles island of Saint-Christophe, now known in English as St Kitts, was the first French slave-based colony in the Caribbean. Slaves of African origin were in contact with the French on the island from about 1626. A study of censuses and other contemporary historical records allows the language contact situation on the island to be described in detail. It is argued that the linguistic environment of Saint-Christophe produced a creole, defined not by typological factors, but by demographic and socio-historic ones. The linguistic environment of the colony was transmitted to other colonies in the region and is the origin of the French Antillean creoles.

1. HISTORY

Saint-Christophe was a French slave-owning colony for only three generations, yet the origin of all lexically French-based Caribbean creoles. The island was also the first English colony in the Caribbean, being settled by Thomas Warner in January 1623 (Moreau 1992:187) following the success of an experimental crop of tobacco planted there in 1622 (Hamshere 1972:27). The French did not show any interest in the island until the Norman privateer d'Esnambuc emulated Warner in 1625, and sailed with a tobacco harvest to Paris, leaving "eighty men and also about forty slaves" on the island to safeguard his interests (Petitjean Roget 1978:1496). The slaves, France's first African slaves in the Americas, must have been taken from the Spanish by the privateer. Most of the eighty men were Normans, for d'Esnambuc's crew, when he left France in 1623, was mainly from Normandy.

D'Esnambuc's tobacco profit led Richelieu to give him official backing to found a colony on Saint-Christophe in late 1626. The Norman decided that the settlement's workers would be recruited from his home province, rather than taken from Africa (Margry 1863:28). Labourers signed on for a three-year term, with the promise of an allotment in the colony at the end of their indenture time. Two-thirds of the *engagés* (indentured labourers) whose origin is known were Normans; 10% were Parisians, 7% from French ports, and the rest, bar one Portuguese, from other French provinces (Petitjean Roget 1978:61).

Three ships left Le Havre in February 1627 with 532 people, mainly *engagés*, on board. They landed at Saint-Christophe in May, and the crew immediately made peace with the English (Abenon 1992:21). Both sides feared the superior numbers of the Caribs on the island, but overcame their apprehension in 1628 by massacring them in a surprise attack

Jan Tent and France Mugler, eds *SICOL: proceedings of the Second International Conference on Oceanic Linguistics:* vol.1, *Language Contact*, 59–68.
Pacific Linguistics, C–141, 1998.
© William Jennings

Jennings, W. "Saint-Christophe: the origin of French Antillean Creoles". In Tent, J. and Mugler, F. editors, *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.*
C-141:59-68. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.59*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.

(Pigneſk 1682:3). A Spanish raid in the same year forced the abandonment of Saint-Christophe (Coppier 1645:37), and while most of the French and English returned after a few months, some of the French founded a settlement on Tortuga, which led ultimately to the colony of Saint-Domingue (Haiti).

In the early 1630s Saint-Christophe grew rapidly. Indentured labour proved ineffective, since the labourers protested that they were treated like slaves and severely punished for the slightest misdemeanour (Casta-Lumio 1906:67). Indeed, one historian has described them as "white slaves" (Goslinga 1971:354). The planters' reaction was to buy black slaves to replace them. Thousands of African captives were taken across the Atlantic and sold in Saint-Christophe. Most were carried by Dutch traders, who dominated the Caribbean in commerce, if not territory, at that time (Coppier 1645:32; Du Tertre 1667–71, II:462). With the rush of settlers from France came the first missionaries. They were the principal chroniclers of daily life in the colonies, for official correspondence tended to dwell on military and economic matters. Father Bouton (1640) was the first to publish an account of life in the French Antilles and other priests followed, notably Pelleprat (1655) and du Tertre (1654, 1667–71), each devoting chapters to slavery.

A shortage of arable land led d'Esnambuc to expand his colony to Martinique in 1635 (Petitjean Roget 1978:5–6). In the same year, an expedition from France settled in Guadeloupe (Abenon 1992:25–26). Overproduction of tobacco caused prices to tumble and tobacco-growing on Saint-Christophe was banned in 1639; many planters went to Guadeloupe with their slaves and planted tobacco there. Saint-Christophe reached its carrying capacity in the early 1640s. The slave population would have reached its maximum in the late 1640s or 1650s (see Table 1 below), and slavers would have turned to other colonies to sell their human cargo. Once this maximum was reached, only a few slave ships a year were needed; enough for planters to replace the slaves who had been worked to death. The lack of land on the island no doubt contributed to the rebellion of 1645 that forced some settlers to flee; they went as far as the Guianas in attempts to found their own colonies.

By the 1650s there were slaves, settlers and former indentured labourers from the island scattered throughout the French Caribbean. Although Saint-Christophe was experiencing zero growth, unlike Martinique, Guadeloupe and other colonies, it still had the prestige of being the first settlement and the capital of the Antilles. However, this situation was to change with the arrival of new immigrants and a boom in a new crop: sugar. The loss of Holland's Brazilian colony in 1654 created Portuguese Jewish and Dutch refugees, many of whom were experienced sugar refiners (see Jennings 1995). Several hundred of them arrived with slaves in Guadeloupe and were well received. Sugar refining was a complex process, requiring a much higher labour input and larger plantations than other crops. It boomed in Guadeloupe and Martinique, but on overcrowded Saint-Christophe, the transition from tobacco to sugar was very difficult.

War in Europe spread to the Caribbean in 1666, forcing a battle between English and French forces for Saint-Christophe that saw the English expelled from the island. Several thousand slaves fought alongside their owners on each side, according to a possibly exaggerated eyewitness account (C10B/1, juin 1666).[1] However, the 1667 Treaty of Breda stipulated that France must return to the English the land it had won (Crouse 1966:83). It took until the early 1670s to iron out financial and legal obstacles, and only then did the two

[1] C10B/1 refers to register 1 of sub-section C10B of the French National Archives in Paris. Similar references also indicate archival sources.

colonies coexist as before, although for only nineteen years. The French expelled the English in 1689, but their victory was short-lived, for the English conquered the entire island the following year. "The greater part" of the French were sent to Saint-Domingue, where it was hoped they would settle and not try to return to the Lesser Antilles (Crouse 1966:166). The "lesser part"—mainly the governor and powerful landowners—were sent to Martinique (C8A/6). The French colony of Saint-Christophe was essentially at an end, aside from a brief occupation several years later. The slave-based plantation society which had existed there since the late 1620s, and which had had a reasonably stable population for three generations, would not exist again under French rule.

2. DEMOGRAPHY

The demography of a community is, like its history, a key to the description of its linguistic environment. In the case of Saint-Christophe, information on the numbers of inhabitants in the colony during the three generations that saw the creole arise is derived from three surviving censuses, as well as occasional reports.

TABLE 1: THE POPULATION OF FRENCH SAINT-CHRISTOPHE

Year	French	Slaves	Source
1626	80	40	1627 contract (in Petitjean-Roget 1978:1496)
1636	4,000–5,000	500–600	Margry (1863:44)
1640	8000		Petitjean-Roget (1978:611)
1666	5,000?	4,000?	C10B/1, juin 1666
1671	3,461	4,518	G1/471
1685	1,784	2,761	C10B/1 [for part of the colony only]
1687	3,192	4,470	C10B/1

Table 1 shows how the colony rapidly reached its carrying capacity of 8–9,000 people. The French population attained this level less than a generation after colonisation. It decreased after indentured labourers were replaced by African slaves. The first generation of Africans on Saint-Christophe were heavily outnumbered by the French, and children born to these slaves would have learned to speak in a linguistic environment that was almost entirely French. As these children grew up, they saw the African part of the population increase to the point where it was nearly on a par with the French population. Although there would eventually be more slaves than colonists, neither side would dominate significantly in terms of numbers.

After the initial dominance by the French population, there were roughly equal numbers of Africans and Europeans for about two generations, from about the late 1640s. Such a demographic history is significantly different from that of other French colonies where a creole language emerged. A typical colony would buy slaves in great numbers once the settlement had been established, and the colonists would find themselves heavily outnumbered within a decade or two of the founding of the settlement. In Cayenne, for example, slaves outnumbered their owners by about four to one within two decades of the arrival of the first slaver (Jennings 1993:32).

TABLE 2: THE 1671 AND 1687 CENSUSES

Year	Colonists				*Engagés*		African slaves			'Free coloureds'	Amer-Indians
	M	W	B	G	M	W	M	W	Ch		
1671	998	577	696	574	583	33	1,653	1,677	1,188	76	25
	Total: 3,461 Europeans						Total: 4,518 slaves			Total: 101 others	
1687	900	588	773	850		81	1,694	1,665	1,111	120	?
	Total: 3,192 Europeans						Total: 4,470 slaves			Total: 120+ others	

Of the three surviving censuses, the most comprehensive is that of 1671, which gives a breakdown of slaves by household. The other two censuses, from 1685 and 1687 (C10B/1), give totals only, but the 1685 census is not complete, and probably refers to only one of the French quarters of the island. In Table 2, the 1671 and 1687 totals are shown. Differences of classification can pose some problems: unmarried men may be counted under the heading B 'boys' or under M 'men'; likewise, unmarried women may be under G 'girls' or W 'women'. However, these difficulties are minor in comparison with the wealth of information the censuses provide.

The slave population of French Saint-Christophe in 1671 was almost identical to that of 1687, which shows that few slavers called at Saint-Christophe during that period. Mongin confirms the rarity of slavers when he wrote in May 1682 that slave ships called only a few times a year (Chatillon 1984:133–134). The colony did not need any more slaves, except to replace those who had died. If the slave population decreased by about 5% per annum on Saint-Christophe (see Debien 1964:27), them some 200–300 captives were bought by the colonists each year. Such a low rate of new arrivals would have favoured the stability of the island's contact language, for the new slaves could not have had much influence on the linguistic environment. They were pidginisers who sought to communicate with the people around them without trying to change the contact language.

There are two reasons for the equal ratio of African men to women in 1671 and 1687. Firstly, the French believed initially that slave numbers could be maintained through natural growth, and sought equal numbers of men and women slaves, hoping to establish a source of labour independent of the growing African slave trade, foreign interests and piracy. Contemporary observers on Saint-Christophe noted that planters encouraged slaves to form families. It was apparently not unusual to see an entire slave family over two generations working on the same plantation (Clodoré 1671:46). However, when it became apparent that the brutal working conditions of the slaves resulted in deaths greatly exceeding births, male captives came to be in much higher demand, and slavers would often cross the Atlantic with twice as many men as women on board (see Klein 1978:150). Secondly, once the population had reached its maximum, the low numbers of new arrivals meant that there was a high proportion of locally-born slaves who were, of course, born in roughly equal ratios.

Saint-Christophe also had a high number of European women from the 1630s onwards (Margry 1863:58), showing France's intention to use Saint-Christophe not so much as a get-rich-quick plantation settlement as a base for their colonial enterprises in the Caribbean. Colonies were usually male-dominated to the point where it became normal to take young women (*filles du Roy*) from prisons and orphanages in France and ship them across the

Atlantic where they were married within days of landing. Only among the *engagés* was there a high ratio of men to women; however, this system had almost ceased by 1687. Although it was significant for only half a century, the indentured labour system existed at a time when a language of communication between French and Africans had developed. The predominantly Norman labourers, who often went on to become landowners, worked with the African slaves and played an important role in developing this language. At the end of the three-year term, a labourer was often given a plot of land and remained in the colony as a settler; Coppier is one example of this (see Coppier 1645; Dampierre 1904:88), and many former labourers became important figures in Saint-Christophe or in other colonies.

Economic information from the 1671 and 1687 censuses also supplies information on the population of Saint-Christophe. There were only seventeen plantations with fifty or more slaves in 1671; the majority of slaves worked on smaller holdings of thirty slaves or less. In 1687 there were 101 *sucreries* ('sugar estates') and 62 indigo plantations, and an average of 27 slaves per plantation. Sugar mills were highly labour-intensive and needed at least about 50 slaves each to be efficient. This number was rarely attained in 1671 or 1687; big sugar mills with over a hundred slaves were an exception on Saint-Christophe. Furthermore, the colony was too small for plantations to be isolated, so there would have been a lot of mixing of slaves. This contrasts with other colonies where slaves could easily find themselves working on a large isolated plantation in a remote district.

In short, Saint-Christophe had a high population density with, after a generation of a predominantly French population, roughly equal populations of Europeans and slaves for some two generations. The smallness of the plantations and the low numbers of new slave arrivals suggest slaves had excellent access to the language of their owners compared to slaves in other French colonies.

3. LANGUAGE CONTACT

From the late 1620s, Saint-Christophe had a population of African-speaking slaves working in a French-speaking environment. The mutual incomprehensibility of the slaves' languages and the dialects used by the French resulted in a problem of communication which needed to be solved urgently. For the first few years there were only a few dozen slaves, who worked alongside Norman French labourers on small holdings. The variety of French the African captives heard in Saint-Christophe would have been heavily influenced by Norman French, and to a lesser extent by Maritime French (see Hull 1979) and may be referred to as colonial French. The slaves would have had good access to the language of their owners and co-workers, and the initial result of language contact would have been that the first slaves on Saint-Christophe learned to speak a variety of French as a second/third...language.

Later, when thousands of Africans were taken to Saint-Christophe during the 1630s and 1640s, the language contact situation changed. The first contact with French for many of these captives was on Saint-Christophe, for the Dutch dominated the trade at the time. It would appear in fact that captives had no knowledge of French even if transported by a French slaver; Mongin (in Chatillon 1984:133–134) and Clodoré (1671:43–44) state that interpreters were needed when the captives were first landed in the colony, and that newly-arrived slaves were placed with slaves who spoke the same African language (i.e. non-native speakers of French).

The urgent problem of communication was rapidly resolved through a compromise between the colonial French of the labourers and slave-owners, and the Africans' "way of speaking" (Pelleprat 1655:53). Bouton (1640:100) had noted a generation before Pelleprat that the French of the slaves lacked articles and "other particles". Indeed, the only articles present in attestations from Saint-Christophe are from stock-in-trade religious expressions like *le Bon Dieu* 'the Good Lord' or *la Bonne Vierge* 'the Holy Virgin' (Mongin 1682, in Chatillon 1984:104). The description of "other particles" may well include copula deletion, as in example (1), from a 14-year-old boy on Saint-Christophe in 1655:

(1) *Moy bien fâché.*
1SG very angry
I am very angry.

(Saint-Christophe: Pelleprat 1655:63)

Bouton may also have been referring to deletion of the unstressed clitic *ne*, which in French precedes the negated verb while *pas* follows it. Examples of negation on Saint-Christophe show that *pas* or an early variant *point* was retained, although it preceded the predicate.

(2) *Toi pas connaître moi.*
2SG NEG know 1SG
You don't know me.

(Saint-Christophe 1682; Chatillon 1984:96)

The boy whose speech was recorded by Pelleprat used *point* as the negative marker in five clauses; it always precedes the verb, but in some cases seems to negate the whole clause and not just the verb, as in examples (3) and (4).

(3) *Point luy iurer.*
NEG 3SG swear
He hasn't sworn.

(Saint-Christophe: Pelleprat 1655:63)

In fact, the word order of certain negated phrases may reveal a non-French influence in the language of the slaves of Saint-Christophe.

(4) *Point aller luy à femme d'autre.*
NEG go 3SG to woman of.other
He hasn't been with another woman.

(Saint-Christophe: Pelleprat 1655:63)

As far as the missionaries were concerned, however, the slaves spoke a jargon based purely on a simplified version of French:

> It consists of the infinitive of the verb, never conjugated, to which are added a few words to explain the tense and the person discussed. For example, if they wish to say 'I want to pray to God tomorrow', they will say 'me pray God tomorrow', 'me eat yesterday', 'you give food to me' and so on. This jargon is very easy to teach to the slaves and also to the missionaries so they can instruct them, and so it is used all the time. (Mongin 1682, in Chatillon 1984:34–35)[2]

---

2 "[Le jargon] est par l'infinitif du verbe, sans jamais le conjuger, en y ajoutant quelques mots qui font connaître le temps et la personne de qui l'on parle. Par exemple s'ils veulent dire: je veux prier Dieu demain, ils diront moi prier Dieu demain, moi manger hier, toi donner manger à moi, et ainsi en toutes

The linguistic compromise of the French-speaking inhabitants of Saint-Christophe may have been limited to the use of 'foreigner talk', but within the African-speaking community, the compromise was much greater, for there were immense linguistic differences among its members. Slavers bought captives from a region covering millions of square kilometres of the African continent; Pelleprat noted thirteen different languages in the slave community (1655:53), while Mongin suggested ten or a dozen (in Chatillon 1984:134). The chaos of so many different languages meant a language of compromise was needed within the slave community. The language used to communicate with French-speakers was the obvious choice, and would no doubt have been employed for interethnic communication within the slave community. This additional use of the French jargon would have led to the language being rapidly conventionalised, especially in a colony too small to allow any dialects to form, and in turn nativised by the children born to slaves on the island.

The language contact situation in Saint-Christophe can be summarised as follows: in the first generation of settlement, colonial French was spoken by all the settlement's inhabitants, including the few African slaves present. Later, the arrival of thousands of Africans in the space of a decade or two led to a jargon based principally on colonial French, with some structural features from other sources. This jargon was used for communication within the slave community as well as between slaves and owners and, as the next section shows, was transmitted to other French settlements in the Americas.

4. TRANSMISSION TO OTHER COLONIES

France's Caribbean colonies were not isolated entities. Ships crossing the Atlantic would often call at more than one colony before returning to France, and there is evidence of interaction among the colonists and slaves of different colonies and of a consequent linguistic transmission from Saint-Christophe to France's other colonies. Firstly, Saint-Christophe was the capital of the French Caribbean for three decades. Its frequent dealings with other French colonies of the region ensured homogeneity of the linguistic environment until the 1650s. The jargon used between slaves and French-speakers in the capital would have been the same used elsewhere in the French Caribbean in those early years. An attestation from Martinique in 1671 (Carden et al. 1990) shows many similarities with the language described by Mongin in 1682 and Pelleprat in 1655. The order of object pronouns is a case in point.

(5) *Toi pas connaître moi.*
2SG NEG know 1SG
You don't know me.
(Saint-Christophe 1682; Chatillon 1984:96)

(6) *Moi pas voir li.*
1SG NEG see 3SG
I didn't see him.
(The perfective aspect is context-specific)
(Martinique 1671; Carden et al. 1990:3)

---

sortes de choses. Ce jargon est fort aisé à apprendre aux nègres et aux missionnaires aussi pour les instruire, et ainsi ils le donnent à entendre pour toutes choses."

In the French versions of (5) and (6), the object pronoun precedes the verb. Furthermore, the strategy employed for negation in the Martinique text is the same as that used in Saint-Christophe; deletion of the French clitic *ne* while retaining and shifting the stronger *pas*.

When the colonies began to diverge in their demography the language situation would have changed. While numbers on Saint-Christophe remained stable for reasons of limited space, other islands saw a large increase in the slave population. The colonial French of the settlers would have remained stable, but the 'jargon' would have changed as increasing arrivals of slave ships brought thousands of African captives to Martinique, Guadeloupe and other colonies. The larger size of Guadeloupe and Martinique allowed dialects to arise in isolated areas. In Guadeloupe, the future marker had several variants until the late nineteenth century, which Hazaël-Massieux (1986:120) assigns to demographic differences. Abbé Goux's description (1842:20) of Martiniquan creole describes "the most widely-spoken language", which in itself is evidence of dialects. The 1671 Martiniquan text was based on an interview with three slaves in an isolated southern part of Martinique and differs in several respects from Mongin's and Pelleprat's attestations on Saint-Christophe. These differences are occasional non-French features where the Saint-Christophe texts retain a French structure. In Saint-Christophe, the French reflexive clitic pronoun *me* is used:

(7) *Moi me dépouiller.*
1SG 1SG undress
I undressed.
(The perfective aspect is context-specific)
(Saint-Christophe 1682; Chatillon 1984:104)

Three of the four Saint-Christophe slaves whose speech was recorded by missionaries use a reflexive, and in all cases the structure is identical to that of example (7). In the Martinique text, the structure differs markedly from French. Two of the three slaves interviewed use a plain pronoun as a reflexive.[3]

(8) *Li caché li.*
3SG hide 3SG
He hid himself.
(The perfective aspect is context-specific)
(Martinique 1671; Carden et al. 1990:4)

Another factor supporting linguistic transmission from Saint-Christophe to other colonies is that new colonies were often founded by people with experience on Saint-Christophe. The best example of this is the settlement of Martinique in 1635. Experience played a key role in the survival of a colony; for every successful settlement, there were countless failures, and the colonies that failed tended to be those founded by pilgrims and naïve adventurers who sailed from Europe.

The leading families of many of France's colonial administrators could trace their origins back to Saint-Christophe. Auger, a company commander in the 1671 census, was governor of Marie-Galante in the late 1680s. The Orvilliers family, which governed French Guiana for three generations in the 1700s, had its origins in Saint-Christophe; the first of the governors was stationed there, and the second was born and raised there. Many of the names encountered in the archives of Saint-Christophe are also to be found in the archives of other colonies. This family connection in the Caribbean is important in the context of linguistic

3 For a discussion of reflexives in French creoles, see Carden (1989).

transmission when it is remembered that the richest families had the most slaves. When an administrator was posted to another colony, many of the slaves went with their owners, and were in this way able to introduce some of the Saint-Christophe jargon to the slaves of the new colony.

Slaves were useful in other ways to the founders of new colonies. Experienced and reliable slaves were needed to get the settlement off to a good start, and also to act as interpreters for newly-arrived captives from Africa. The Jesuit priest Father Jean Grillet, who was sent to French Guiana in 1667, brought slaves with him from Saint-Christophe so that he might baptise and instruct the slaves of Cayenne before they died (Archives Françaises de la Compagnie de Jésus, FGu5).

The greater the distance from Saint-Christophe a colony was, the smaller the linguistic transmission was likely to be. Cayenne creole was, according to Barrère (1743:40), "less ridiculous" than—and therefore different to—the creole of the Islands. For this creole, a scenario of independent genesis with influences from Saint-Christophe would appear to be the case. In Louisiana, the creoles are separated by both distance and time from Saint-Christophe, and would appear to have been relatively independent in nature (Speedy 1994).

5. CONCLUSION

Good access to the colonial French of the small colony of Saint-Christophe meant that a relatively stable vehicular language of enslaved Africans arose rapidly. The locally-born children of the first slaves acquired this language, and it was transmitted to other French Caribbean colonies. As these colonies evolved separately, the contact language diverged and became the group of interrelated French Antillean creoles. Colonies further afield, such as Louisiana and Cayenne, may have been influenced by this initial creole, but to a lesser extent.

REFERENCES

Abenon, L-R., 1992, *Petite histoire de la Guadeloupe*. Paris: L'Harmattan.

Archives Françaises de la Compagnie de Jesus, Paris. (FGu series dealing with missionary correspondence from French Guiana.)

Barrere, P., 1743, Nouvelle relation de la France Equinoxale. Paris: Piget.

Bouton, J., 1640, *Relation de l'establissement des francois depuis l'an 1635 en l'Isle de Martinique, l'une des antilles de l'Amerique*. Paris: Cramoisy.

Carden, G., 1989, Reflexive constructions in *Atipa* and other early Guianese texts. In M. Fauquenoy, ed. Atipa *revisité ou les itinéraires de Parépou*, 161–182. Paris: l'Harmattan.

Carden, G., M. Goodman, R. Posner and W. Stewart, 1990, A 1671 French Creole text from Martinique. Paper presented at the Society for Pidgin and Creole Linguistics meeting, 5 January 1990.

Casta-Lumio, L., 1906, *Etude historique sur les origines de l'immigration réglementée dans nos anciennes colonies de la Réunion, la Guadeloupe, la Martinique et la Guyane*. Doctoral dissertation, Faculté de Droit de Paris.

Chatillon, M., 1984, *L'évangélisation des esclaves au XVIIe siècle: lettres du R.P. Mongin*. Extract of *Bulletin de la Société d'histoire de la Guadeloupe, 61 & 62*.

Clodoré, J. de, 1671, *Relation de ce qui s'est passé dans les Isles et Terre-Ferme de l'Amérique*. Paris: Clouzier.

Coppier, G., 1645, *Histoire et voyage aux Indes occidentales, et de plusieurs autres régions maritimes et esloignées.* Lyon: Huguetan.

Crouse, N., 1966, *The French struggle for the West Indies, 1665-1713.* New York: Octagon.

Dampierre, J. de, 1904, *Essai sur les sources de l'histoire des Antilles françaises (1492-1664).* Paris: Picard.

Debien, G., 1964, *Sur une sucrerie de la Guyane en 1690.* Dakar: Université de Dakar.

Debien, G. et al., 1978, Journal d'un voyage de traite en Guinée, à Cayenne et aux Antilles fait par Jean Barbot en 1678-1679. *Bulletin de l'Institute Français d'Afrique Noire* 40/2:235–395.

Du Tertre, J.-B., 1654, *Histoire générale des Isles de Christophe, de la Guadeloupe, de la Martinique et autres dans l'Amérique.* Paris: Langlois.

1667–71, *Histoire générale des Ant-isles de l'Amérique.* Paris: Jolly.

Goslinga, C., 1971, *The Dutch in the Caribbean and on the Wild Coast, 1580-1680.* Gainesville: University of Florida Press.

Goux, Abbé M., 1842, *Catéchisme en langue créole.* Paris: Vrayet de Surcy.

Hamshere, C., 1972, *The British in the Caribbean.* Cambridge, Mass: Harvard University Press.

Hazaël-Massieux, G., 1986, La créolisation est-elle un phénomène limité dans le temps? l'expression du futur en Guadeloupe. *Etudes créoles* 9/1:114-126.

Hull, A., 1979, On the origin and chronology of the French-based Creoles. In Ian F. Hancock, ed. *Readings in creole studies*, 201–216. Ghent: Story-Scientia.

Jennings, W., 1993, *La genèse du cayennais: étude de sa démographie et de l'évolution de son système verbal.* MA dissertation, University of Auckland.

1995, The first generations of a creole society: Cayenne 1660-1700. In Philip Baker, ed. *From contact to creole and beyond*, 21–40. London: University of Westminster Press.

Klein, Herbert S., 1978, *The middle passage:comparative studies in the Atlantic slave trade.* Princeton: Princeton University Press.

Margry, P., 1863, *Origines transatlantiques. Belain d'Esnambuc et les Normands aux Antilles.* Paris: Faure.

Moreau, J-P., 1992, *Les Petites Antilles de Christophe Colomb à Richelieu (1493-1635).* Paris: Karthala.

Pelleprat, P., 1655, *Relation des missions des PP. de la Compagnie de Jésus dans les Isles, et dans la Terre Ferme de l'Amérique Meridionale.* Paris: Cramoisy.

Petitjean Roget, J., 1978, *La société d'habitation à la Martinique: un demi-siècle de formation 1635-1685.* Doctoral dissertation, Université de Paris VII.

Pignefk, Rabbi, 1682, Lettre à son épouse sur Saint-Christophe. Paris: Bibliothèque Nationale.

Speedy, Karin, 1994, *Mississippi and Tèche Creole: a demographic and linguistic case for separate genesis in Louisiana.* MA dissertation, University of Auckland.

C10B/1, Section: Colonies, Sub-section C [correspondence from the colonies to France], Series 10B [Saint-Christophe]. French National Archives, Paris.

G1/471, Section: Colonies, Sub-section G1 [census reports from colonies to France]. French National Archives, Paris.

Jennings, W. "Saint-Christophe: the origin of French Antillean Creoles". In Tent, J. and Mugler, F. editors, *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.*
C-141:59-68. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.59*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.

UNEXPECTED SHARED FEATURES IN MELANESIAN PIDGINS/CREOLES: IS BROKEN A MELANESIAN PIDGIN/CREOLE?

ERNEST W. LEE

1. INTRODUCTION[1]

Torres Strait Creole or Broken (BR) has intrigued me ever since I heard the comment a number of years ago that speakers of BR consider it to be more like Solomon Islands Pijin (SP) than either Bislama (BI) or Tok Pisin (TP). It was not until 1993 that I learned about Shnukal's *Broken: an introduction to the creole language of Torres Strait* and was able to begin comparing SP and BR. A couple of obvious items shared by BR and SP which differ from BI and TP are the use of *nekstumora* (BR) and *nekes tumoro* (SP) 'day after tomorrow' and the use of *po* (BR) and *fo* (SP) 'for, in order to' where TP and BI use *blong*. There are, on the other hand, features shared only by BR and TP or BR and BI.

The substrate languages of BR are Meriam Mir, a Papuan language of the eastern islands, and Kala Laga Yau, of the Australian Pama-Nyungan family of the central and western islands. A large number of lexical items from both languages have found their way into the dialects of BR spoken in these areas and concerning the superficial resemblance of BR to English (ENG). Shnukal (1988:4) observes: "On a deeper level, however, both the system of meanings and the way the language is used resemble the traditional languages of Torres Strait much more than English." The eastern group has a smaller population, but the only specific information that Shnukal gives is that there are 2,500–3,000 first language speakers in the islands and roughly four times that many second language speakers, so that there may be as many as 15,000 total speakers, including many speakers in towns on Cape York in Queensland. BR first creolised in the east on Erub and Ugar in the 1890s and about fifteen or twenty years later in the west at St Paul's Anglican Mission on Moa following half a century of contact with varieties of pidgins. According to Shnukal, the majority of the men in these

---

1 This paper is dedicated to the memory of Pastor Bill Camden who laboured for many years along with the other consecrated members of the Kokonas Project to translate the Bible into Bislama. Pastor Bill's contribution to the advancement of Bislama, however, was far broader than producing the Bible. His 1977 work *A descriptive dictionary: Bislama to English* is an example of this contribution. For their labours and enthusiasm for Vanuatu and the Church, I pay tribute to Pastor Bill who died in December 1994 and to his wife Sue who survives him.

I want to thank Gabriel Lovanitili and Erastus Otairobo for supplying SP data which substantially helped in fleshing out the examples and also John Lynch for helpful comments on an earlier draft of this paper.

Jan Tent and France Mugler, eds *SICOL: proceedings of the Second International Conference on Oceanic Linguistics:* vol.1, *Language Contact*, 69–83.
Pacific Linguistics, C–141, 1998.
© Ernest W. Lee

Lee, E.W. "Unexpected shared features in Melanesian pidgins/creoles: is Broken a Melanesian pidgin/creole?". In Tent, J. and Mugler, F. editors, *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.*
C-141:69-83. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.69*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. **A sealang.net/CRCL** initiative.

areas were Pacific Islanders, most of whom had local wives. Shnukal does not discuss specifically the influence of Queensland Pidgin or Early Melanesian Pidgin on BR.

This paper raises the question as to whether BR is a Melanesian pidgin/creole. Superficially, it looks more different from TP, SP and BI than any of the latter look from each other. Keesing (1988:8) clearly includes Broken as a Melanesian pidgin in his introduction in the statement "...other [besides Tok Pisin] Melanesian Pidgin dialects—Bislama, Solomons Pidgin, Torres Strait Creole...". Harris (1986:36) accepts Early Melanesian Pidgin as a precursor to Torres Strait Creole (1986:7, 36, 292). And Clark's (1979:48) tree diagram shows a closer connection between Cape York Creole (my BR) and New Hebrides Pidgin (my BI) and Solomon Islands Pidgin (my SP) than any of these with New Guinea Pidgin (TP). Others, however, generally consider TP, BI and SP as comprising present-day Melanesian Pidgin (e.g. Crowley 1992).

Although the evidence presented in this paper does not alone demonstrate that BR should be included as a Melanesian pidgin/creole, the large number of shared but unexpected features of the commonly accepted present-day Melanesian pidgins (MPE), namely, TP, SP and BI with BR suggest that BR does have much to support it as a Melanesian pidgin/creole. For the balance of this paper I include BR as a variety of MPE, but at the same time recognise that many of the shared features predate Early Melanesian Pidgin.

By unexpected shared features, I mean features which one would not expect on the basis of ENG as the source language. Of the shared features considered below, most are probably because of a common core vocabulary, but some may come through diffusion with contact or through substrate influence, and still others may be independent or parallel innovations. Some, especially some of the phonological features, are too unusual to be likely the result of localised substrate influence or independent innovations such as the loss of 'v' in BR *manggru*, TP *manggro*, and SP *mangguru* 'mangrove'. I conclude that most of the various sets reflect older (nineteenth or early twentieth century) sharings rather than recently developed ones.

The comparisons made in this paper are based mostly on dictionary sources and my own knowledge of SP, which primarily represent the varieties as spoken today although with some indication of older usages which have been or are being lost. These comparisons focus on shared phonological, morphological and semantic features not expected from ENG source words. This does not mean that this approach is superior to other approaches. Working from all nineteenth and early twentieth century written sources, as Baker (1995) has done, has considerable merit. There is, however, a place for working from the present as well, which at least Crowley (e.g. 1992) has already demonstrated. Many currently shared uses never show up in documents since domains of the written sources only include a part of the full range of the life experience of speakers.

Clark's approach (1987)—of looking at shared lexical items which one would reasonably expect the Melanesians of plantation areas to have known before the mid-1880s and which TP, SP and BI share—is also valid.[2] This has support from my approach since the large number of unexpected shared features would indicate that there were also numerous shared features not reflecting unexpected differences from ENG.

Further dialect investigation and recording all available uses will fill in some of the gaps in the materials presented here as well as provide evidence for many more sets of words. Tryon

2 I have not seen Clark's paper. The information here is based on Crowley (1989:87).

(1991) has rightly pointed out that many regional conservatisms lurking in the various dialects need to be utilised. In addition, as Crowley (1989:87) has noted, there are many loans into vernaculars which attest the existence of words at an earlier stage.

Section 2 focuses on unexpected phonological considerations, that is, when two or more of the languages reflected a phonological form that would not have been expected from the ENG source language. Crowley (1991:6), for example, has pointed out the unexpected vowel *a* reflected in the forms from ENG 'help' in MPE (see set (7) below), but which is not in the printed records available. It is in the area of phonology that the historical documents give us the least evidence. An occasional deliberate non-English spelling will give cues as will materials written by someone trying to write pidgin who does not approach it as a speaker of ENG (see Keesing's (1988:101) description of Bislama written by the French Pionnier). Sections 3 and 4 focus on shared morphological items and semantic features which are different from ENG. In this presentation *most* of the sets in each subsection below are *not* exhaustive. Only the more unusual ones are. (A detailed list of all sets will appear in a fuller version of this paper.)

The slant of the paper is clearly towards BR since the purpose is to highlight the shared features of BR with one or more of the other varieties of MPE. Consequently all sets of words are listed with the BR form first, if there is one. In some sections, I also list separately one or more sets in which BR does not share the feature. Sets are alphabetised within each group according to the first form in each set.

Items in BR which are listed in Shnukal as being old, I have so indicated. Words in other varieties which are obsolete, archaic or restricted, I have attempted to so indicate. Where there are otherwise competing forms, if I know which is more frequent, I have listed it first. I have attempted to copy words accurately, but occasionally my familiarity with SP may cause me to incorporate SP spellings into other varieties.

A word on the role of BI and SP. Although BI and SP have many significant differences, their history and current closeness indicate that they are a subgroup of MPE. I have not listed sets of words where only BI and SP share the feature in question.

Finally, in the sets of reflexes I have listed the form for a specific word if the form is cognate with the others even though it may not exhibit the feature under consideration. To make it apparent which forms do exhibit the feature, however, I have underlined them.

2. PHONOLOGICAL CONSIDERATIONS

In this section, I treat some of the phonological features I have found shared by Broken with one or more of the other MPE varieties and which are different from what would be expected from the ENG sources.

2.1 VOWELS

All of the varieties have five simple vowels. For the most part the pattern of reflexes of vowels in the four varieties is that ENG high front vowels become /i/; mid front vowels become /e/; low front vowel and mid and low central vowels become /a/; high back vowels become /u/; and the non-high back vowels become /o/. Most variations from this are

considered to be unexpected. (The fuller version of the paper includes shared diphthongisation of *æ*, loss of unstressed initial schwa, and other minor shared features.)

2.1.1 /ɪ/ > /e/

A number of words in each of the varieties of MPE have a mid front vowel /e/ where the ENG has the higher vowel /ɪ/ and one would expect /i/ in the reflexes. This is more common in BR than any of the others.[3] Where there are competing vowels, the lower one is probably the older one.

	Broken	Tok Pisin	S I Pijin	Bislama	English
(1)	*em*	*em*	*hem*	*hem*	he, she, it, etc.
(2)	*liba/leba*	*lewa*	*liva/leva* (arch.)	*leva*	liver
(3)	*melk*	*melek*	*melek/milk*	*melek/milk*	milk (of coconut), milk (except TP)

2.1.2 /ɛ/ > /a/ BEFORE ALVEOLAR CONSONANTS

For most words of ENG origin, /ɛ/ is reflected as a mid front vowel, but in a good number of words where the vowel is followed by an alveolar consonant, the vowel is lowered to /a/.

	Broken	Tok Pisin	S I Pijin	Bislama	English
(4)	*plente/plande* (arch.)	*planti/plenti*	*plande/plenti*	*plante*	plenty
(5)	*sele/sale*	*salim*	*salem*	*salem*	to sell
(6)	*yalo*	*yelo*	*ialo*	*yala/yelo*	yellow[4]

Not shared by BR:

(7)	*elpe*	*halivim*	*helpem/halavem*	*halpem/helpem*	help

2.1.3 NON-HIGH BACK VOWEL > /a/

	Broken	Tok Pisin	S I Pijin	Bislama	English
(8)	*antap*	*antap*	*antap/on top*	*antap*	on top, above
(9)	*stap*	*stap*	*stap*	*stap*	to stay, remain < stop
(10)	*wata*	*wara*	*wata*	*wota/wora*	water

It is possible that these may bear American ENG influence where /a/ would be the norm although both *on* and *water* have variants with low back vowel.

---

3 There are also instances in BR where /e/ reflects the /i/ of Meriam Mir, such as *merer* 'pandanus leaf' from *mirir* (Shnukal 1988:135).

4 Concerning (6), there is a ditty in the United States with 'green and yaller'. I am inclined, however, to think that (6) reflects *yellow* rather than *yaller*, which is much less common.

2.1.4 REFLEXES OF ENG STRESSED *ir, ur*

This sequence, which is manifested in most dialects of American ENG as a syllabic *r* and which is written as *ɘr* in my Collins dictionary (1981) of British ENG, is variously reflected in BR and the varieties of MPE. In BR the most common reflex is /a/. In the other languages there seems to be more variation of /a/ with /e/ and /o/, but it would appear that /a/ was most common at earlier stages. Some examples with consistent use of /a/ in the four varieties of MPE are:

	Broken	Tok Pisin	S I Pijin	Bislama	English
(11)	*pastaim*	*pastaim*	*fastaem/festaem*	*fastaem*	first
(12)	*tane*	*tanim*	*tanem*	*tanem*	to turn

Other sets which are consistent and presumably old are /oi/ (either spelled *oi* or *oe*) in reflexes of *burn*, the /e/ in reflexes of *girl*, and the /o/ in reflexes of *church.*

	Broken	Tok Pisin	S I Pijin	Bislama	English
(13)	*ban/boin*	*boinem*	*bon(em)/ boen(em)* (arch.)	*bon(em)/ boen(em)*	to burn, singe, scorch (only transitive listed for TP)
(14)	*gel*	*gel*	*gele/gel*	*gel*	girl
(15)	*sos*	*sios*	*sios*	*jyos/joj*	church (TP recent in many areas (so Mihalic 1971:175))

Set (13) is very suggestive of American influence since *burn* is similarly pronounced in some dialects. The significant point here, however, is that it shows up in all of the varieties including BR.

2.1.5 LOSS OF FINAL VOWEL OF CLUSTER

	Broken	Tok Pisin	S I Pijin	Bislama	English
(16)	*amas*	*hamas/haumas*	*haomas/hamas*	*hamas*	how much
(17)	*kwait*		*kuaet*	*kwaet*	quiet

Not shared by BR:

(18)	*poizen*	*posin/poisin*	*poesen/posen*	*posen*	sorcery < poison

In light of the complex vowel clusters in many of the substrate languages, it seems significant that the three varieties with reflexes of *quiet* (17) drop the last vowel. The expected reflex of this would have been BR **aia* or BI and SP **aea.* Most Solomon Islanders speaking ENG treat *quiet* and *quite* as homophones.

2.1.6 SCHWA OF 'CANOE'

	Broken	Tok Pisin	S I Pijin	Bislama	English
(19)	*kenu*	*kanu*	*kanu/kiniu*	*kinu/kenu/kanu*	canoe

Although TP, SP and BI all have variants with /a/ in the first syllable, all the varieties except TP also have a variant with either /e/ or /i/ or both (BI). The /a/ is what I would have expected.

2.2 CONSONANTS

The normally expected changes in consonants include devoicing (especially in word final position), zero reflection of initial *h* in areas where there is no *h*, substitution of stops for fricatives, change of affricates to simple sibilants, addition of vowels to final consonants (especially in SP), deletion of last consonant of final consonant clusters. (Additional shared features in the fuller version of the paper include medial cluster reduction and voicing of *t* and *k* following nasal consonants.)

2.2.1 INTERVOCALIC /d/, /t/ /ð/ > /ɾ/

	Broken	Tok Pisin	S I Pijin	Bislama	English
(20)	*gerap*	*kirap*	*girap*	*girap*	wake up, get up < get up
(21)	*sarap/sadap*	*sarap*	*sarap*	*sarap/satap*	shut up!
Not shared by BR:					
(22)	*nada(-)*	*nara(-)/ara(-)*	*nara(-)*	*nara(-)*	another, different
(23)	*Satade*	*Sarere*	*Sarere/ Satade*	*Sarere/Sarede/ Satede*	Saturday

There are also sporadic reflexes of /d/ and /t/ as /ɾ/ such as BI, SP *bari* 'body' and BI *blari* 'bloody', *griri* 'greedy' and *parel* 'paddle'. These may reflect wider usages in the past which have been lost.

One wonders about the source of /ɾ/ for /t/, /d/ and /ð/ (set (22) is the only example reflecting ENG /ð/ and it is assumed that /ð/ became /d/ and then /d/ > /ɾ/). It was suggested to me that the /ɾ/ could be the influence of Australian aboriginal languages, but that doesn't answer the question of why the sounds would be reanalysed as /ɾ/ in the first place when both /t/ and /d/ were available. The other possibility that presents itself is the influence of American speakers in trading in the South Pacific. The most common American pronunciation of intervocalic /t/ and /d/ is a flap and is phonetically closer, for example, to SP /ɾ/ than to SP /d/ or /t/. I conclude that the forms with /ɾ/ are the older forms but the American connection needs further investigation.

2.2.2 /-tl-/, /-dl-/ > /l/

	Broken	Tok Pisin	S I Pijin	Bislama	English
(24)	*lelebet*	*lelebet*	*lelebet*		small amount, a little bit
(25)	*neil/nil*	*nil*	*nila*	*nil*	nail < needle (see below)

Not shared by BR:

(26)	*nidel*	*nil*	*nila*	*nidil*	needle

It looks likely that there was an early confusion between *nail* (25) and *needle* (26) since none of the forms for *nail* except the more common one for BR (*neil*) have the expected vowel /e/ (or /ei/). If, however, they all derive from *needle*, it would account for the vowel and the loss of /d/. BR *nidel* and BI *nidil* are presumably of more recent origin following the current pattern for words derived from ENG.

2.2.3 BILABIAL CONSONANT PLUS VOWEL—AN UNUSUAL CHANGE

In the two following examples, /v/ and /w/ plus the following vowel have unusual developments, but the pattern is clear (BR and SP >/u/, TP>/i/, BI>/wi/) although TP shares only one of the words.

	Broken	Tok Pisin	S I Pijin	Bislama	English
(27)	*kanbis/kanus* (arch.)	*kanis*	*kanus*	*kanwis*	canvas (see Crowley 1992:6)
(28)	*langus*		*langguis/langgus/ lanus* (restricted)	*lanwis*	vernacular language

In set (28), the SP *langguis* is the newer form. BI *lanwis* could also be recent since /i/ is the normally expected vowel from ENG 'language' /længgwɪj/, but it does follow the same pattern as *kanwis*.

2.2.4 PALATALISATION

	Broken	Tok Pisin	S I Pijin	Bislama	English
(29)	*gyaman/gyeman /geman*	*giaman*	*giaman*	*gyaman/ giaman*	to lie, deceive < gammon
(30)	*tomyok*	*tamiok*	*tomeo*	*tamiok*	axe (usually small) < tomahawk

2.2.5 CONSONANT LOSS—FINAL

There is a tendency in the MPE varieties to drop the second consonant of a word-final cluster (e.g. *wes* 'west' in all of them). Although likely due to substrate influence, it could be early. The only clear loss of a single post-vocalic consonant across the board, however, is the final consonant /ŋ/ of *long* and *blong* (*bilong* in TP) in all varieties (sets (31) and (32)).

This could be an independent development in each of the four varieties since phonological reduction of function particles is a common feature of languages and one we would expect with creolisation. In all except BR, which has been creolised for four generations, the longer form is generally written. BR also has loss of the final /-m/ of the transitive suffix (see §3.1), has *olse* along with *olsem* 'like, as', and other final consonant losses in function words.

	Broken	Tok Pisin	S I Pijin	Bislama	English
(31)	*blo/blong*	*bilong/bilo*	*blong/blo*	*blong/blo*	possessive marker
(32)	*lo/long*	*long/lo*	*long/lo*	*long/lo*	generic preposition
(33)	*mangru*	*mangro*	*mangguru*		mangrove

The loss of final /v/ of *mangrove* (33) is undoubtedly connected with an early shared feature. According to the *American heritage dictionary*, *mangrove* ultimately derives from Portuguese *mangue* which is in turn from Taino (extinct West Indies language) *mangle* and the final part of the word is influenced by ENG *grove*. ENG, nonetheless, is clearly the source for the varieties here; the Portuguese form which has no /v/ would require an intrusive /r/.

3. MORPHOLOGICAL CONSIDERATIONS

The first three morphological features treated in this section are based on features of the four varieties themselves, but the last two are based on features of the source language, namely, the occasional reflexes of plural ENG forms for singular and of past tense forms of verbs rather than the basic form of the verb. (The fuller version also includes compounds and/or fixed phrases where an expected transitive suffix is missing and verbs which reflect ENG adjectives.)

3.1 DOUBLE TRANSITIVE AFFIXES

The transitive marker (whether simple transitive or causative) has been widely generalised in each of the four MPE varieties under consideration. The suffix in all four have some variant with an *-m*. Although the suffix *-em* occurs in BR, the normal reflex is reduced to simply *-e* or sometimes *-i*. Shnukal does not (as best I can determine) distinguish between the use of the normal reduced form and the full form, but it appears from the many examples in her dictionary that the full form *-em* is somewhat regularly used when there is no overt object following the verb as in (34b) and (34c), although I noted a few exceptions, such as (34d), which has neither *-m* nor an overt object. The transitive affix is in bold typeface in each example below.

(34)a. ***Kese*** *diswan ya!*
catch this here
Take hold of this one!

b. ***Kesem*** *pas!*
Catch first
Hold this for a moment!

c. *Wanem yu kesem?*
what you catch
What did you catch?

d. *Ai piget po gibi yu.*
I forget to give you
I forgot to give it to you.

There are, however, a good number of words in BR in which *-m-* occurs regularly before the compound verb suffixes *-ap* and *-aut*. There are a few words in which it occurs with other compound verb suffixes as *stomwei* 'to hide, stowaway'. The compound suffixes *-ap* and *-aut* (and only these two) are always followed by a second transitive suffix as *klinmape* (*klin-m-ap-e*) 'to clean up, out' and *klinmaute* (*klin-m-aut-e*) 'to clean out'. It is these words that are of interest here since both TP and SP have words which also have the transitive marker occurring twice with similar compound verbs. It is exactly the same two affixes in TP and SP which can have two transitive affixes as in BR. The only difference is that in BR both are required **if** the main part of the verb has *-m-* whereas in TP and SP they are less stable, especially with *-aut* (TP) and *-aot* (SP). There is no double use of the transitive suffix with any other compound verbs such as with *daun/daon*, *wel/wei*, *raon* in any of the three.[5] BI does not use double transitive affixes with any compound verbs. Double underlining is used in the sets below only for reflexes with two transitive suffixes. The specific forms with two transitive affixes are not necessarily all old, but the pattern appears to be old since it is shared by three varieties of MPE.

	Broken	Tok Pisin	S I Pijin	Bislama	English
(35)	*hange*	*hangamapim* (transitive)	*hangemap/ hangemapem*	*hangemap*	to hang something
(36)	*lekmape*	*litimap/litimapim/ liptimapim*	*liftimap/ liftimapem*	*leftemap*	to lift up
(37)	*pulmape*	*pulimap/ pulimapim*	*fulumap/ fulumapem*	*fulumap*	to fill
(38)	*sermaute*		*searaotem/ seremaot*	*seraotem/ searemaotem*	to share out
(39)	*somape*	*samapem*	*somap/somapem soingemap/ soingemapem*	*somap*	to sew
(40)		*telimautim*	*talemaot/ talemaotem*	*talemaot*	to reveal
(41)	*tekmaute*	*tekimaut/ tekimautim*	*tekemaot/tekaotem/ tekemaotem*	*tekemaot*	to take out

Note the unusual consonant reduction of ENG *ft* in set (36). TP has variants which simply drop the *f* and one variant changes it to a stop, but the combination has been changed to *k* in BR. Also note that some of the TP forms do not have the expected vowel before the first

5 Mihalic (1971:24) says there is such a use with *ewe* but gives no examples. One likely exception is TP *tromoi/tromoim* 'throw, throw away' which is apparently derived from 'throw' plus 'away'.

transitive affix of the word as, for example, *hangamapim* 'to hang up' (35). It is also interesting that the TP word *hangemap* (not listed above) with only the medial transitive affix is in fact intransitive 'to be hanging'. These last two points both support the longstanding occurrence of the medial suffix.

3.2 REDUPLICATION

Reduplication occurs in all of four varieties of MPE, although it is extremely rare in ENG. The uses of reduplication tend to reflect the substrate languages, but there are some patterns which are shared by two or more of the MPE varieties. All except BR share a reduplication of intransitive verbs, as opposed to transitive verbs, which are not reduplicated, and BR has one possible example. SP and BI share several animal names which are reduplicated such as *dakdak* 'duck' and TP also shares *sipsip* 'sheep'. These languages also share a few excretory words which are reduplicated. Each of these are discussed in the fuller version of this paper. The one pattern which BR clearly shares with the other varieties is the reduplication of adjectives.

There is a pattern of reduplication attested in BR (apparently primarily in the East) which is not a pattern of reduplication in Meriam Mir, the substrate language. The only morphological process for which Meriam Mir uses reduplication is to derive adjectives from nouns with the resulting meaning of 'having [+noun]' (Shnukal 1988.22). The single pattern which survives in BR is not an active process and is probably quite old. Of the ones listed by Shnukal (she lists only six which is about a third of them) BI and SP have cognate reduplicated forms for two and TP for one, although TP *liklik* is semantically related and so also included.

	Broken	Tok Pisin	S I Pijin	Bislama	English
(42)	kalakala	*kalakala*[6]	*kalakala*	*kalakala*	variegated, multicoloured
(43)	*smolsmol*	*liklik*[7]	*si-smol*	*smosmol/smolsmol*	small (frequently distributive)

All of the languages have unreduplicated reflexes of *small.* SP *si-smol* has only reduplication of the initial consonant plus an epenthetic vowel.

3.3 PRONOMINAL SYSTEM

The pronominal systems of TP, SP and BI have only very minor differences so only the BR and TP systems are charted here. The related forms of BR and TP are underlined.

---

6 Mihalic (1971:102) does not make a separate entry for *kala kala* 'many-coloured', but his only illustrative sentence for *kala* 'colour' contains it.

7 One wonders if it is possible that TP formerly had **smolsmol* which *liklik* replaced. The vernacular source word given for *liklik* in the dictionary is not reduplicated. The unreduplicated form *lik* in TP means 'leak'.

	Person	Singular	Dual	Trial	Plural
BR:	1EXC	*ai* (recent as subject)/*mi*	*mitu*	*miplatri*	*mipla*
	1INC	----	*yumi/yumitu*	*yumplatri/yumitri*	*yumpla*
	2	*yu*	*yutu*	(none listed)	*yupla*
	3	*em*	*demtu*	(none listed)	*dempla/ol* (subj.)/*em* (obj.)
TP:	1EXC	*mi*	*mitupela*	*mitripela*	*mipela*
	1INC	----	*yumitupela*	*yumitripela*	*yumi*
	2	*yu*	*yutupela*	*yutripela*	*yupela*
	3	*em/en* (after *long*, *bilong*)	*tupela*	*tripela*	*ol*

Although the pronominal systems of the other MPE varieties are closer to each other than BR is to any of them, the similarities of BR with them are striking. All have the non-ENG categories of dual and trial as well as contrast of first person inclusive and exclusive manifested in significantly parallel ways. The only missing forms in BR are part of the trial forms, but the occurrence of some in the dictionary may indicate that the missing forms were overlooked. On the other hand, the dictionary also lists inclusive and exclusive forms for *the four of us* (SP also has occasional forms referring to four). Differences in BR are:

(a) the 'recent' first singular subject form *ai*;

(b) an 'archaic' *wi* for all non-singular first person forms;

(c) the 'lack' of *-pla* in the dual (BI, however, also has optional *yumitu* for first person dual inclusive);

(d) the use of *yumi* for the first person dual inclusive when the other three use the unaffixed form *yumi* (TP, BI) or *iumi* (SP) for the first person plural inclusive (BR uses *yumpla*);

(e) *dem-* in the non-singular third person forms *demtu* and *dempla*.

BI and SP differ from both BR and TP in that the third singular is *hem* and the third plural is BI *olgeta* and SP *olketa/olgeta*.

3.4 REFLEXES OF PLURAL FOR SINGULAR

Crowley (1992:9) has already pointed out the occurrence of shared reflexes of ENG plural forms for the singular. These are mostly small things or things often thought of collectively so the plural may be more common in ENG. The point here, however, is that all the varieties reflect only the plural for most of these words.

	Broken	Tok Pisin	S I Pijin	Bislama	English
(44)	*anis*	*anis*	*anis*	*anis*	ant
(45)	*masis*	*masis*	*masis*	*mases*	match
(46)	*tit*	*tit*	*tit*	*tut*	tool

Not shared by BR:

(47)	*hasis*	*has/hasis*[8]	*hasis/ases*	hatch (nautical)

3.5 REFLEXES OF PAST TENSE OF VERB

Normally verbs are reflexes of the simple stem of the verb in ENG. A few, however, reflect the past tense of the verb.

	Broken	Tok Pisin	S I Pijin	Bislama	English
(48)	*brok*	*bruk/brukim*	*brek/brok* (arch.)	*brok*	broken, to break
(49)	*kapsaid/ kapsaiz*	*kapsait*	*kapsaet*	*kapsaet/ kafsaet*	turn over, spill

4. SEMANTIC CONSIDERATIONS

In this last section, I list a sampling of the sets of reflexes which are semantically significantly different from ENG or where a less common word from ENG is reflected rather than a more common one. Some of these are quite significant such as the use of *bagarap* in all four varieties for 'damaged' or 'ruined', with no vulgar connotation (see set (53)). The same, however, is true of my variety of American English. I use it in the same way as Solomon Islanders and was an adult long before I learned that it also had a vulgar connotation. Apart from *bagarap*, no items which Baker (1995) treated are included below with the same meaning he gives. (Most of his list is included in the fuller version as are various items from Crowley.)

For each set below it is the unexpected shared meaning that is specified in the ENG gloss. If a word in one of the languages comes from the same ENG source but without the specified meaning or very close to it, it is listed but not underlined. In a few cases the ENG source may also have the specified meaning in certain contexts but restricted so that it is not what would be expected in MPE.

	Broken	Tok Pisin	S I Pijin	Bislama	English
(50)	*ambag*	*hambak*	*hambag*	*hambag*	be a nuisance (except SP), illicit sex (except TP) < humbug
(51)	*ap*	*hap*	*haf*	*haf*	part, piece of, half
(52)	*bagarap*	bagarap	*bagarap*	*bagarap*	damaged, ruined < bugger up
(53)	*brase*		*brasim*	*brasem*	to cut grass or undergrowth
(54)	*broke* (skin)		*brekem*	*brokem/ brekem*	have sexual intercourse (BR), deflower (SP, BI) < break

8 I was told by a sailor that they use *has* or 'hatch' and that *hasis* is used by landlubbers.

(55)	*dabelskin*		*dabolskin*	*dabolskin*	scabies
(56)	*ded*	*dai*	*dae*	*ded*	numb
(57)	*gerap*	*kirap*	*girap*	*girap*	wake up (and all except BR sexually aroused) < get up
(58)	*ilpis*		*ilfis*	*ilfis*	eel < eel + fish
(59)	*kapa*	*kapa*	*kapa*	*kapa*	corrugated roofing iron < copper
(60)	*kapsaiz/ kapsaid*	*kapsaitim*	*kapsaet*	*kapsaet/ kafsaet*	ejaculate[9], urinate (SP, BI), pour, spill (all)
(61)	*langus*		*langgus/ langguis/ lanus*	*lanwis*	vernacular (excludes ENG, pidgins/creoles)
(62)	*luz*	*lus*	*lus*	*lus*	to die (euphemism)
(63)	*nadakain*	*narakain*	*narakaen*	*narakaen*	unusual (BI: implied by one of Camden's examples)
(64)	*nekstumora*		*nekes tumoro*		day after tomorrow
(65)	*nogud sik*	*sik nogut*	*sik nogud*	*sik nogud*	venereal disease
(66)	*poret*	*poret*	*fored*	*fored*	prow, bow or ship < forehead
(67)	*samting*	*samting*	*samting*	*samting*	thing, euph. for genitals
(68)	*sem*	*sem*	*sem*	*sem*	embarrassed, disgraced (see Crowley 1992:6) < shame
(69)	*smolaus*	*smolhaus*	*smolhaos*	*smolhaos*	outhouse, toilet
(70)	*stiya*	*stia*	*stia*	*stia*	rudder, helm < steer
(71)	*storiyan*		*stori/ storian* (arch.)	*storian/ storeyan*	to tell a story < story yarn
(72)	*Tru God!*		*Tru God!*	*Tru God!*	So help me God! (oath)
(73)	*yumi* (dual)	*yumitupela*	*iumitufala*	*yumitufala*	we (dual inclusive) used as a plural in public speaking

Not shared by BR:

(74)		*ating*	*ating*	*ating*	perhaps < I think
(75)		*bikhet*	*bikhed*	*bighed*	stubborn

9 Mihalic (1971:106) also lists the transitive in: *Mi iet mi kapsaitim wara/strong/blut bilong mi* which he glosses as 'I pollute myself.' I assume this is intended to include a sexual connotation.

(76)	*bigmaut*	*bikmaus*	*bikmaos*	*bigmaot*	to yell (BR 'talkative')
(77)		*burumbut* (obs.)	*purubut*	*purumbut/ purput/ frumbut*	to tread on < put your boot (TP from Crowley 1992:7)
(78)	*gris*	*gris*	*gris*	*gris*	body fat
(79)	*melk*	*melek*	*melek/milk*	*melek*	semen
(80)	*poizen*	*posin/ poisin*	*poesenem/ posenem*	*posenem*	to put a curse on < poison
(81)	*tane*	*tanim*	*tanem*	*tanem*	to translate < turn
(82)		*traut*	*toraot*	*traot*	vomit < throw out (not up)
(83)	*tru antap*	*tru antap*	*tru antap*		Honestly! (oath, affirmation)

Set (56): Although reflecting differing forms of ENG 'die', it would appear that *ded, dai* and *dae* 'numb' all share a common origin. There may at one time have been variants in some stage of earlier MPE.

Set (72) with set (83): The phrase *tru antap* (83) glossed as 'Honestly!' derives from *true on top.* As Crowley (1989:105) has rightly noted, this is a euphemism or avoidance term for *tru God* (72) which for some people at least is blasphemous. *Antap* is itself a euphemism for *heaven*, and to the Jewish people of New Testament times *heaven* was a euphemism for *God.* BR has only *tru God* and TP has only *tru antap.*

Set (73): The use of the dual inclusive for plural inclusive is a striking example and is used to make personal contact between the speaker and the individuals in the audience. It occurs frequently in preaching and is shared by all (BI source is from ni-Vanuatu students at Bishop Patteson Theological College) except Tok Pisin in the data available. Although the form in BR differs from BI and SP, it is nonetheless dual and the function is identical (see §3.3).

5. CONCLUSION

Evidence here does not demonstrate conclusively that BR is a Melanesian creole. It shows a lot of sharing, much of which is probably at least from Early Melanesian Pidgin sources and some of which is from Queensland Pidgin or other earlier sources. Much of the morphological and semantic considerations may be already traced or traceable in historical documents, but it is unlikely that we will find significant evidence of the phonological considerations apart from what has been recently recorded in a sound or phoneme based orthography or which is extant and not yet recorded. The phonological sharings as they stand suggest a strong affinity between Broken and the other MPE varieties. Further work needs to be done before we can state whether BR is closer to one of the MPE varieties than another, or whether in some ways it has closer affinity to one of them than they have to each other (apart from BI and SP). Detailed dialect studies and evidence from loans into vernaculars would be great potential sources of further evidence.

REFERENCES

The American heritage electronic dictionary, 1992. Boston: Houghton Mifflin. (Based on *The American heritage dictionary of the English language*, 3rd edn)

Baker, Philip, ed. 1995, *From contact to creole and beyond.* London: University of Westminster Press.

Camden, Pastor Bill, 1977, *A descriptive dictionary: Bislama to English.* Port Vila: Maropa Bookshop.

Clark, Ross, 1979, In search of Beach-la-mar: towards a history of Pacific Pidgin English. *Te Reo* 22:3–64.

1987, The lexicon of Early Melanesian Pidgin. Paper presented at the Seventh New Zealand Linguistics Conference.

Collins pocket dictionary of the English language, 1981. Wrotham, England: William Collins.

Crowley, Terry, 1989, Referential and expressive expansion in Bislama. *English World-Wide* 10/1:85–118.

1990, *An illustrated Bislama-English and English-Bislama dictionary.* Port Vila: Pacific Languages Unit and the Vanuatu Extension Centre of the University of the South Pacific.

1991, Varieties of Melanesian Pidgin: separate identities vs. European stereotypes. *Language and Linguistics in Melanesia* 22:51–65.

1992, Derivational morphology and structural complexity in nineteenth century Melanesian Pidgin. *Te Reo* 35:3–18.

Harris, John W., 1986, *Northern Territory Pidgins and the origin of Kriol. PL*, C–89.

Keesing, Roger M., 1988, *Melanesian Pidgin and the Oceanic substrate.* Stanford, California: Stanford University Press.

The Macquarie dictionary (2nd edn), 1991. Sydney: Macquarie Library.

Mihalic, Francis, 1971, *The Jacaranda dictionary and grammar of Melanesian Pidgin.* Milton, Queensland: Jacaranda Press.

Shnukal, Anna, 1988, *Broken: an introduction to the creole language of Torres Strait. PL*, C–107.

Simons, Linda and Hugh Young, 1978, *Pijin blong yumi: a guide to Solomon Islands Pijin.* Honiara: Solomon Islands Christian Association.

Tryon, Darrell, 1991, Wanem Bislama? In Robert Blust, ed. *Current trends in Pacific linguistics: papers on Austronesian languages and ethnolinguistics in honour of George W. Grace*, 509–519. *PL*, C–117.

Lee, E.W. "Unexpected shared features in Melanesian pidgins/creoles: is Broken a Melanesian pidgin/creole?". In Tent, J. and Mugler, F. editors, *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.*
C-141:69-83. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.69*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.

COMPARING OLD AND NEW INFORMATION IN BISLAMA: NOMINAL DELETION WITH *OLSEM*

MIRIAM MEYERHOFF

This paper considers the strategies available in Bislama for comparing two NPs using the preposition *olsem*, which means 'like, as'.[1] I will show that Bislama allows two structures, one with *olsem* preceding a noun and one with *olsem* following the noun. However, I will suggest that even though this looks like freedom of movement in the placement of *olsem*, *olsem* in fact consistently precedes nouns. However, whether the nominal argument of a prepositional phrase headed by *olsem* surfaces overtly is constrained by the information status of the two nominals being compared.

For the purposes of this paper I will make my focus quite narrow and consider the use of *olsem* in only one kind of comparison, that is, similes or comparisons where two nominals are being equated or instantiated (e.g. a light like a diamond; teams like the All Blacks; Clark, as leader of the Labour Party; etc.). I will not be considering other uses of *olsem* in Bislama, for instance, it is also used to compare verb phrases, and to introduce reported or direct speech, nor will I discuss other kinds of comparison in Bislam (e.g. comparisons of greater than or less than relationships).

The data used in this investigation are equative comparisons using *olsem* found in written Bislama. Most of the examples discussed in this paper are taken from an exhaustive sample of the equative comparisons in the Bislama language pages of the national newspaper, the *Vanuatu Weekly/Hebdomadaire*, over two months in 1994–95. I have retained the original spellings and punctuation used in the *VWH* even though this results in some inconsistencies between examples and means that the examples in this paper do not correspond to the standards recently established in Vanuatu for written Bislama. However, in my discussion of examples I have followed the new standards. A written database of Bislama proved to be the best environment for gathering tokens of *olsem* in equative comparisons, since this use of *olsem* turns out to be proportionately more common in written Bislama than it is in spoken Bislama. In spoken Bislama, *olsem* serves a range of different (though semantically related) syntactic and discourse functions. These functions are outlined in Meyerhoff and Niedzielski (1995).

1 I would like to thank the Wenner-Gren Foundation (grant #5742) for assisting me to conduct this research, also the Vanuatu Cultural Centre. I would also like to thank Sharon Morrie Tabi whose judgements in Bislama aided me greatly in this analysis of *olsem*. Also, members of the audience at the Second Oceanic Linguistics Conference, Suva, especially John Lynch and Robin Hooper.

Jan Tent and France Mugler, eds *SICOL: proceedings of the Second International Conference on Oceanic Linguistics:* vol.1, *Language Contact*, 85–93.
Pacific Linguistics, C–141, 1998.
© Miriam Meyerhoff

Meyerhoff, M. "Comparing old and new information in Bislama: nominal deletion with olsem". In Tent, J. and Mugler, F. editors, *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.* **C-141:85-93.** Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.85*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. **A sealang.net/CRCL** initiative.

The variation we are concerned with takes the following form. The Bislama preposition *olsem* occurs both before and after nouns, as shown in (1) and (2):[2]

NP + *olsem* + NP

(1)a. *I gat* ***plante fres frut olsem popo.***
PRED have plenty fresh fruit like papaya
There are **plenty of fresh fruit like papaya**.

b. *Hemi faenem i had tu blong winim* ***wan***
3SG find PRED hard too to beat a

tim olsem St Louis.
team like St Louis
They also found it hard to beat **a team like St Louis**.

c. *Hemi talem se ino stret blong singaotem* ***trip***
3SG tell that PRED.NEG straight to call trip

blong wan gavman minista olsem turis *from...*
of a government minister like tourism because
He said it's not fair to call **a minister's trip 'tourism'** because...

NP + *olsem*

(2)a. *...from taem hemi smol pupu hemi mekem*
because time 3SG.PRED small grandson 3SG.PRED make

wan nakamal olsem.
a *nakamal* like
...because when he was a small boy he built **a similar *nakamal*** (i.e. ceremonial house).

b. *Ino gud blong stap konfiusum pipol long ol*
PRED.NEG good to HAB confuse people to PL

kaen tingting olsem.
kind thoughts like
It's not right to confuse people with **those kinds of ideas**.

c. *Vanuatu i wan gud ples blong setemap* ***kaen***
Vanuatu PRED a good place to set.up kind

projek olsem.
project like
Vanuatu is a good place to set up **such projects**.

---

2 Abbreviations used in this paper are as follows:

ANT	anterior	IRR	irrealis
CONT	continuous	NEG	negation
DU	dual	PL	plural
EXC	exclusive	PRED	predicate marker
HAB	habitual	SG	singular
INC	inclusive	SPEC	specificity marker

The problem appears to be that *olsem* can be both a preposition and a postposition. It is the postposed forms which stand out most. Bislama is an SVO language: it has prepositions, adjectives usually precede nouns (e.g. *wan gudfela fren* 'a very good friend'), and tense/aspect marking precedes the main V (e.g. *mifala i stap wet* 'we are waiting', *hem i bin askem* 's/he asked'). There are a small number of notable exceptions to this generalisation of right branching. There is a postnominal determiner, *ya*, which marks specificity in Bislama. The placement of *ya* with respect to the NP contrasts with the placement of the other determiners, *ol* 'PL (definite)' and *sam* 'some' and *wan* 'a', all of which are prenominal. Historically, *ya* derives from an adverb (i.e. 'here'), and presumably this partly accounts for why *ya* holds the anomalous position of a postmodifier. In addition, *nogud* 'bad' is a postmodifier of nouns, and people's local affiliations are invariably expressed through N + modifier constructions (e.g. *man Malo* 'person from Malo; a Malo guy'). Thus, it is not a purely logical possibility that *olsem* might be a postmodifier in Bislama. There are limited, though salient, exceptions to the right-branching structure more general to Bislama syntax.

One possible explanation would be that we are dealing with two meanings of *olsem* and that these are systematically distinguished by their placement with respect to the noun. This does not seem to be the case here though, since both have essentially the same meaning and both are used to equate or identify through comparison.

However, closer inspection reveals that the nature of the comparison is not exactly the same in the structures exemplified in (1) and (2). With prenominal *olsem*, as in (1a-c), the second noun being compared is new or non-redundant information which adds to or helps develop a shared mental schema between the writer and reader. For example:

(3)a. *Ol driver i yusum* ***rod ia*** *olsem* ***wan mein strim.***
PL driver PRED use road SPEC like a main stream
The drivers use that road as a main thoroughfare.

b. ***Mi olsem wan ex-studen blong Lycee LAB,*** *mi wandem sherem tingting blong mi...*
I like a ex-student of Lycee LAB I want share thought of me
As a former Lycee LAB student, I'd like to share my thoughts...

c. *Taem yumi lukluk long rigen mo kolosap moa* ***ol brata kaontri blong yumi olsem Solomon o PNG...***
time we.INC look to region and close more PL brother country of we.INC like Solomons or PNG
When we consider the region and, even closer to home, our neighbouring countries like the Solomons or PNG...

In all such cases of NP + *olsem* + NP, the information in the second NP, which I will call the 'comparator', is novel and not necessarily predictable given the first NP, which I will call the 'benchmark'. What the benchmark is being equated or instantiated with is new information to the reader. By providing this new information, the comparator helps to establish a jointly constructed discourse universe of shared knowledge where there had previously been none. In other words, when we see prenominal uses of *olsem* the benchmark is identified or instantiated in new or non-obvious ways by the comparator.

It seems clear to native speakers of Bislama that with postposed *olsem* there is, in fact, an implied comparator but that in these cases the comparator is identical with or entirely predictable given the benchmark and/or the wider discourse. In these cases, what is being instantiated by the benchmark is already well-established or old information in the discourse. I would suggest that because this information is given already, it is perceived to be redundant and can therefore be deleted, as might be predicted from pragmatic principles of relevance (Sperber & Wilson 1986) or the "Given-New Contract" proposed by Clark and Havilland (1977). Both Relevance theory and the Given-New Contract essentially state that in orderly discourse, familiar or redundant information need not be spelt out in full if, for instance, an unambiguous referent can be computed from memory, unless the speaker wishes some other inference or conclusion to be drawn from the spelling out.

Thus, it follows from these principles that a perfectly straightforward NP + *olsem* + NP equative comparison may be transformed into the anomalous looking NP + *olsem*. If the need to instantiate the benchmark (i.e. the first NP) has been trivially satisfied by the discourse as a whole, it becomes unnecessary to spell out the comparator; the discourse context satisfies this need. If you ask native speakers of Bislama to fill out these constructions they can do so, but in doing so they always repeat the benchmark as the comparator, often modifying it slightly to explicitly indicate that the focus is on the benchmark. In other words, speakers treat (4a', b') as underlying (4a, b).

(4)a. *...from taem hemi smol pupu hemi*
because time 3SG.PRED small grandson 3SG.PRED

mekem wan nakamal olsem.
make a *nakamal* like
...because when he was a small boy he built a similar *nakamal.*

a'. *...from taem hemi smol pupu hemi*
because time 3SG.PRED small grandson 3SG.PRED

mekem wan nakamal olsem nakamal ya.
make a *nakamal* like *nakamal* SPEC
...because when he was a boy he built a *nakamal* like this *nakamal.*

b. [*Depo-Provera*] *hemi sef mo ol toktok blong*
[Depo-Provera] 3SG.PRED safe and PL talk of

rod olsem i no tru nating.
road like PRED NEG true nothing
D-P is safe and such rumours are completely unfounded.

b'. [*Depo-Provera*] *hemi sef mo ol toktok blong*
[Depo-Provera] 3SG.PRED safe and PL talk of

rod olsem tufala toktok ya i no tru nating.
road like two talk SPEC PRED NEG true nothing
D-P is safe and rumours like those two rumours are completely unfounded.

The article that (4a) is drawn from is about the opening of a *nakamal* on Pentecost which had been built by an old man using traditional construction methods. Thus, it is this larger discourse topic that determines that the comparator of *wan nakamal olsem*, if made explicit, is singular. Similarly, information provided by the immediately preceding discourse requires

that the comparator of *ol toktok blong rod olsem*, when made explicit, be plural. Two rumoured side effects are raised and addressed:

> *Long saed blong ol toktok long rod we sam mama oli bin stap talemaot abaot stik o method ia se hemi save mekem bodi i fatfat o samfala i save grow mustas, Misis Ronolea i dinaem ol toktok ia.*
>
> Mrs Ronolea denied the rumours that some mothers have been repeating concerning the injection or this method [i.e. Depo-Provera], namely that it can make you put on weight or some women can grow beards.

Other advantages to this analysis of *olsem* are:

(a) it is consistent with other structures in Bislama (which I will illustrate shortly) that are sensitive to whether the referent is hearer-new information or hearer-old, and

(b) it allows us to explain the interpretation of sentences like (5) and (6) and determine how they should best be represented underlyingly.

In (5) and (6) we find examples of what appear on the surface to be nominals postmodified first by *olsem* and then by some other constituent, a PP in (5) or CP in (6). However, again, it seems clear that contextually redundant information has been omitted and speakers of Bislama flesh these implicit comparisons out along the lines of (5a') and (6a').

(5)a. *Long ples ia 3 man i bin ded, 2 yia I*
in place SPEC 3 man PRED ANT dead 2 year PRED

pas nao folem ***ol kaen aksen olsem blong PNG Difens Fos.***
pass now follow PL kind action like of PNG Defence Force
Three people were killed here two years ago following similar activities by the PNG Defence Force.

a'. *...folem ol kaen aksen olsem ol aksen blong PNG*
follow PL kind action like PL action of PNG

Difens Fos tedei.
Defence Force today
...following similar activities to today's by the PNG Defence Force.

(6)a. *Hemia i bin longtaem blongwet long* ***sam help***
3SG.SPEC PRED ANT longtime of wait from some help

olsem we bae i kam long gavman.
like that IRR PRED come from government
It's been a long wait for such help that's come from the government.

a'. *...Blong wet long sam help olsem help ya we bae*
to wait from some help like help SPEC that IRR

baei kam long gavman.
PRED come from government
...to wait for help like this help that has come from the government.

One possible analysis of *olsem* in examples such as these might take the position that *olsem* is not, in these sentences, a preposition. Crowley's (1990:196, 258) discussion of *olsem* distinguishes between its use as a preposition and its use as an adverb, meaning 'thus' or 'so' (e.g. *mifala i wokem olsem* 'we made it thus/like that', or *huia i talem olsem?* 'who said so?'). However, it is easy to show that postposed *olsem* is not functioning as an adverb

in the equative comparisons in (5) and (6). This is ruled out by virtue of constraints on the placement of adverbs within NPs. When an adverb like *longwe* 'there' modifies a noun, it must follow a PP modifying the same noun, as shown in the elicited data in (7):

(7)a. *Yu lukim haos blong mi longwe?*
Have you seen my house over there?

b. **Yu lukim haos longwe blong mi?*

There appears to be more freedom in the placement of *olsem* and a relative clause modifying a noun, however, I interpret the grammaticality of (5), and two other examples of the sequence N + *olsem* + PP found in the corpus as showing that *olsem* in these sentences is not functioning as an adverb.

Perhaps of greater relevance to the argument that apparently postposed cases of *olsem* involve deletion of an argument, are cases where head nouns are modified by both *olsem* and the specificity marker *ya.* For example:

(8) *Fulap famili tedei oli fesem* ***problem olsem ia.***
many family today PRED face problem like SPEC
Lots of families today face problems like these [just cited].

(9) ***Kaen prodak olsem ia*** *hemi* ***tabu tumas blong karem***
kind product like SPEC 3SG.PRED forbidden very to bring
i kam long Vanuatu.
PRED come to Vanuatu
Products like these [under a photo and description] are completely prohibited from being brought into Vanuatu.

It is clear that in sentences like (8) and (9), *ya* is functioning in its well-established role in Bislama as the "demonstrative cum definiteness marker" (Crowley 1990:285). It is not the case, as described by Sankoff and Brown (1976) for Tok Pisin, that the meaning of *ya* has been bleached and that it is functioning simply to mark the rightmost boundary of a clause, or in this case a phrase.

However, placement of *ya* is also revealing of the discourse status of the N it modifies. *Ya* may occur immediately after the head N it modifies or another phrase can intervene between *ya* and the head N. But when *ya* occurs immediately adjacent to the head N, the interpretation is that the N is given information, cf. the following elicited examples:

(10)a. *Mi mas go long ofis longwe ya.*
I must go to office there SPEC
I had to go to the office over there.
(hearer may not know which office exactly)

b. *Mi mas go long ofis ya longwe.*
I had to go to that office over there.
(hearer knows exactly which office, it is given information in the discourse or it has been pointed out)

(11)a. *Long tiket we mi bin nidim ya, oli putum stam*
to ticket that I ANT need SPEC PRED put stamp

blong olgeta.
of 3PL
They put their stamp on the ticket that I needed.
(a specific ticket exists, but is not necessarily identifiable to hearer)

b. *Long tiket ya we mi bin nidim, oli putum stam*
to ticket SPEC that I ANT need PRED put stamp

blong olgeta.
of 3PL
They put their stamp on this (very) ticket, which I needed.
(speaker might be waving ticket)

The problem appears to be the following. Examples (10b) and (11b) show that a reading of maximally given or identifiable information is preferred when *ya* occurs immediately next to the N it modifies. But Bislama speakers' intuitions about strings of N + *olsem* + *ya*, are that the benchmark N, from which the *ya* is separated by *olsem*, is likewise identifiable discourse-old information. This suggests that even though it might be possible to analyse N + *olsem* + *ya* sequences as Ns followed by two postmodifiers, given the way the information status of the nominals is interpreted, this analysis is not the most appropriate.

Thus, I am claiming that the difference between NP + *olsem* + NP and NP + *olsem* constructions is the recoverability of the second element in the comparison string. An interesting question then becomes whether this pragmatic distinction has any structural reflexes. Are there consistent qualitative differences between the benchmark NPs in NP + *olsem* + NP constructions and the benchmark NPs in NP + *olsem* constructions? And are there consistent qualitative differences between the comparators in the two constructions?

The second question is simultaneously easy to answer and impossible to test. The account I have argued for here is built around my assertion that there is a qualitative difference between the form of the comparator in the two constructions. I noted that Bislama speakers interpret NP + *olsem* constructions as having an implied comparator that is always discourse-old, and showed that, when asked to, speakers consistently spell out the deleted comparator with a specific NP, usually of the form N + *ya*. Thus, I cannot compare the form or information status of the comparators in the two constructions, as I have analysed one set of comparators as being of invariant form.

However, we can compare the form of the **benchmark** nouns in the two constructions, and the results of this comparison for the 201 equative comparisons found in the *Vanuatu Weekly* are presented in the following table. Definite and specific benchmarks have been grouped together and these form a natural class with zero arguments, since Bislama appears to allow license zero arguments largely on pragmatic grounds, for example, when the referent is retrievable from prior utterances (though this remains to be systematically tested). Obviously, definiteness and specificity are determined by the discourse, and are not purely formal measures, thus the examples provided in this column (i.e. a noun followed by *ya* and *ol* with a head noun) are by no means exhaustive. A relative clause or a prepositional phrase could make a head noun specific, for instance.

TABLE: COMPARISON OF THE FORM OF THE **BENCHMARK** NOUNS IN THE CONSTRUCTIONS

NP + *olsem* AND NP + *olsem* + NP

		FORM AND REFERENTIAL PROPERTIES OF BENCHMARK			
	No. of tokens	definite/specific, e.g. N *ya*, *ol* N	zero	indefinite,e.g. *wan*, *sam* N	generic N or abstract N
NP + *olsem*	66	32% (21)	0% (0)	33% (22)	35% (23)
NP + *olsem* + NP	135	53% (71)	18% (25)	15% (20)	14% (19)
Total	201	46%	12%	21%	21%

We can see that the benchmark noun in NP + *olsem* constructions (where the comparator is discourse-old information) is more likely to be an indefinite, abstract or generic NP than it is in NP + *olsem* + NP comparisons. For example:

(12) *Hemi wan blong ol hae panismen we Kot i*
3SG.PRED a of PL high punishment that court PRED

givimlong wan rep keis olsem.
give to a rape case like
This is one of the heaviest punishments the court can give in a rape case like [this].

(13)a. *Be hemi pruvum se tede woman tu i save*
but 3SG.PRED prove that today woman too PRED can

holem top position olsem.
hold top position like
But this proves that today women too can hold senior positions like [this one].

b. *Situesen olsem ia nao yumi save talem se ol*
situation like SPEC now we.INC can tell that PL

pleia oli lusum konsentresen blong olgeta.
player PRED lose concentration of 3PL
In a situation like this [situation], we can say the players have lost their concentration.

Note that (13b), in which the comparator that has been deleted is something like *situesen ya* 'this situation', indicates that it is important to maintain a distinction between hearer-old information and discourse-old information (Walker & Prince forthcoming). Deletion of the comparator is possible in this example because it is given by the discourse (every event is situated in time and space), that is, it is discourse-old, not because there has been prior use of the phrase *situesen (ya)*, which is a requirement for something to be considered hearer-old.

In NP + *olsem* + NP comparisons (pace the examples given in (1)), the benchmark is more likely to be definite/specific than indefinite or generic, as shown in (14), and in this *olsem* construction, the benchmark is more likely to be definite/specific than the benchmark in NP + *olsem* constructions.

(14)a. *Tufala i askem olgeta kastom jif blong karem taem ia*
3DU PRED ask 3PL custom chief to take time SPEC

olsem wan taem blong fogetem ol rabis tingting...
like a time to forget PL rubbish thought
They asked the traditional chiefs to make this time a time for forgetting bad feelings...

b. *Mifala i stap lukaot hem olsem wan stret pikinini*
we.EXC PRED CONT care.for him like a real child

of *mifala.*
blong we.EXC
We're looking after him like he was one of our own children.

In conclusion, then, I have argued that what looks like freedom in the placement of *olsem* with regard to benchmark nouns in equative comparisons is, in fact, a regular pattern of nominal premodification. However, the presence or absence of the comparator is constrained by pragmatic factors, namely whether or not the comparator, which functions as an instantiation of the benchmark, is discourse-old and is therefore easily retrievable by the hearer/reader. I have discussed constraints on the interpretation and placement of *olsem* when it occurs with adverbs and the specificity marker *ya* and shown that two other possible analyses of NP + *olsem* constructions, namely one in which *olsem* is an adverb and one in which it is a nominal postmodifier similar to *ya* are ruled out on syntactic and pragmatic grounds respectively. In other words, use of *olsem* in Bislama seems to be another instance of syntactic variation that is best explained by the structure of the discourse and the information status of the referents being compared.

REFERENCES

Clark, Herbert and Susan Haviland, 1977, Comprehension and the Given-New Contract. In Roy O. Freedle, ed. *Discourse processes: advances in research and theory*, vol.1: *Discourse production and comprehension*, 1–40. Norwood, NJ: Ablex.

Crowley, Terry, 1990, *Beach-la-Mar to Bislama: the emergence of a national language in Vanuatu*. Oxford: Clarendon Press.

Meyerhoff, Miriam and Nancy Niedzielski, 1995, On the discourse functions of *olsem* in Bislama. Paper presented at NWAVE 24, Philadelphia, October 1995.

Sankoff, Gillian and Penelope Brown, 1976, The origins of syntax in discourse: a case study of Tok Pisin relatives. *Language* 52/3:631–666.

Sperber, Dan and Deirdre Wilson, 1986, *Relevance: communication and cognition*. Oxford: Blackwell.

Walker, Marilyn and Ellen Prince, forthcoming, A bilateral approach to givenness: a Hearer-Status algorithm and a Centering algorithm. In J. Gundel and T. Fretheim, eds *Reference accessibility*.

Meyerhoff, M. "Comparing old and new information in Bislama: nominal deletion with olsem". In Tent, J. and Mugler, F. editors, *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.*
C-141:85-93. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.85*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.

SOUTH INDIAN LANGUAGES IN FIJI: LANGUAGE CONTACT AND ATTRITION

FRANCE MUGLER

1. SOUTH INDIANS AND DRAVIDIAN LANGUAGES IN FIJI[1]

Nearly half of Fiji's present population are of Indian origin, and most Indo-Fijians are descendants of indentured labourers transported between 1879 and 1916, mainly to work on sugarcane plantations. Of the 60,000 or so indentured labourers who came to Fiji, a little over three quarters were recruited in North India, through the emigration depot in Calcutta. About 15,000, or nearly 25%, came from South India, starting in 1903 when recruitment extended to Madras (Table 1).[2]

TABLE 1: ORIGINS OF FIJI'S INDENTURED LABOURERS
(Source: Lal 1983:50)

Recruited through	Number	Percentage
Calcutta	45,439	75.02
Madras	**15,132**	**24.98**
Total	60,571	100

1 My thanks go to the informants, who gave generously of their time, to the interviewers, my colleague Veena Khan and students Urmila Chandra, Nirwani Devi, Rukeshma Lata, Vasu Deo Maharaj, Mohammed Taufil Omar, Anita Prasad, Geeta Prasad, Ruven Prasad, Chandra Rao, Estha Sami, to my colleague Jan Tent, and my collaborators in India, Sam Mohan Lal, V. Saratchandran Nair and G. Vijayasarathi. Data gathering for the survey and recorded language samples was made possible by two research grants from the University of the South Pacific (USP), and the analysis of recorded data was carried out from September to December 1994 while I was on study leave from USP and affiliated to the Central Institute of Indian Languages (CIIL) in Mysore, India. The support of both institutions is gratefully acknowledged. Thanks also to Rekha Sharma, Sudhir Mathur, O.N. Koul, N. Nadaraja Pillai, Uma Rani for their comments and interest, and Rajeshwari HS for computer assistance. I am particularly grateful to E. Annamalai, then Director of CIIL, for his comments and his support throughout.

2 Very small numbers of South Indians came to Fiji before recruitment started from Madras: according to Lal, 0.2% of indentured labourers recruited in the North were of South Indian origin, and there were probably some South Indians among the 600 or so labourers who reindentured from other colonies, especially Natal and Mauritius (Lal 1983:44–54).

Jan Tent and France Mugler, eds *SICOL: proceedings of the Second International Conference on Oceanic Linguistics:* vol.1, *Language Contact*, 95–108.
Pacific Linguistics, C–141, 1998.
© France Mugler

Mugler, F. "South Indian languages in Fiji: language contact and attrition". In Tent, J. and Mugler, F. editors, *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.*
C-141:95-108. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.95*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.

Most indentured labourers recruited in South India spoke a Dravidian language as their first language—mainly Tamil, Telugu or Malayalam—while a few others spoke Hindustani or Marathi. It is difficult to identify the language of the migrants, which was not indicated on emigration passes, except in 1903, and their district of origin is not always a reliable guide to identifying their first language (Siegel 1987:145).

2. LANGUAGE SHIFT

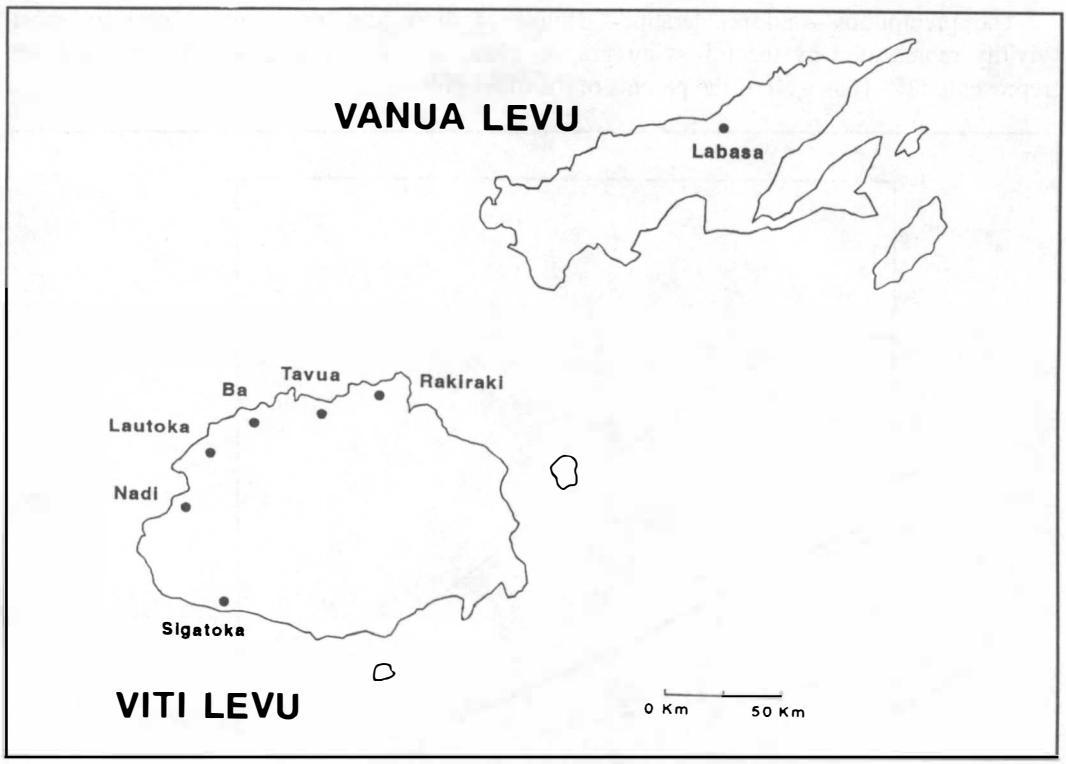
It is also difficult to estimate current numbers of Indo-Fijians of South Indian origin—let alone numbers of speakers of Tamil, Telugu, or Malayalam—since all persons of Indian ancestry are lumped together in the Fiji censuses, which also normally do not contain any information on language. Only in 1956 and 1966 did the census include a question which asked Fiji Indians to indicate the main language spoken by members of their household. Table 2 shows the number of households where a South Indian language was reported as the main language used at home. A comparison of results in the two censuses reveals a decline for all three languages, as shown in the table by the percentage of loss over the 10-year period, calculated by Siegel:

TABLE 2: HOME LANGUAGES OF INDO-FIJIANS (1956 and 1966 censuses)
(Source: Adapted from Siegel 1987:205)

	NUMBER OF HOUSEHOLDS		
Language	1956	1966	Percentage of loss
Tamil	1,498	999	33.3
Telegu	797	301	64.9
Malayalam	134	47	62
Total	2,429	1,347	44.5

The decline of the Dravidian languages probably started much earlier, however. South Indians have been a minority among Fiji Indians since the beginning, and they started arriving nearly 25 years after the first North Indians. By then, a koinê based on various dialects of Hindi and Hindustani, now known as *Fiji Baat* or Fiji Hindi, had become established as the lingua franca on plantations. Although a number of interpreters worked with Dravidian languages during indenture, there was also an unofficial policy of encouraging Hindi at the expense of other Indian languages (Siegel 1987:162). It is likely then that Dravidian languages underwent a sudden and extensive shrinking of domains and were soon reduced, as Pillai (1971:3) claims, to "domestic languages".

Dravidian languages in Fiji continue to lose speakers, as is shown by the results of a 1993 survey of nearly 500 Indo-Fijians of South Indian descent in the country's sugarcane belt, in the West of the main island, Viti Levu, and the North of the second largest island, Vanua Levu (for details, see map and Mugler and Tent, this volume). Table 3 shows the percentage of speakers of Tamil, Telugu and Malayalam in the sample surveyed. The results are displayed according to five age cohorts: three represent the age-groups surveyed among participants themselves: 15–25 years old, 30–45 years old, and over 50; the other two cohorts are the parents and grandparents of the older group in the sample, since questions were included about the knowledge of Dravidian languages of the participants' forebears.



MAP: AREAS SURVEYED IN 1993

TABLE 3: 1993 SURVEY—PERCENTAGE OF SPEAKERS FOR FIVE AGE COHORTS

Language	Grandparents	Parents	50+ years	30–45 years	15–25 years
Tamil	78.2	73.2	59.5	31	11.7
Telegu	42.4	42.6	25.2	7.7	1.2
Malayalam	13.4	15.2	10.4	2.6	.6

The loss of speakers over the generations is clear. Table 4 shows the percentage of speakers each language has lost between the three age-groups surveyed: between two generations, the older and the middle-aged groups, then the middle-aged and the younger group, and over three generations, between the older and younger groups:

TABLE 4: 1993 SURVEY—PERCENTAGE OF LOSS BETWEEN THREE AGE-GROUPS

Language	Between 50+ & 30–45	Between 30–45 & 15–25	Between 50+ & 15–25
Tamil	48	62	80
Telegu	69	84	95
Malayalam	75	76	94

The precipitous—and accelerating—decline of Dravidian languages is perhaps most vividly represented by the following graph, where the three age-groups interviewed are represented (2–4) as well as the parents of the older group (1).

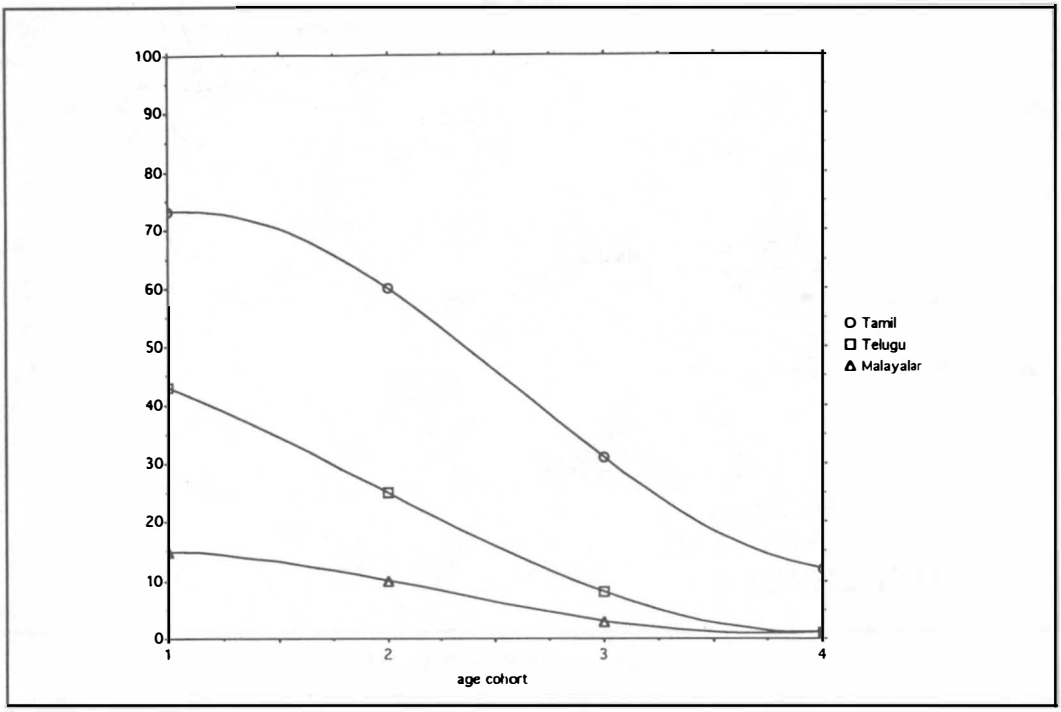


FIGURE: REPORTED KNOWLEDGE OF A DRAVIDIAN LANGUAGE BY FOUR AGE COHORTS

All the survey figures above are based on the participants' responses to a general question on whether they know one or several Dravidian languages. A question about whether the language of their ancestors is spoken at home yields even lower figures, while answers to specific questions about language use at home with various interlocutors indicate that there is very little use of Dravidian languages (Table 5). These figures show that use of a Dravidian language is correlated not only with the age of the speaker, but also that of the interlocutor: Dravidian languages are more likely to be used at home with parents than with a spouse or siblings, and least likely when addressing children.

TABLE 5: PERCENTAGE OF LANGUAGE USED WHEN SPEAKING WITH...

Language	Parents	Spouse	Siblings	Relatives	Friends	Children
a Dravidian language	14	8	8	6	3	2
Fiji Hindi	78	67	89	94	96	75
English	.8	.6	1	.6	1.5	2

Table 5 also shows the overwhelming preponderance of Fiji Hindi among descendants of South Indians. Even in this most intimate domain—home—the use of possible competing languages (i.e. Dravidian languages and English) is extremely limited. The high figures for

Fiji Hindi represent a number of different factors: some of the informants do not know any Dravidian language at all, some have limited competence or just a passive knowledge, while others who may have sufficient competence are the only members of the household to speak the language of their forebears. The shift from Dravidian languages to Fiji Hindi is almost complete then, and Fiji Hindi has become the first language of nearly all descendants of South Indians. Not only is it the first language of those who know no Dravidian language at all, but for most of those who still do have some knowledge of one, Fiji Hindi is in almost every case the language they feel most competent in and use most frequently.

3. THE STRUCTURE OF DRAVIDIAN LANGUAGES IN FIJI

Given the steady loss of speakers of Dravidian languages, several questions arise: Do the varieties spoken in Fiji exhibit any structural phenomena of attrition? What is the impact on these varieties of the languages they have been in contact with in Fiji, particularly the dominant language of the community, Fiji Hindi? Can we legitimately call these varieties Fiji Tamil, Fiji Telugu, Fiji Malayalam?

In an attempt to answer these questions, the 1993 survey was followed by another stage of research, in which two kinds of data were recorded. Firstly, a series of semi-structured interviews were conducted with some of the older people who had been administered the survey questionnaire. These interviews contained questions about the geographical origins in India of the participants' ancestors, their own language use as they were growing up, and questions gauging their attitudes towards the Dravidian language of their ancestors and towards current maintenance efforts. In a separate exercise, samples of conversations in Tamil, Telugu and Malayalam were recorded between an older person and a relative or neighbour, some featuring participants previously interviewed while others were of older descendants of South Indians known to the interviewers.[3]

Over 40 samples have been recorded so far, at least a dozen for each of the three languages. Although the older age-group (over 50) was targeted, many conversations recorded include relatives or neighbours from different age-groups, so that participants' ages range from 10 to 78. Sample length goes from 5 to over 30 minutes, and the total time of recoded data is a little over six hours. While this is not sufficient for any generalisations about each variety of language as a whole, it is enough to give an indication of structural trends and of what avenues of inquiry are worth pursuing.

The recorded samples were analysed over a period of three months in late 1994 at the Central Institute of Indian Languages in Mysore, India, in cooperation with three Dravidian specialists: Sam Mohan Lal for Tamil, V. Saratchandran Nair for Malayalam, and G. Vijayasarathi for Telugu.[4] What follows is a sketch of findings to date.

Firstly, the Tamil, Telugu and Malayalam spoken by our Fiji informants are each clearly identifiable as overseas varieties, distinct from any variety spoken in India. Secondly, and not unexpectedly, one of the characteristics of all three Fiji varieties is the retention of

---

3 The interviewers were part of the team who had administered questionnaires during the survey: ten senior linguistics students at the University of the South Pacific and one colleague from the Department of Literature and Language (see fn.1).

4 For details see Mugler and Mohan Lal (1995), Mugler and Nair (1995), and Mugler and Vijayasarathi (1996).

archaisms, and of dialectal or other features which allow for identification of the background of informants: North Arcot, Trichur, Rayalaseema, Malayali Muslim, etc. Such clues are found at the phonological, syntactic, lexical and semantic levels. Most interesting from the point of view of language contact is the influence of Fiji Hindi, now the dominant language of the community.

3.1 LANGUAGE CONTACT

3.1.1 LEXICO-SEMANTICS

The Tamil, Telugu and Malayalam spoken in Fiji have borrowed a number of words from the languages they have been in contact with in their new sociocultural environment, namely Fiji Hindi, English and Fijian, in that order of numerical importance.

The most common borrowing from Hindi is *accha* 'okay/fine', which has become a pan-Indian feature in the languages of India, and is commonly used in all three Dravidian languages present in Fiji. In some Tamil informants' speech *accha* alternates with Tamil *sēri* while in that of others it replaces it.[5]

Identifying borrowings from Fiji Hindi is not always straightforward since Dravidian languages in India themselves have borrowed liberally from Hindi and from Sanskrit. Some words may have once been used in, say, the Tamil of India, such as *bāśa* 'language', which has long been archaic in India and therefore, in the Tamil of Fiji, is more likely to have been borrowed from Fiji Hindi. The origin of other borrowings is less ambiguous, such as *lēkin* 'but', *bimāri* 'sickness', *barābar* 'sufficient', *rasta* 'road' in the Telugu data but absent from the Telugu of India, or Tamil *kēti* 'farm' *kalas* 'finished', *karab* 'bad', all unknown in Indian Tamil, as is *band* 'closed/stopped', in *vikkalu band āy paōccu* 'the hiccoughs stopped'.

Fiji Hindi borrowings are adapted to the phonology of the borrower language, so that the initial [k] in Tamil *kēti*, *kalas* and *karab*, for example, is deaspirated. Similarly, *ekdam* 'absolutely', a word unknown in the Malayalam of India, is realised as [ʸekdam], with the obligatory on-glide of Malayalam. On the other hand, many nouns are 'Dravidianised' with the suffix *-m*, as are Sanskrit and Hindi borrowings in the Dravidian languages of India. Thus we find, for instance, *dēram* 'duration' in the Telugu of Fiji.

In some cases the phonological shape or the particular meaning of a borrowing indicates its Fiji Hindi origin. The word *dyānam*, from Sanskrit, means 'knowledge' in the following sentence from the Tamil data:

(1) *Tamiḻle dyānam varudulla avugaḷukku.*
They do not get knowledge in Tamil.

In Indian Tamil the form is *ñānam* and the meaning 'meditation'. Here, the initial cluster *dy-* of Hindi, along with the word's meaning, indicates that the word is more likely to have been borrowed from Fiji Hindi—and nativised with the usual Dravidian nominal suffix *-m*, than to have undergone a semantic shift from Indian Tamil.

Similarly, in the Malayalam data, *śikṣa*, also found in the Malayalam of India, here has the primary Hindi sense of 'education' rather than the primary Malayalam sense of 'punishment':

---

5 One informant who normally alternates between the two words uses both concurrently on one occasion, as a phrase realised with a single intonation contour, *accha sēri*.

(2) *Accanammamāṛ nalla mātiri sīkṣa koṭuttu.*
Our parents gave us a good education.

This usage is also commonly found in the speech of second language learners of Malayalam whose first language is Hindi (Nair 1986:294, 1994:81).

The influence of Fiji Hindi on lexico-semantics can be illustrated with a few idioms from the Telugu data:

(3)		
	vivahasamskaram	marriage ceremony
	dhanyavādamulu iccunānu	to give thanks
	kṣāma ceyyāni	to pardon
	bajanam pādatāru	to sing *bhajans*

The meaning of *samskara* in the Telugu of India is restricted to 'funeral rites', while here in *vivahasamskaram* (lit. 'marriage rites') the word is used in the more general Hindi sense of 'rite, ritual, ceremony'. Similarly, in *dhanyavādamulu iccunānu* '[I remain,] having given thanks', the verb *telupū* (lit. 'to tell') which would be used in the Indian Telugu idiom, is replaced by *icci* (lit. 'to give') a calque of Fiji Hindi *dhanyavad dena.* Another calque is *kṣāma ceyyāndi* (lit. 'to do/make pardon') where the verbaliser *incu* of Telugu—*kṣāmindadi* 'to pardon'—is replaced with the very productive verb *karna* 'to do', as in Fiji Hindi *śama* (or *māf*) *karna.* The Fiji Telugu form is also morphologically simpler than the Indian Telugu idiom. In *bajanam pādatāru* (lit. 'they (always) sing *bhajans*') Indian Telugu would have *cēstāru* 'do', since *bhajans* in Andhra involve dancing, while when singing is involved, *kirtana* is used, rather than *bajana.* Finally, *sep* 'tell' (cf. Standard Telugu *ceppu*) appears instead of *tsaduvu* 'read' in the context of reciting mantras, a usage probably influenced by Fiji Hindi, where both *psṛhe* 'to read' and *bole* 'to tell' can occur in this idiom.

A number of words have been borrowed from English. In Malayalam we find, among others, *class*, *meeting*, *high school*, *English*; in Tamil *fast*, *mistake*, *fees*; in Telugu *doorstep*, *town area*, *medical*, *family*; and in all three, *doctor*, *government*, *master*, *hospital*, *committee*. All words of English origin present in the data also exist in Fiji Hindi and it is likely that most were borrowed through Fiji Hindi, now the dominant language of the community, rather than directly from English.[6] The word *school*, for example, which has also been borrowed by Tamil in India as [skul], is realised in the Fiji Tamil data as [iskul], with the obligatory prothetic vowel which precedes the consonant cluster [sk] in Fiji Hindi. Fiji Hindi *girmit* 'indenture', originally from English 'agreement', refers specifically to the experience of the immigrant labourers and is unknown in India. The Dravidianising *-m* ending is occasionally extended to nouns borrowed from English, as in Fiji Telugu *designam* 'design'.

Very few Fijian borrowings appear, and this may simply be due to the limited nature of the data. In Telugu, for instance, we find *koro* 'Fijian village', *kaibiti* (< Fijian *kaiviti*) 'Fijian', here 'Fijian language'. All these also exist in Fiji Hindi, as does *naqona* 'kava', which appears in its Fiji Hindi form (< *na yaqona* 'the kava') rather than the original Fijian

6 Many of these borrowings from English are also found in the speech of educated speakers of Tamil, Telugu and Malayalam in India nowadays. This is a fairly recent phenomenon however, and in the Fiji data they occur in everyone's speech, regardless of level of education.

yaqona.[7] As in the case of borrowings from English, Fijian words are likely to have come into the Dravidian languages of Fiji via Fiji Hindi, the dominant language of the informants.

3.1.2 PHONOLOGY

The influence of Fiji Hindi may extend, beyond lexico-semantics, to certain aspects of phonology, in particular syllable structure, although the evidence is sometimes ambiguous. Here are a few examples from the Tamil data, contrasted with the corresponding forms in the Tamil of India:

(4)	*nāl piḷḷaihaḷ*	*nālu piḷḷaihaḷ*	'four children'
	mul muyaṛeci	*mulu muyaṛeci*	'complete effort'
	nal mādiri	*nallā mādiri*	'good manner'
	pāṭ pāḍa	*pāṭṭu pāḍa*	'sing a song'
	mūṇ pēru	*mūṇu pēru*	'three people'

In each case the word-final, unstressed, vowel of Indian Tamil is deleted, resulting in a consonant cluster across word boundaries. Unstressed vowel deletion also occurs internally, as in *evḷavu* for *evvaḷavu* 'how much', *selvu* for *seluvu* 'cost', or the reduplicated forms *cincinna* for *cinnacinna* 'very small', *kaskasa* for *kasakasa* 'something fishy', *koñckoñca* for *koñcokoñcō* 'just a little bit'. This pattern of vowel deletion also exists in the Northern dialects of Tamil in India, especially with [u], which in unstressed position is normally realised as unrounded but often is instead lost altogether. But it is also typical of the Tamil spoken by first language speakers of Hindi, and it is difficult to attribute this feature unambiguously to retention or to the influence of Fiji Hindi.

The same pattern of unstressed vowel deletion leading to consonant clusters is found in the Telugu data:[8]

(5)	*dagra*	for	*daggara*	'at, near'
	bōdnam		*bōjanam*	'food'
	bidlu		*biḍḍalu*	'children'
	manku		*manaku*	1PL.INC.DAT 'to us'
	wāḍki		*wāḍiki*	3SG.DAT 'to him'
	walki		*waḷḷaki*	3PL.DAT 'to them'

In contrast to Tamil, this feature is not known to exist in any dialect of Telugu in India. However, it also is found in the speech of Telugu learners whose first language is Hindi.

---

7 Kava, a traditional drink used for ceremonial and social purposes, is made from the root of *Piper methysticum*.

8 Abbreviations used in this paper are as follows:

DAT	dative	PERF	perfect
INC	inclusive	PL	plural
INF	infinitive	SG	singular
NEG	negative	VP	verb phrase
PART	participle		

3.2 LANGUAGE ATTRITION

At the morphosyntactic level, a number of structures are characterised by loss of marking or great variation in how certain distinctions are marked, which points to the possibility of attrition.

3.2.1 SIMPLIFICATION OF SANDHI

The sandhi rules that apply in the Indian varieties of the three Dravidian languages are often not observed, so that a case marker is added directly to the bare stem of the noun. In the Tamil data, for instance, we find *saṅgamle* 'in the Sangam', rather than *saṅgattale*, from which the sandhi increment *-tt-* is absent. Other examples, contrasted with the Indian forms, are:

(6)	*būlōgamle*	*bulogattile*	'in the world'
	sāyaṅgalamendu	*sāyaṅgalattilirindu*	'since evening'

Similarly, we have *kaṣṭam paḍa māṭṭāṅga* 'they won't face any difficulty', rather than *kaṣṭappaḍa māṭṭāga*, in which Indian Tamil drops the final *-m* and geminates the initial *-p* of the next word. This is clearly a case of morphological simplification.

3.2.2 NUMBER MARKING

In the Tamil data the plural marker *kaḷ* is generalised to all count nouns, including collectives: in *makkaḷka* 'children', for example, *makkaḷ* is treated as monomorphemic rather than already marked for plural, and is therefore suffixed with *kaḷ*.

Similarly in the Malayalam data, the plural *kār* is overgeneralised in *bāśakkār* 'languages', *malayāḷakkār kuṭṭikal* 'Malayali children', *malayāḷakkārakkum telungukkārakkum* 'for Malayali and Telugu [people]'. This last example is morphologically less complex than the Indian Malayalam form *malayāḷikaḷkkumtelungaṛkkum*. But this overgeneralisation is also found among second language learners of Malayalam (Nair 1986:288). On the other hand, in *mūnna penkuṭṭi, mūnna aṅkuṭṭi* (lit. 'three girl, three boy') the plural marker is absent—another simplification since, with the presence of a number, the nominal plural marker is redundant.

Another Malayalam plural marker, *mār*, normally used only with kinship terms, is extended by analogy, as in *keraḷanmārellam naḍakkaṇam* 'Keralites should follow'.

The Telugu data is particularly rich with respect to variation in number marking. Collective markers for human nouns are often omitted after numerals, as in Telugu *magapillōlu nālugu*, 'four boys' or *nāku tommidi pilakāyalantā kaligæ*, 'I had nine children'. The Standard Telugu collective markers are *guru* with numbers 3–7 and *mandi* with 8 and above, while the Rayalaseema dialects—features of which are abundant in the data—have *mandi* for 3 and above. The second example, besides missing any collective marker, has clear Rayalaseema features: *pilakāyalantā*rather than Standard *pillalu* for 'children' and, at the end of *kaligæ*, the absence of *-ru* concord, an optional marker in Rayalaseema. But the same speaker produces later the phrase *antā tommidimandi* 'all nine', this time with the collective marker. In Fiji Telugu the collective marker may have become optional, and may be in the process of being lost. Neither Tamil nor Kannada—both of which have influenced the Rayalaseema dialects—have this marker, which is also absent from Fiji Hindi. While

there is no loss of distinction since the information carried by the marker is redundant, the loss of this collective does result in morphological simplification.

There is a great deal of variation in the use of number concord markers. This is particularly extreme in the speech of two Telugu informants, who often use a plural noun with a singular concord marker on the verb, or suffix a nominal number marker to a verb:

(7) *Janam lēdu.*
people.PL EXIST.NEG.SG (for *lēru*, plural)
There are no people. (lit. There is no people.)

(8) *Aidu pillōlu unḍādi.*
five child.PL EXIST.SG (for *undaru*, plural)
There are five people here. (lit. Five children is here.)

In the same informant's speech, *andaru unḍādi* (lit. 'all people is...') with the singular verb ending *-di* with a plural subject, alternates with *andaru unḍālu*, with the plural nominal ending *-lu* suffixed to the verb, rather than the plural verbal ending *-ru*.

While the use of singular rather than plural concord markers may be explained by an influence of the invariability of such constructions in Fiji Hindi (as well as of Kannada and Tamil on some varieties of Telugu in India), the use of nominal rather than verbal markers cannot, and the extreme variation can probably best be explained as a sign of an incipient breakdown of the morphological system in the marking of concord.

A similar confusion between noun and verb endings is found in *wāḍki erukunu* 'he only knows', which has the pronoun in the dative and a verb (with the concord *-nu*), whereas Indian Telugu can have either *wāḍu erugunu*, with the pronoun in the nominative and a verb, or *wāḍiki eruka*, with the pronoun in the dative and a noun.

3.2.3 CASES

In all three Fiji varieties there is a great deal of variation in case endings and deviation from the Indian varieties. In the Tamil data, one case is entirely missing (the genitive marker for nouns, which is replaced by the dative or the accusative), the accusative is omitted in some environments, and there is extreme variation in the use of the dative, which appears where Indian Tamil would have the genitive, the accusative, the locative or even the nominative.

In Malayalam, we find some genitives, locatives and accusatives missing, the accusative often overgeneralised to non-humans, the sociative omitted with communicative verbs, and the dative sometimes used instead of the accusative, the locative and the ablative. In Telugu the most regular differences between the Fiji data and the Telugu of India are the use of the dative for the instrumental/sociative, the sociative for the ablative, and the nominative for the genitive. A few examples in each of the three languages will illustrate this extreme variation in case marking:

(a) Tamil

(a_1) Dative for genitive

(9) *Kaavuukku pēre eḍukkuṛē.*
I take the name of God.

The genitive *-uṭaya* does not appear at all in the data and is often replaced by the dative, as it is here in *kaḍavuḷukku* instead of *kaḍavuḷuṭaya*. While the use of the dative for genitive is said to be found in some dialects of Indian Tamil, such as in the Kanyakumari area, Mohan Lal (pers. comm.) argues that the realisation of the two case markers is in fact different, with [ikku] for the dative and [ikka] for the genitive.

(a_2) Zero for accusative

(10) *Oru payyẽ n ̣pākṛẽ.*
I'm looking after a boy.

The noun in direct object position is [+ human] and therefore has an obligatory accusative ending in Indian Tamil. The zero marking here may be an overgeneralisation of the optional marking of [- human] direct objects.

(a_3) Dative for locative

(11) *Ide patti tān manasukku eppavum neneccu koṇḍu irukkiadu.*
In our minds, we always think about this [going to India].

Here Indian Tamil would have the locative *manasule* rather than the dative *manasukku.*

(b) Malayalam

(b_1) Zero for locative

(12) *Fiji vannu.*
I came to Fiji.

But the same speaker does use the locative ending later in a similar environment:

(13) *Fijiyil balaṛe malayāḷakkāṛuṇḍa.*
In Fiji there are many Malayali.

The locative is also missing in:

(14) *Kerala nial pōyi.*
You went to Kerala?

(15) *Ippo muslīm skūlāna niṅṅala paḍippikuṇṇata.*
Now it's the Muslim school you're teaching in.

(b_2) Accusative with non-human objects

(16) *Ente makkaḷ tanṛe bhāśaye paṛayān.*
My children would speak the language.

The accusative *bhāśaye* is used where Indian Malayalam would have *bhāśa.* Here the accusative of [+ human] direct objects is overgeneralised—a feature found in the speech of L_2 speakers of Malayalam (Nair 1986:302). This is a mirror image of the overgeneralisation of the optional zero marking of [- human] direct objects in Tamil (see above).

(b_3) Zero for sociative with communication verbs

(17) *Itṛayum paṛaññu ñān santōśam.*
(lit. I told this much with happiness.)

The expected sociative would be *santōśatōde.*

(b_4) Dative for locative

(18) *Enṛe accan kannikkōṭṭaninna fijikke vannu.*
My father came from Cannicut to Fiji.

Generally, the use of one case for another is commonly found among second language speakers of Malayalam (Nair 1986:293, 298, 302–303).

(c) Telugu

(c_1) Dative for sociative

The dative *-ki* is used instead of the instrumental/sociative *-to* in this phrase:

(19) *Dēwuḍiki prātana cēs.*
pray to God

(c_2) Sociative for ablative

(20) *Iṭuwanṭi pīnugatō gūḍa ḍabbu tīstaru.*
Even from a dead body they make money.

Here Indian Telugu would have, instead of the sociative *pīnugatō*, the ablative *pīnuganunci* (or the postposition *mīda* 'on'). On the other hand the sociative *-tō* alternates with the postposition *kūḍā* in *famli kūḍā*, then *famlitō* 'with the family'.

(c_3) Nominative for genitive

In *ī illu wāllu* 'the people of this house', the nominative *illu* 'house' is used rather than the expected genitive of Indian Telugu *inṭi*. This simplification could well be influenced by the corresponding Fiji Hindi phrase *ghar wale*.

Generally, the replacement of one case marker by another often occurs in the speech of second language learners of the three Dravidian languages in question, including those whose first language is Hindi. Fiji Hindi has even less case marking than Standard and other Indian varieties of Hindi, and since Fiji Hindi is now the dominant language of most speakers of Tamil, Telugu and Malayalam in Fiji, case substitution and/or case marking loss may well be attributable at least in part to the influence of Fiji Hindi. Nonetheless, the loss of case marking in particular results in morphological simplification, which is a well-attested characteristic of language attrition. This simplification then is overdetermined, in that the influence of the dominant language of the community and processes of attrition can only reinforce each other.

3.2.4 SYNTAX

The most morphologically and syntactically complex constructions, such as causatives and relative participles, seem to be the ones which are most often and most drastically simplified. Ocasionally word order also differs from that of the Indian varieties. In the following example from the Tamil data, both verb and noun morphology are simplified:

(21) *En piḷḷaikaḷukku nān paḍittu vaittu irukkiryen.*
my children.DAT I study.VP make.VP be.PERF.PART
I made my children study [Tamil].

Where Indian Tamil has the accusative case for the object, the infinitive for the lexical verb and the conjugated, tensed form of the causative verb *vai*, this sentence has the dative of the noun and the participle of both the lexical and the causative verbs. There is a distinction in Indian Tamil between the causative (INF + *vai*) and an action done for future use (PART + *vai*), and the use of participle in causative constructions in the data may indicate that this distinction is being lost.

In the following causative sentence from the Malayalam data, there is some confusion between the transitive and intransitive:

(22) *Ñaṅṅaḷkka ī malayāḷam paḍippikkān nalla oru māsṫarkku bēnam*
We want a good teacher to teach us Malayalam.

In Indian Malayalam the sentence would be *ñaṅṅaḷe ... paḍippikkān*, with the pronoun in the accusative and the causative form of the verb, or *ñaaḷkka ... paḍikkān*, with the pronoun in the dative and the transitive form of the verb. The first sentence could be translated as 'we want a teacher to teach us Malayalam', and the second 'we want a teacher for us to learn Malayalam', since *paḍikkān* is 'to learn' and *paḍippikkān*, the causative, 'to teach' (lit. 'to cause to learn'). These two sentences of course have the same basic meaning. The sentence used by the informant then, compared to the Indian Malayalam sentence, has either a deviant case (dative for accusative) or a deviant verb form (causative for transitive). This transitive/intransitive confusion is another feature of the Malayalam of native speakers of Hindi (Nair 1986:298–299, 293).

A particularly dysfluent sentence in the Telugu data also exhibits greatly simplified structures:

(23) *Nēnu peḷḷijēsi...peḷḷijēsinānu...paddhenimidisamatsaramulu.*
I got married 18 years ago. (lit. I marr...married...18 years.)

The reflexive *kun* is omitted, so that this sounds as if the speaker performed the marriage, while the intended meaning is clearly that of 'getting married', which would be in Indian Telugu *peḷḷicēsukunnānu*. There is no postposition, such as *kritam* (in Standard Telugu) or *mundu* (in the Rayalaseema dialects) 'before, ago', to connect the phrase for '18 years' with the preceding clause. If there was the word for 'year' would end with the genitive *-la* rather than the nominative *-lu*.

4. CONCLUSION

The Tamil, Telugu and Malayalam spoken in Fiji are readily identifiable as diasporic varieties, and it is not surprising to find in these transplanted languages which have been isolated for almost a hundred years such characteristics as retention of archaisms and of dialectal features which point to the geographical origins in India of the informants' ancestors. But one of the most noticeable and pervasive influences has been that of the major language the three Dravidian languages have been in contact with—Fiji Hindi. The lingua franca on plantations for nearly a quarter of a century before the arrival of indentured labourers recruited in South India, Fiji Hindi has continued to both displace and influence the Dravidian languages ever since. Indeed, for most descendants of South Indians, Fiji Hindi has *re*placed the Dravidian language of their ancestors, while for those who still have some knowledge of that language, Fiji Hindi has become the dominant language. Thus, if for many informants the Dravidian language of their forebears is their 'mother tongue' in the

sense of their ancestral language and often the language they feel the greatest emotional attachment and loyalty to, Fiji Hindi is just as clearly their first language, for many in the chronological sense (the first language they acquired as children), and for nearly all others in terms of competence, ease and fluency (their 'best' language) and frequency of use.

The influence of Fiji Hindi can be detected in the phonology, lexicon and semantics of the three Dravidian languages, while their morphology and syntax show a good deal of simplification and reduction of marking, a common sign of attrition not unexpected in languages whose domains and frequency of use have been drastically curtailed. Since Fiji Hindi is substantially less marked than Dravidian languages, it is often not possible to unambiguously attribute a particular case of reduction in marking to language attrition rather than to the influence of Fiji Hindi as the major contact language. Clearly both processes go in the same direction, that of reduction of morphosyntactic complexity. Indeed, we note that many such features found in the data also occur in the speech of second language speakers of Tamil, Telugu or Malayalam—and not only among those whose first language is Hindi. Of course, as mentioned earlier, many of our informants are, in terms of their competence, much more like second language speakers of, say, Tamil, than like first language speakers. And simplification is a common feature not only of language attrition but also of language acquisition, of pidginisation and generally of situations of restricted input—hence the importance of studies of the structural erosion of languages under stress.

The limited data available so far precludes anything more than a tentative conclusion. In particular, it is difficult at this stage to ascertain to what extent the features discussed here are typical of the whole language varieties or of the particular individuals recorded. Data gathering continues and will, it is hoped, include more middle-aged and younger informants. Fiji speakers of all three Dravidian languages are fewer with each generation, and it is important to record the speech of those who still have some knowledge of these languages, in the hope of throwing some light on the processes of structural change which affect minority languages undergoing drastic shift.

REFERENCES

Lal, Brij, 1983, *Girmitiyas: the origins of the Fiji Indians*. Canberra: Journal of Pacific History.

Mugler, France and Sam Mohan Lal, 1995, Fiji Tamil: the structure of a language under threat. *International Journal of Dravidian Linguistics* 24/2:118–133.

Mugler, France and V. Saratchandran Nair, 1995, Notes on Fiji Malayalam. Paper presented at the Twenty-Third All India Conference, Trivandrum, India, 22–25 June.

Mugler, France and G. Vijayasarathi, 1996, Telugu in Fiji. Paper presented at the Twenty-Fourth All India Conference, Kuppam, India, 20–22 June.

Nair, V. Saratchandran, 1986, A study of the errors of the second language learners in Malayalam. In K. Rangan, ed. *Proceedings of the Thirteenth All India Conference of DravidianLinguistics*. Thanjavur: Tamil University.

1994, Lexical simplification in interlanguage: a case study of L2 learners of Malayalam. *South Asian Language Review* 4/2:80–92.

Pillai, Raymond C., 1971, The emergence and decline of the T.I.S.I. Sangam. MS. Pacific Collection, University of the South Pacific, Suva, Fiji.

Siegel, Jeff, 1987, *Language contact in a plantation environment: a sociolinguistic history of Fiji*. Cambridge: Cambridge University Press.

Mugler, F. "South Indian languages in Fiji: language contact and attrition". In Tent, J. and Mugler, F. editors, *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.*
C-141:95-108. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.95*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.

SOME ASPECTS OF LANGUAGE USE AND ATTITUDES IN FIJI

FRANCE MUGLER AND JAN TENT

1. INTRODUCTION

A survey of language use and attitudes in Fiji was conducted by the authors in mid-1993.[1] Its aim was to identify the use and functions of, and attitudes towards, several of the languages of Fiji; and to ascertain in particular the extent of language shift and/or maintenance. The survey involved administering almost 1,000 questionnaires to a stratified sample of people in two major geographical regions. Firstly, in the capital, Suva, Fijians and Indo-Fijians were surveyed with respect to their use of and attitudes towards Fijian (Standard Fijian and other communalects), Hindustani (both Standard—or *Shudh*—Hindi and Fiji Hindi)[2] and English. Secondly, on the western side of the main island of Viti Levu and in the north of the second largest island, Vanua Levu, descendants of South Indians were surveyed

1 We express our gratitude first to all respondents and to the students who conducted the interviews: Shannon Coalala, Nicholas Cornelius, Premila Devi, Asinate Gadolo, Manasa Kolivuso, Megnath Krishnan, Rusila Racule, Samina Sadiq, Shobhna Sen, Sangeeta Singh, John Soso, Tuipulotu Vugakoto ('Suva' Survey); Urmila Chandra, Nirwani Devi, Rukeshma Lata, Vasudeo Maharaj, Muhammed Taufil Omar, Anita Prasad, Gita Prasad, Ruven Prasad, Chandra Rao, Estha Sami ('South Indian' Survey).

Many others helped with various aspects of the research project and, even if we did not always follow their advice, we would like to thank them: Mike Delena, Donn Bayard, Janet Holmes and David Blair for the initial questionnaire design; Barry Richardson and Roland Schultz for advice on methodology and questionnaire design; Epeli Waqavonovono and Vijay Naidu for information on socioeconomic issues; Paul Geraghty and Pio Manoa for the Fijian version of the questionnaire; Ravin Prasad and Veena Khan for the Fiji Hindi versions of the questionnaires; Judy Adams for questionnaire preparation; Tiroam Neeri, Susan Ricketts and George Nicole for computer data entry; Donn Bayard, Nicola Walmsley, Josalice Caballes and Patrick Griffiths for advice on statistics.

Special thanks go to Veena Khan, who participated in most aspects of the survey, especially translation, interviewing and monitoring, and to Robin Taylor, for his invaluable assistance with the statistical analysis.

We are also grateful to many other students and colleagues for their interest and encouragement. This research project was funded by a University of the South Pacific research grant # 0702-9203. The support of USP is hereby gratefully acknowledged.

2 Fijian is characterised by a great deal of regional diversity, with about 300 communalects. We use 'communalect' here as "a variety spoken by people who claim they use the same speech" (Geraghty 1983:18), and 'Fijian' as a cover term for both Standard Fijian and these communalects. Fiji Hindi is a local variety of Hindi, which evolved from a koinê of various dialects of the Hindustani lingua franca of North India (Siegel 1987:187–203).

Jan Tent and France Mugler, eds *SICOL: proceedings of the Second International Conference on Oceanic Linguistics:* vol.1, *Language Contact*, 109–134.
Pacific Linguistics, C–141, 1998.
© France Mugler and Jan Tent

Mugler, F. and Tent, J. "Some aspects of language use and attitudes in Fiji". In Tent, J. and Mugler, F. editors, *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.*
C-141:109-134. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.109*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.

with respect to their use of, and attitudes towards, the Dravidian language of their ancestors (Tamil, Telugu or Malayalam), Hindi (both varieties) and English.

With this survey we attempt to study two distinct situations. One involves the two major languages of Fiji, Fijian and Hindi, their different varieties, and the interplay between these and the dominant official language and lingua franca, English, including the possibility of language shift. The other involves the relationships between minority languages which are known to have been undergoing shift in their community (the Dravidian languages), the dominant language of that community, Hindi, and, again, English. While these two situations are clearly different, there is also much overlap between them, and it was deemed that a single survey, using essentially the same instrument and the same methodology, would yield useful data, while each component of the survey could be further analysed separately.

2. BACKGROUND

2.1 THE LANGUAGE SITUATION IN FIJI

The two major languages in Fiji, in terms of numbers of native speakers, are Fijian and Fiji Hindi. While the term 'Standard Fijian' commonly designates the particular variety of Fijian[3] used as the lingua franca amongst Fijians, most Fijians speak a communalect as their first language. Among Indo-Fijians, who at the time of the survey made up about 47% of the population (Bureau of Statistics estimate), nearly everyone's first language is Fiji Hindi, while Shudh Hindi is used for some official functions.

Most Indo-Fijians are descendants of labourers who came to Fiji during the indenture period (1879–1916). Three quarters of indentured labourers came from the north of India and most spoke varieties of Hindi, while most of those who came from the south were speakers of Tamil, Telugu or Malayalam, all Dravidian languages (Lal 1983; Siegel 1987).

2.2 PREVIOUS SURVEYS

Four previous surveys dealing with language use and attitudes have been conducted in Fiji, each with at least some implications for language shift, each also with limitations in terms of methodology, geographical region(s) surveyed or issues investigated. They are Adam (1958), White (1971), Siegel (1973) and May (1990).

Adam (1958) reports on a large-scale survey (close to 5,000 responses) conducted in 1950 by the Education Department, with a focus on the attitudes of parents of school children towards the main language to be used in school. Results showed that a majority of parents advocated English, and that preference for English was related to place of residence, age, level of education and occupation.

The other three studies were on a much smaller scale and each limited to a single geographical region: White and May studied Suva, while Siegel surveyed the urban and rural population in and around Nadi. These surveys were more comprehensive in scope than Adam's, as they investigated not only attitudes towards language but also patterns of language use.

---

3 'Standard Fijian' evolved in towns and possibly in church institutions or boarding schools where Fijians from different regions came together for the first time. It is based on the communalects of south-eastern Vitu Levu (Paul Geraghty, pers. comm.).

White (1971) conducted a survey of language use as part of a larger survey of low income households in the Raiwai area of Suva. The language survey of 404 Indo-Fijians and Fijians dealt mainly with the use of English, and was conducted in English. Among results with implications for language shift, White found that English was the home language of a surprisingly high number of households, and that it was the main language used for communication between the two major ethnic groups, although bilingualism in Fijian and Fiji Hindi was also present and, he claimed, a typical urban phenomenon.

Siegel (1973) conducted a survey in the Nadi area of 145 Indo-Fijians who were interviewed in Fiji Hindi. He found that Shudh Hindi was being displaced by English for most formal language functions, such as reading, writing and public speaking. Concurrently, a comparison of the 1956 and 1966 censuses revealed that Fiji Hindi was displacing the Dravidian languages still used in Fiji. Contrary to White, Siegel also found that most Indo-Fijians communicated with Fijians in Fijian, rather than in English, and that bilingualism in Fijian and Fiji Hindi was therefore not exclusively an urban phenomenon. Thus, the pattern of language use was often quite complex, while language shift was not limited to a simple 'gain' by English.

May (1990) surveyed 834 people in the Suva area about their perception of their language skills. He reports that the answers to the few preliminary questions on language use indicate that young people were shifting towards greater use of English in the home.

Each of the four surveys briefly reviewed here suffers from limitations. Adam addresses himself only to the issue of language attitudes, and that in the specific context of education. White limits himself to the Suva City area and administered the questionnaire in English. Siegel's study deals only with Indo-Fijians in the Nadi area and he acknowledges that his sample was skewed in terms of sex and that some of his questions were too general. May's survey has very few questions dealing with language use and they also are extremely general. Three of the studies are also dated now.

3. THE 1993 SURVEY

3.1 HYPOTHESES

Our survey covered a fairly wide geographical region, a larger sample than most of these previous surveys, and a more detailed and wider array of issues related to both language use and attitudes. A stratified sample of the population was surveyed so as to establish which sections of the population are most closely involved in any shift.

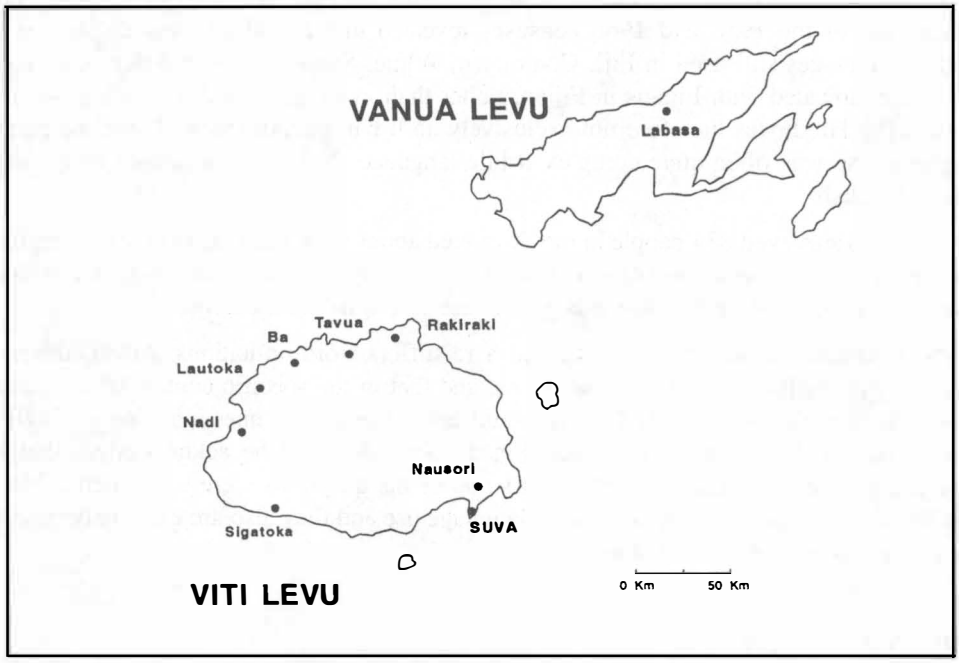
The major hypotheses tested are as follows:

- language shift is taking place;
- two different language shifts are occurring: an overall shift towards greater use of English, and a shift for speakers of Dravidian languages towards Fiji Hindi;
- the shift towards English implies, in most cases, an increase in bilingualism rather than a loss of the first language, which is being displaced in some domains, rather than replaced overall;
- the shift of speakers of Dravidian languages towards Fiji Hindi, on the contrary, implies a loss of the first language over time.

3.2 METHODOLOGY

3.2.1 REGIONS SELECTED AND PARTICIPANTS

The two regions surveyed were chosen for different reasons. The Suva-Nausori area was selected because, as the largest urban area in Fiji, it is the most likely to exhibit trends towards language shift. The west of Viti Levu and north of Vanua Levu, known as the 'sugarcane belt', are the areas most heavily populated by Indo-Fijians in general, and descendants of South Indians in particular.[4] The specific regions surveyed were urban and rural areas in and around Sigatoka, Nadi, Lautoka, Ba, Tavua and Rakiraki on Viti Levu, and Labasa and the Macuata rural area around Labasa on Vanua Levu.



MAP: REGIONS COVERED IN THE SURVEY

About 1,000 informants participated in the survey. Two samples of 504 informants each were designed for each region surveyed.[5] Both samples were stratified in terms of sex and age, with equal numbers of male and female respondents, reflecting closely proportions in the general population, and equal numbers in three age-groups (15–25, 30–45, 50+), since the behaviour of different age-groups is often related to the presence or absence of language shift. Information on education was also gathered, in particular the highest level of education attained, another variable thought to be related to language use and attitudes.

---

4 The Fiji censuses do not distinguish among Indo-Fijians on the basis of their Indian ancestors' origin, so that the geographical location of descendants of South Indians is not documented. Nonetheless, some regions of Fiji are well known as being home to a relatively high proportion of descendants of South Indians (Rakiraki, for instance), while the rest are scattered among the general Indo-Fijian population.

5 Each half of the survey was designed to be administered by a team of 12 interviewers, but only 11 were available in July for the SI survey. One interviewer administered a number of additional questionnaires in November, yielding a total of 481 respondents rather than the 504 planned.

The Suva sample was also stratified in terms of ethnicity and income, with equal numbers of Fijians and Indo-Fijians, the two major ethnic groups in Fiji, and equal numbers of respondents in the low and mid to high income groups (Table 1). The cut-off point between low and mid-high groups was set at an annual income of F$10,000 per household, the figure normally used for Fiji by sociologists (Vijay Naidu, pers. comm.).

TABLE 1: DESIGN OF THE SUVA SURVEY

Income group	Fijians		Indo-Fijians	
Mid-high income group	*males*	*females*	*males*	*females*
	15–25	15–25	15–25	15–25
	30–45	30–45	30–45	30–45
	50+	50+	50+	50+
Low income group	*males*	*females*	*males*	*females*
	15–25	15–25	15–25	15–25
	30–45	30–45	30–45	30–45
	50+	50+	50+	50+

(Note: The sample was designed to have 21 respondents per cell.)

The sample for the 'South Indian' survey[6] was stratified in terms of geographical region and area of residence, with equal numbers of respondents from the west of Viti Levu and the north of Vanua Levu, and two-thirds urban residents versus one-third urban dwellers, a proportion which reflects figures for the whole country (Table 2). Stratification by area of residence was selected rather than income—although information about income was also gathered—as it was deemed a more important variable, particularly in relation to the possible presence of language shift, urban dwellers being thought more likely to be shifting away from the South Indian language of their ancestors.

TABLE 2: DESIGN OF THE SOUTH INDIAN SURVEY

Residency	Viti Levu		Vanua Levu	
Urban residents	*males*	*females*	*males*	*females*
	15–25	15–25	15–25	15–25
	30–45	30–45	30–45	30–45
	50+	50+	50+	50+
Rural residents	*males*	*females*	*males*	*females*
	15–25	15–25	15–25	15–25
	30–45	30–45	30–45	30–45
	50+	50+	50+	50+

(Note: The sample was designed to have 28 respondents per rural cell and 14 per urban cell.)

6 The term 'South Indian' survey ('SI survey', for short) is used in this paper, reluctantly, as shorthand. The accurate term, unfortunately rather unwieldy, would be 'survey of Indo-Fijians of South Indian descent'.

3.2.2 INSTRUMENTS

Two slightly different questionnaires were designed: one for the Suva survey, one for the South Indian (SI) survey. The Suva questionnaire was printed in three versions—English, Standard Fijian, and Fiji Hindi—and the SI questionnaire in two—English and Fiji Hindi. Both questionnaires, in all language versions, were piloted, then slightly modified.

Each version of the questionnaire contains approximately 220 questions. Besides the requisite demographic data, the questionnaires included questions on:

- **Language background**, e.g. languages spoken by family members, as well as when and where English or (for the SI survey) Fiji Hindi was learned.
- **Language use** in different domains, for different purposes or topics, with different interlocutors, e.g. writing a letter to a relative, friend or for business; discussing work, religion, or politics with a boss, subordinate, co-worker, a customer or a teacher; and the language(s) used in the home with various family members.
- **Language attitudes**, e.g. whether a particular language should be taught in school, or used more extensively on the radio or television; whether a particular language is considered beautiful or pleasing to the ear; which language is regarded as the most useful; whether code-switching is acceptable; and attitudes towards the respondents' own competency in the various languages they know.

3.2.3 PROCEDURE

Questionnaires were administered over two weeks in July 1993 by a team of assistants consisting of 22 University of the South Pacific linguistics students and one lecturer. Twelve assistants worked on the Suva survey and eleven on the SI survey, each being assigned to administer 42 questionnaires. The assistants were matched with the people they interviewed in terms of ethnicity and, whenever possible, sex.[7] Thus, Fijians interviewed Fijians and Indo-Fijians interviewed Indo-Fijians in the Suva area, whilst interviewers for the SI survey either were themselves descendants of South Indians or had close relatives who were. All interviewers worked in an area where they had close kinship ties. Interviewers and respondents, therefore, belonged to some of the same social networks, so that initial contact was natural, agreement to take part in the survey was probably more·forthcoming, and the authenticity of responses more likely.

Initial contact was made in the community language shared by interviewer and repondent (Fijian or Fiji Hindi). Once the person contacted agreed to participate in the survey, the questionnaire was administered verbally in the language chosen by the respondent (Fijian or English, Fiji Hindi or English) and the interviewer recorded answers on the appropriate printed version. Interviewers were monitored during the two weeks of the survey.

---

7 As there were more female than male interviewers, males interviewers interviewed *only* male respondents, while some female interviewers interviewed both female and male respondents

4. RESULTS

Not all 220 or so questions were applicable to every respondent. The detailed section on language use outside the home was subdivided into workplace, high school and tertiary institution, and normally only one category was applicable to a given respondent. Some respondents were unemployed and did not answer any question in that section. Others were illiterate in English, their vernacular or both, and questions on reading and writing did not apply to them. Thus the average number of responses obtained for each respondent was about 150.

The data set consists of three files, one for SI respondents, and two for Suva respondents—one for Fijians and one for Indo-Fijians. Separate data files were necessary because some of the questions asked each group were slightly different or dealt with different languages. Nonetheless there is a great deal of overlap, and most answers are comparable.

A full analysis of the data has yet to be completed, and in this paper we concentrate on a few salient results. The dependent (linguistic) variables we consider here are:

language use:

- language of questionnaire
- language use in the home
- incidence of code-switching

attitudes towards language:

- the most pleasing language
- the most useful language
- preferred language for answering questions about oneself
- is code-switching *good* or *bad*?

The independent (social) variables are:

- sex
- age
- income
- highest level of education reached
- ethnicity (for the Suva survey)
- area of residence—urban/rural (for the SI survey)

Chi-square tests were administered to determine the level of statistical significance for all variables analysed. Preliminary analysis of responses obtained indicates no statistical significance for sex in relation to language use or attitudes. In the survey as a whole, income and education level are closely linked, but education level is the more significant of the two. Similarly, in the SI part of the survey, urban versus rural residence and income are closely linked but, as predicted, area of residence is the more significant of the two.

4.1 LANGUAGE USE

4.1.1 LANGUAGE OF THE QUESTIONNAIRE

Since respondents could answer questions in their vernacular or in English, their choice became our first data point. Overall, a slight majority of respondents, 56.5%, answered in their vernacular (Fijians in Fijian, Indo-Fijians in Fiji Hindi). There are, however, interesting differences between the different sections of the population surveyed. In the Suva survey, the relationship between ethnic group and language of questionnaire is highly significant, with Fijians answering in the vernacular and Indo-Fijians in English in roughly the same proportions, so that results for both groups are almost a mirror image of each other (Table 3).[8]

TABLE 3: LANGUAGE OF QUESTIONNAIRE BY ETHNIC GROUP
(Suva Survey)

Ethnicity	Vernacular	English
Fijians	58	44
Indo-Fijians	44	56

(Chi-square = 8.65, df = 1, p = .0033)

In the SI survey a slightly higher percentage of respondents chose the vernacular, 62%. However, there is a highly significant difference between rural and urban dwellers, resulting again in a near mirror image, with over three quarters of rural dwellers answering in Fiji Hindi while 69% of urbanites answered in English (Table 4).

TABLE 4: LANGUAGE OF QUESTIONNAIRE BY AREA OF RESIDENCE
(SI survey)

Residence	Vernacular	English
Urban	31	69
Rural	77	23

(Chi-square = 95.10, df = 1, p = .0001)

The relationship between choice of language and age is also highly significant, so that the younger the respondents are, the more likely they are to answer the questionnaire in English (Table 5).

TABLE 5: LANGUAGE OF QUESTIONNAIRE BY AGE-GROUP

Age	Vernacular	English
15–25	41.6	68.4
30–45	51.5	48.5
50+	77.2	22.8

(Chi-square = 92.41, df = 2, p = .0001)

8 For ease of readability, all figures in the tables, starting with Table 3, represent percentages of responses. Chi-squares are calculated, however, on the basis of frequency counts

The relationship between choice of language and educational level is even more dramatic. The more highly educated respondents are, the more likely they are to answer the questionnaire in English. Nearly 90% of primary-educated respondents answered in the vernacular, compared to less than half of those with a secondary education, and only a quarter of those with a tertiary education (Table 6).

TABLE 6: LANGUAGE OF QUESTIONNAIRE BY EDUCATION LEVEL

Education	Vernacular	English
Primary	89	11
Secondary	42.5	57.5
Tertiary	24.4	75.6

(Chi-square = 144.07, df = 2, p = .0001)

These results may reflect both ability in, and attitude towards, English. Exposure to English takes place primarily in the classroom, where it is the official medium of instruction, and therefore ability in English is normally lower the less educated one is. English is also considered prestigious, particularly for anything regarded as having some official aura, which answering a questionnaire may well have for the respondents.

4.1.2 LANGUAGE USE AT HOME

The questionnaire section on language use covers two broad domains, home and work (or school, for younger respondents). The home section has a series of questions about the respondent's choice of language when addressing various interlocutors: 'What language did you mostly use at home when you last talked with your [parents, husband/wife, children, brothers and/or sisters, other relatives, friends and visitors]?' Although responses from each community vary in some respects, common trends emerge. The most important is that in all cases the predominance of vernaculars is overwhelming, with an overall average of over 96% (Table 7).

TABLE 7: LANGUAGE USE AT HOME: VERNACULAR VERSUS ENGLISH

Survey	Vernacular	English
Suva Survey	93.8	6.2
SI Survey	98.7	1.3

(Chi-square = 3.33, df = 1, p = .0682)

The difference between the Suva and the SI parts of the survey is not highly significant. In the Suva survey, figures for both ethnic groups are very high, over 90% (Table 8). The slightly greater percentage of use of the vernacular among Indo-Fijians than among Fijians is not highly significant.

TABLE 8: LANGUAGE USE AT HOME: VERNACULAR VERSUS ENGLISH
(Suva Survey)

Ethnicity	Vernacular	English
Fijians	91.9	8.1
Indo-Fijians	95.6	4.4

(Chi-square = 1.17, df = 1, p = .2798)

As for differences in language use according to interlocutor, in both communities the vernacular is most likely to be used with relatives, and English with friends (Tables 9 and 10). Percentages for these categories of interlocutors in the two groups are very close, with figures for English marginally higher among Fijians than among Indo-Fijians. On the other hand, percentages of use of English with children, and especially with spouses, are significantly higher among Fijians (p = .00749 and p = .0013, respectively).

TABLE 9: LANGUAGE USE AT HOME: FIJI HINDI VERSUS ENGLISH
(Suva Indo-Fijians)

Language	with relatives	with spouse	with parents	with children	with siblings	with friends
Fiji Hindi	98.4	98.2	97.9	97.1	94.3	87.6
English	1.6	1.8	2.1	2.9	5.7	12.4

(Note: Columns are ranked in order of percentage of Fiji Hindi usage.)

TABLE 10: LANGUAGE USE AT HOME: FIJIAN VERSUS ENGLISH
(Suva Fijians)

Language	with relatives	with parents	with siblings	with children	with spouse	with friends
Fijian	98	96.6	91.6	90	89.7	85.3
English	2	3.4	8.4	10	10.3	14.7

(Note: Columns are ranked in order of percentage of Fijian usage.)

Paul Geraghty (pers. comm.) considers the relatively high figures for English among Fijians unreliable. He argues that in fact few Fijians speak English at home (especially with their spouses, for instance), but that speaking English at home is commonly viewed as a sign of sophistication and education. The relatively high incidence of reported English use would be more likely then to be a reflection of this attitude than an accurate picture of behaviour. However, this purported high status of English by Fijians is not shown to the same extent in the responses to our questions on language attitudes (see §4.2). The discrepancies between the two communities in language choice with some interlocutors, particularly children and spouses, remain unexplained.

The figures for Fijian in Table 10 are in fact an aggregate of different responses ('Standard Fijian' and 'communalect'). Separate figures appear in Table 11, showing the great vitality of Fijian communalects, which have for most interlocutors a range of percentage of use only slightly lower than that of Standard Fijian (30%–50%). The communalects are spoken the

most to parents, siblings and relatives, less to spouses and children, and very little to friends. It is to be expected that the communalects should be most commonly spoken with kinfolk. In the city, on the other hand, friends are not necessarily from one's communalectal region, which explains the relatively high figures for both Standard Fijian and English. The figures for spouses may be explained by marriage between speakers of different communalects, and their consequent use of Standard Fijian as a lingua franca at home. The similar figures for children may indicate that, for those families, Standard Fijian may be, to some extent, displacing the communalects.

TABLE 11: LANGUAGE USE AT HOME: STANDARD FIJIAN, COMMUNALECT AND ENGLISH (Suva Fijians)

Language	with friends	with children	with spouse	with relatives	with parents	with siblings
Standard Fijian	72.2	58.6	57.6	46.8	38.8	38.3
Communalect	13.1	31.4	32.1	51.2	57.8	53.4
English	14.7	10	10.3	2	3.4	8.4

(Note: Columns are ranked in order of percentage of Standard Fijian usage.)

In the SI survey, the figures for the vernacular are even higher and the use of English negligible, probably because the people surveyed include both rural and urban dwellers while the Suva survey, by definition, has only city residents. Differences in language use between different interlocutors are much smaller, at 2% versus around 11% for the Suva survey (Table 12).

TABLE 12: LANGUAGE USE AT HOME: VERNACULAR VERSUS ENGLISH (SI Survey)

Language	with relatives	with spouse	with parents	with siblings	with friends	with children
Vernacular	99.4	99.2	99.1	98.7	95.5	97.4
English	0.6	0.8	0.9	1.3	1.5	2.6

(Note: Columns are ranked in order of percentage of vernacular usage.)

Among descendants of South Indians the overwhelming bulk of communication is reportedly conducted in Fiji Hindi, while the use of Dravidian languages is quite low, and that of English even lower (Table 13). The relationship between age and language use is obvious, as respondents tend to use a Dravidian language more often with their spouses and parents than with their children. The use of Dravidian languages with friends and visitors is naturally low as many of these are not necessarily descendants of South Indians themselves and have no knowledge of these languages. In fact the rankings by interlocutors for Fiji Hindi and Dravidian languages are a mirror image of each other.[9]

9 Note that there is not a single instance of reported use of Shudh Hindi.

TABLE 13: LANGUAGE USE AT HOME: FIJI HINDI, DRAVIDIAN LANGUAGE AND ENGLISH (SI Survey)

Language	with friends	with children	with relatives	with siblings	with spouse	with parents
Fiji Hindi	95.8	94.5	93.8	90.8	88.3	83.9
Dravidian language	2.7	2.9	5.6	7.9	10.9	15.2
English	1.5	2.6	0.6	1.3	0.8	0.9

(Note: Columns are ranked in order of percentage of Fiji Hindi usage.)

4.1.3 CODE-SWITCHING

Respondents were asked whether they switched languages during their last conversation with someone from their speech community. This kind of self-report may of course be as indicative of attitudes as of behaviour, if not more. (A later question probed attitudes towards code-switching directly, and results will be discussed in §4.2.4).

There were four possible answers: (1) 'all the time', (2) 'quite a lot', (3) 'not a lot' and (4) 'not at all'. The overall majority of respondents report some code-switching, but here again, there are interesting differences between sections of the community. In the Suva sample, nearly 79% report switching between English and their vernacular. But Table 14, in which answers have been combined into two categories (some switching, 'yes', versus no switching, 'no'), shows a very significant difference between the two ethnic groups, with Indo-Fijians reporting considerably more code-switching.

TABLE 14: CODE -SWITCHING (Suva Survey)

Ethnicity	Yes	No
Fijians	69.2	30.8
Indo-Fijians	88.5	11.5

(Chi-square = 25.68, df = 1, p = .00005)

Two factors may account for this higher figure among Indo-Fijians. One is that it indicates a more tolerant view of code-switching, a conclusion supported by the results of our direct question on attitude (see §4.2). The other is that 'code-switching' is often interpreted as occurring simply when a number of English borrowings are present in speech. Although this interpretation is common to both communities, the higher percentage of borrowings from English in Fiji Hindi may account for the higher reported incidence of code-switching among Indo-Fijians.

For the SI survey the question related to the incidence of switching between Fiji Hindi and a Dravidian language. This question of course does not apply to over half the respondents, who do not know the Dravidian language of their ancestors. Among those who do, 51.1% reported code-switching in their last conversation with a fellow 'South Indian'. More than a greater dislike for code-switching, this probably reflects the marginal competence of many in their Dravidian language. This explanation is supported by other results in the survey, which show for most respondents both a low self-evaluation of

competence and a drastically reduced use of Dravidian languages in most domains (see, for example, Table 13).

Another important finding about code-switching is the highly significant relationship between reported incidence and age, but the significance is different in the two halves of the survey. In Suva, the younger the respondents are, the more likely they are to report code-switching during a conversation with a member of the same speech community, although the difference between the over 50 and the under 45 is much greater than that between the young and the middle-aged (Table 15).

TABLE 15: CODE-SWITCHING BY AGE-GROUP
(Suva Survey)

Age	Yes	No
15–25	87	13
30–45	82.4	17.6
50+	64.9	35.1

(Chi-square = 25.3, df = 2, p = .00005)

Age is also significant in the SI survey, but here the relationship is more complex, with the middle group much more likely to code-switch than expected, and the older group relatively less (Table 16).

TABLE 16: CODE-SWITCHING BY AGE-GROUP
(SI Survey)

Age	Yes	No
15–25	48	52
30–45	69	31
50+	42	58

(Chi-square = 12.1, df = 2, p = .0023)

There is also a highly significant relationship between reported code-switching and education level in the Suva survey, so that the higher the level of education respondents have attained, the more likely they are to report switching (Table 17). This is not surprising, since English is the official medium of instruction and a higher level of education means more exposure to English and, normally, greater competence in it (see §4.1.1 and §4.1.2). Respondents who are more competent and more comfortable with English are more in a position to switch than respondents with less competence.

TABLE 17: CODE-SWITCHING BY EDUCATION LEVEL
(Suva Survey)

Education	Yes	No
Primary	52.7	47.3
Secondary	82.3	17.7
Tertiary	89.7	10.3

(Chi-square = 46.99, df = 2, p = .00005)

The same is not true for the SI survey, where the question was about switching between Fiji Hindi and a Dravidian language. The lack of relationship between reported code-switching and education level is explained by the fact that switching is between two languages which can both be considered 'vernacular', rather than between a vernacular and English, the official medium of education and a language widely perceived as prestigious in most official spheres.

4.2 ATTITUDES

4.2.1 MOST PLEASING LANGUAGE

Responses to the question 'Which language do you think is the most pleasing to the ear?' reveal that an overwhelming majority of the respondents (overall, 80%) have a much greater attachment to their own language(s) than to English. If the aesthetic evaluation of English is similar across groups, however, the evaluation of the other language varieties known to the respondents varies considerably between Suva's Fijians and Indo-Fijians, and between Suva's Indo-Fijians and the west's Indo-Fijians of South Indian descent.

In the Suva survey, the striking difference between Fijians and Indo-Fijians is in their evaluation of the standard variety and the vernacular. While Fijians rate their regional communalect higher than Standard Fijian, Indo-Fijians rate Shudh Hindi higher than Fiji Hindi (Table 18).

TABLE 18: MOST PLEASING LANGUAGE BY ETHNIC GROUP
(Suva Survey)

Ethnicity	Vernacular	Standard language	English
Fijians	60.3	23.4	16.3
Indo-Fijians	29.1	45.8	25.1

(Chi-square = 51.37, df = 2, p = .0001)

The vitality of the Fijian communalects has already been demonstrated in the section on language use (see Table 11), and the strength of speakers' attachment to them is obvious here. The importance of Standard Fijian both as a standard and as a lingua franca among Fijians from different regions does not lessen the status of those communalects, which is very high at the local level, since they are used for everyday informal communication and as an expression of regional solidarity, and also often for ceremonial purposes.

The high rating of Shudh Hindi among Indo-Fijians, on the other hand, can be explained by the extremely low status of Fiji Hindi. In spite of the very restricted domains of Shudh

Hindi in Fiji and the admitted low competence in it of most Indo-Fijians, the contempt Fiji Hindi is held in is nearly universal and its denigration the subject of many letters to the editor in the local press. Some deny Fiji Hindi's very existence, others consider it a bastardised form of standard Hindi, and any attempt at recognition is received with a storm of outraged protest.[10]

A comparison of results for Suva's Indo-Fijian respondents and the Indo-Fijians of South Indian descent reveals very different attitudes towards Shudh Hindi and Fiji Hindi. While Suva Indo-Fijian respondents rate Shudh Hindi highest, it is rated lowest by descendants of South Indians, who rate Fiji Hindi highest, and Dravidian languages in second place (Table 19).[11]

TABLE 19: MOST PLEASING LANGUAGE
(Suva Survey)

Fiji Hindi	Dravidian language	English	Shudh Hindi
45.6	20.8	17.3	16

Since most respondents in the SI survey either do not speak a Dravidian language at all or have marginal competence in it, this second place rating is quite high and indicates a strong attachment to languages which, in spite of their very restricted use, are still considered by many as a badge of group identity. The completely different evaluation of Shudh Hindi and Fiji Hindi by the two groups can also be explained. Tamil, Telugu and Malayalam belong to the Dravidian language family and are not related to Hindi, a member of the Indo-European family, and descendants of South Indians have not inherited other Indo-Fijians' prejudices against Fiji Hindi or their awe for Shudh Hindi. Descendants of South Indians, then, clearly have, as a group, different language attitudes, and this confirms that they have a distinct identity within the Indo-Fijian community.

In the survey overall, there is a strong correlation between language rating and two social variables, age and education level. Generally, the younger the respondent, the more likely he or she is to rate English highest. The rating of English also increases with the highest education level attained by respondents.

The relationship between age and the rating of English is very clear among Suva's Indo-Fijians, for example (Table 20). There is an even stronger correlation between age and the rating of Fiji Hindi: the older the respondent is, the more likely he or she is to rate Fiji Hindi highest. The rating of Shudh Hindi is more constant across age-groups.

10 One such case was when a Peace Corps volunteer compiled a dictionary of Fiji Hindi, using the Roman orthography, primarily to provide a resource for fellow volunteers needing to learn the language for their work (Hobbs 1985).

11 Urdu was named by one respondent.

TABLE 20: MOST PLEASING LANGUAGE BY AGE-GROUP
(Suva Indo-Fijians)

Age	Shudh Hindi	Fiji Hindi	English
15–25	43.5	21.2	35.3
30–45	49.4	29.4	21.2
50+	45.7	37	17.3

(Chi-square = 10.01, df = 4, p = .04)

While age is also correlated to language rating among Suva Fijians, the pattern is different. Although all age-groups rate their communalect highest, the older group rates English significantly lower than the other two groups (Table 21).

TABLE 21: MOST PLEASING LANGUAGE BY AGE-GROUP
(Suva Fijians)

Age	Standard Fijian	Communalect	English
15–25	23.8	57.1	19
30–45	20.2	57.1	22.6
50+	26.2	66.7	7.1

(Chi-square = 8.27, df = 4, p = .082)

Among descendants of South Indians, the significance of age is even stronger than among other Indo-Fijians, particularly with respect to their evaluation of Dravidian languages and of English. Dravidian languages are rated most highly by older respondents, while English is rated most highly by the young (Table 22). This is no doubt related to the fact that more of the older respondents know any Dravidian language, while very few of the young do.

TABLE 22: MOST PLEASING LANGUAGE BY AGE-GROUP
(SI survey)

Age	Shudh Hindi	Fiji Hindi	Dravidian language	English
15–25	14.2	44.4	4.9	36.4
30–45	20.7	56.8	13.6	9
50+	13.5	36.2	43.6	6.1

(Chi-square = 126.70, df = 6, p = .0001)[12]

The correlation between responses and educational level is also strong, particularly with respect to the rating of English. The overall tendency is that the rating of English increases with the educational level of respondents, not a surprising fact in a country where English is the official medium of instruction after the first three years. This tendency is very clear among Suva Fijians, for example (Table 23). As for the other language varieties, primary-educated respondents tend to rate their communalect significantly more highly than those with more education, which probably corresponds to their greater use of communalect. Standard Fijian, on the other hand, is rated highest by secondary-educated respondents, and lowest by

12 One respondent in the 50+ group (0.61%) named Urdu.

those educated up to tertiary level. The two groups give their communalect similar ratings, so the difference in their rating of Standard Fijian mirrors their different rating of English. Now, while exposure to English increases with the level of education because of its status as medium of instruction, there is also some de facto use of Fijian as a medium in many primary and secondary schools. In higher education, however, English is not just the official language of instruction, it is, de facto, the only medium.

TABLE 23: MOST PLEASING LANGUAGE BY EDUCATION LEVEL
(Suva Fijians)

Education	Standard Fijian	Communalect	English
Primary	22.9	75	2.1
Secondary	27.5	56.3	16.2
Tertiary	14.5	58.1	27.4

(Chi-square = 15.91, df = 4, p = .003)

The correlation between education level and attitude towards language is also very strong among Suva's Indo-Fijians (Table 24). If the pattern is not very clear when it comes to attitudes towards English, there are two obvious differences in the respondents' evaluation of the other two language varieties: primary-educated respondents are significantly less likely to name Fiji Hindi as the most pleasing language, while Shudh Hindi is named least by tertiary-educated respondents.

TABLE 24: MOST PLEASING LANGUAGE BY EDUCATION LEVEL
(Suva Indo-Fijians)

Education	Shudh Hindi	Fiji Hindi	English
Primary	34.1	11.4	54.5
Secondary	33.6	27.6	38.8
Tertiary	12.7	27.3	60

(Chi-square = 15.05, df = 4, p = .0046)

In the SI survey a number of respondents, particularly in rural areas, reported having had no formal education, hence the extra category, 'none' (Table 25). English is named 'most pleasing language' in increasing proportion to the education level of the respondent—although the difference between the secondary- and the tertiary-educated is not significant—while the reverse is true for Dravidian languages. The figures for Fiji Hindi, however, are very similar across education levels. So are those for Shudh Hindi, except that this language is named by significantly fewer respondents with no education, perhaps understandably since the major locus of exposure to Shudh Hindi is precisely the school.

TABLE 25: MOST PLEASING LANGUAGE BY EDUCATION LEVEL
(SI survey)

Education	Shudh Hindi	Fiji Hindi	Dravidian language	English
None	5.3	45.7	46.8	2.1
Primary	15.7	48.8	25.3	9.6
Secondary	21.2	43	5.7	30
Tertiary	18.5	44.4	11.1	25.9

(Chi-square = 102.54, df = 9, p = .0001)[13]

4.2.2 MOST USEFUL LANGUAGE

Responses to the question 'Which language do you think is the most useful?' yield a strong rating for English, a rating which is nearly identical among the three groups surveyed, averaging around 64% (Tables 26 and 27).

TABLE 26: MOST USEFUL LANGUAGE
(Suva Survey)

Ethnicity	Vernacular	Standard language	English
Fijians	63.9	18.1	18.1
Indo-Fijians	65.1	24.2	10.7

(Chi-square = 7.03, df = 2, p = .0297)

TABLE 27: MOST USEFUL LANGUAGE
(SI Survey)

English	Fiji Hindi	Dravidian language	Shudh Hindi
63.7	27.3	5.8	2.9

While Suva Fijians and Indo-Fijians rate English almost identically highly, Fijians are evenly divided as to the usefulness of Standard Fijian and the local communalects. Although Indo-Fijians acknowledge the usefulness of Fiji Hindi, Shudh Hindi's rating as most useful language by 11% of the respondents is relatively high and almost certainly more a reflection of its high status than of any objective usefulness. This is confirmed by the much lower—and probably more realistic—percentage of descendants of South Indians rating Shudh Hindi as most useful (2.9%). The low percentage of respondents rating a Dravidian language as most useful (5.8%) is also likely to be realistic, but contrasts markedly with the relatively high rating on the question on the most pleasing language (20.8%, see Table 19). In other words, descendants of South Indians recognise that the usefulness of the Dravidian language of their ancestors is now very circumscribed, but their attachment to that language nonetheless remains high.

13 Urdu was named by one primary-educated respondent (0.61%)

As in the case of responses to the question on the most pleasing language, there is a strong correlation between responses to the question on the most useful language and age and level of education. The correlation between responses and age is very clear overall, with the rating of English decreasing as the age of the respondent increases—and percentage ratings for English among the three age-groups are quite close across the three communities surveyed—while that of the other languages or language varieties, on the contrary, increases (Tables 28–30). Thus, among Suva Fijians English is rated as most useful by the young, while Standard Fijian and communalects are rated most useful by the old (Table 28). Similarly, Fiji Hindi is rated increasingly highly by Suva Indo-Fijians as they get older, while the correlation between age and the rating of Shudh Hindi is much weaker (Table 29). A similar pattern obtains in the SI survey, with Fiji Hindi and Dravidian languages rated increasingly more highly the older respondents are (Table 30).

TABLE 28: MOST USEFUL LANGUAGE BY AGE-GROUP
(Suva Fijians)

Age	Standard Fijian	Communalect	English
15-25	9.5	14.3	76.2
30-45	17.9	17.9	64.3
50+	27.2	22.2	50.6

(Chi-square = 12.77, df = 4, p = .0125)

TABLE 29: MOST USEFUL LANGUAGE BY AGE-GROUP
(Suva Indo-Fijians)

Age	Shudh Hindi	Fiji Hindi	English
15-25	9.1	11.8	78.8
30-45	9.4	20	70.6
50+	13.4	41.5	45.1

(Chi-square = 24.85, df = 4, p = .00005)

TABLE 30: MOST USEFUL LANGUAGE BY AGE-GROUP
(SI Survey)

Age	Shudh Hindi	Fiji Hindi	Dravidian language	English
15-25	3.1	19.6	1.2	76.1
30-45	3.9	23.4	3.9	68.8
50+	1.9	38.9	12.4	46.3

(Chi-square = 44.47, df = 6, p = .0001)

As was the case for the question on the most pleasing language, the correlation between responses to the question on the most useful language and education level is even stronger than that with age. At the same time the effect of education level is stronger for responses to this question on usefulness than to the one on most pleasing language. The correlation is in

the same direction, with the rating of English increasing with the highest level of education attained by the respondent (Tables 31–33). The rating of the respondent's vernacular is a mirror image of that of English, and decreases as the level of education rises. Among Suva Fijians, this is the case for both Standard Fijian and the communalects, both extensively used in this urban environment (Table 31), while among Indo-Fijians, it is the case for Fiji Hindi (Tables 32 and 33).

TABLE 31: MOST USEFUL LANGUAGE BY EDUCATION LEVEL
(Suva Fijians)

Education	Standard Fijian	Communalect	English
Primary	38.3	34	27.7
Secondary	15.7	15.7	68.6
Tertiary	8.1	11.3	80.6

(Chi-square = 36.01, df = 4, p = .00005)

TABLE 32: MOST USEFUL LANGUAGE BY EDUCATION LEVEL
(Suva Indo-Fijians)

Education	Shudh Hindi	Fiji Hindi	English
Primary	11.1	48.9	40
Secondary	12.5	21.7	65.8
Tertiary	5.5	10.9	83.6

(Chi-square = 24.88, df = 4, p = .00005)

TABLE 33: MOST USEFUL LANGUAGE BY EDUCATION LEVEL
(SI survey)

Education	Shudh Hindi	Fiji Hindi	Dravidian language	English
None	2.2	50.5	10.8	36.6
Primary	1.2	32.1	18.5	57.6
Secondary	4.7	15.5	2.1	77.7
Tertiary	3.6	3.6	0	92.9

(Chi-square = 72.61, df = 9, p = .0001)

It is interesting that among Suva Indo-Fijians, Shudh Hindi has quite a low rating among tertiary-educated respondents (Table 32). One would expect this highly educated group to have the most use for the prestigious, standard variety of their community language. But this is clearly not the case, and the fact that these are also the people who give English its highest rating indicates that there is no real competition between English and Shudh Hindi, and that the domains of use of Shudh Hindi must be very restricted. We interpret this result as confirming Siegel's observation (1973) that English has encroached on many formal domains in which Shudh Hindi was previously used.

At this point we would not expect the same ratings of Shudh Hindi among descendants of South Indians. As for their rating of Dravidian languages, there is quite a difference between respondents with at most a primary education, who tend to see some usefulness in them, and those better educated, who do not (Table 33).

4.2.3 PREFERRED LANGUAGE FOR ANSWERING QUESTIONS

Respondents were asked: 'What language would you prefer to use when answering questions about yourself?'. This was intended as a reliability check so that results could be compared with the language in which the respondent chose to answer the questionnaire. However, if we compare figures for English for these two items, it is clear that far fewer respondents said they prefer to answer questions about themselves in English than actually answered the questionnaire in that language (Table 34; see also §4.1.1).

TABLE 34: PREFERRED LANGUAGE FOR ANSWERING QUESTIONS VERSUS LANGUAGE OF QUESTIONNAIRE

Respondents	Preferred language = English	Language of questionnaire = English
Suva Fijians	26.3	42
Suva Indo-Fijians	38.2	56
SI respondents	22.5	38

(Chi-square = .66, df = 2, p = .72)

What accounts for this discrepancy, which occurs in all three parts of the survey? First of all, the options were different. To answer the questionnaire, respondents had to choose between two languages: Standard Fijian or English for Fijians, Fiji Hindi or English for Indo-Fijians (for both the Suva and the SI survey). But the choice was wider when respondents were asked about their preferred language for answering questions: Suva respondents could name English, their community's standard or the vernacular, and those in the SI survey could also select a Dravidian language (Tables 35 and 36).

TABLE 35: PREFERRED LANGUAGE FOR ANSWERING QUESTIONS BY ETHNIC GROUP (Suva Survey)

Ethnicity	Vernacular	Standard language	English
Fijians	33.9	39.8	26.3
Indo-Fijians	51	10.8	38.2

(Chi-square = 58.97, df = 2, p = .00001)

TABLE 36: PREFERRED LANGUAGE FOR ANSWERING QUESTIONS (SI Survey)

Fiji Hindi	English	Dravidian language	Shudh Hindi
64.9	22.5	10	2.5

Figures show, then, that most respondents prefer to answer questions about themselves in the language of their community. Among Fijians, the figures for Standard Fijian are slightly higher than those for the communalects, and English comes in third place, while Indo-Fijians overwhelmingly prefer Fiji Hindi over Shudh Hindi and English is in second place. A more detailed analysis of results to this question would recapitulate much of the information yielded by previous ones and is therefore not included here: the correlation between stated preference for a particular language and both age and education level, the correlation between language preference and area of residence in the SI survey, or the discrepancy between the evaluation of Shudh Hindi by Indo-Fijians of South Indian descent and other Indo-Fijians, for instance.

Besides the difference in options, there is another reason why answers to this question cannot be expected to tally with the respondents' choice of the language of the questionnaire. The phrasing 'questions about yourself' is in fact extraordinarily vague and respondents would have been entitled to request clarification: 'Do you mean questions about myself that have to do with my work, or my family, or with my interests, dislikes, beliefs?'. And while the question has a vaguely intimate connotation, it is likely that a questionnaire administered on behalf of two university lecturers has for most respondents, on the contrary, quite an official aura. Thus, while the answers are interesting in themselves, the question was ill-conceived as a reliability check.

4.2.4 CODE SWITCHING

We saw earlier that most respondents reported code-switching in their own speech (§4.1.3). Respondents were also asked: 'What do you feel about people who switch back and forth between languages or communalects during a conversation?'. Possible answers were (1) 'It's good', (2) 'It's bad' and (3) 'It doesn't matter'. In the tables below, neutral and positive responses have been combined and are labelled 'OK', while negative responses appear under 'bad'.

Results for the Suva survey show that more Fijians than Indo-Fijians have a negative attitude towards code-switching. Fijians are split almost right down the middle, while only 15% of Indo-Fijians consider code-switching bad, a highly significant difference between the two groups (Table 37). Results for the SI survey are very similar to those for Suva Indo-Fijians (Table 38).

TABLE 37: ATTITUDE TO CODE-SWITCHING BY ETHNIC GROUP
(Suva Survey)

Ethnicity	OK	Bad
Fijians	49.8	50.2
Indo-Fijians	84.9	15.1

(Chi-square = 68.99, df = 1, p = .00005)

TABLE 38: ATTITUDE TO CODE-SWITCHING
(SI Survey)

OK	Bad
80	20

Generally, these results go in the same direction as those on reported code-switching: more Indo-Fijians than Fijians report code-switching and find it acceptable. For both Suva groups the figures for reported code-switching are higher than those for a positive attitude towards it, but the difference is much greater among Fijians (Table 39).[14] In other words quite a number of Fijians (nearly 20%, or 1 in 5) 'admit' to code-switching, even though they consider it 'bad'.

TABLE 39: REPORTED CODE-SWITCHING AND ATTITUDE TO CODE-SWITCHING BY ETHNIC GROUP
(Suva Survey)

Ethnicity	Preported switching	Switching is OK
Fijians	87	13
Indo-Fijians	82.4	17.6

(Chi-square = 5.38, df = 1, p = .0203)

Since the data on the practice of code-switching are reported, rather than observed, and since code-switching is sometimes unconscious anyway, these figures are inherently less reliable as an indicator of actual use than of attitudes. To that extent the discrepancy between reported switching and positive attitude among Fijians is in itself indicative of a certain amount of unease about code-switching. As for Indo-Fijians, very few of them consider switching 'bad', and a negligible number of those 'admit' doing it anyway. Why this much greater acceptance of code-switching among Indo-Fijians?

We argue above (§4.1.3) that the common interpretation of code-switching in Fiji as including simply 'using a lot of borrowings', coupled with the higher percentage of borrowings from English in Fiji Hindi, may well account for the greater reported incidence of code-switching among Indo-Fijians. Another factor which may promote the tolerance to this 'code-switching' is the exposure, through films and videos, to varieties of Hindi from India which also feature borrowings from English, while they retain a much higher status than Fiji Hindi.[15]

5. CONCLUSION

The analysis presented here is preliminary, since only some of the results are reported. Moreover, since all data from the questionnaires are self-reported, a degree of caution is in order. Nonetheless, some basic facts emerge. First and foremost is the robust health of the vernaculars, in spite of the obvious impact of English.

---

14 This is not the case for the SI survey, where the much lower incidence of reported switching between Fiji Hindi and a Dravidian language (51.1%) is attributable to the respondents' marginal competence in and use of that Dravidian language.

15 Thanks to Veena Khan for this observation.

While many patterns of use and attitudes are broadly common to Fijians and Indo-Fijians, there are important differences. Both groups report very high use of their vernacular in the home, as well as a positive attitude towards them, while both consider English very useful. But figures for English are higher among Indo-Fijians on all questions but one: the questions on attitudes ('most pleasing', 'most useful', 'preferred' language) and the choice of language in which to answer the questionnaire. Figures are also higher among Indo-Fijians on code-switching (greater incidence and more positive attitude). The only question where figures for English are lower among Indo-Fijians is use in the home, but the difference with the Fijian group is not significant.

English is widely considered in Fiji to be indispensable to higher education and a good job. This is understandable because of the use of English as the main official language of government, business and the media, and also since perhaps the most prominent domain of English is precisely education. English is the official medium of instruction beyond the first three years of primary school and, in spite of the presence of code-switching in many classrooms, the main de facto medium up to the secondary level, and the sole medium at tertiary level.

While this is true for Fijians and Indo-Fijians alike, for the majority of Indo-Fijians education—and therefore English—is considered the *only* avenue towards advancement. This is a consequence of their precarious situation in Fiji, with uncertainties about the future including the viability of the sugar industry and in particular the security of land tenure, as well as constitutional issues and in particular equitable representation in Parliament. These uncertainties, aggravated by a sluggish economy and the threat of recurring calls for 'repatriation' to India by a small but vocal number of Fijian extremists, mean that migrating from Fiji is never very far from the minds of Indo-Fijians—even if it is beyond the ability of many. Most migration from Fiji is to neighbouring developed countries with which Fiji has a historical link, such as Australia and New Zealand, as well as two major countries in the Northern Hemisphere where English is also the main language, Canada and the USA. These countries are attractive, among other things, because their main language is English, and high skills in English in turn mean an enhanced ability to emigrate to these countries and better chances at successful adaptation. For those who cannot afford to migrate, education—and therefore English—is for the children the only passport away from the insecure piece of land whose lease to the family may or may not be renewed.

Does this particular importance of English to Indo-Fijians mean that they are shifting away from their first language? In spite of the strong effect of age on many responses, the very high figures for use of Fiji Hindi at home indicate that any shift is minimal. The impact of English, while very strong, seems limited to certain domains—a hypothesis which appears to be confirmed by a cursory look at results from additional sections of the survey. As for Shudh Hindi, in spite of its very high prestige and the equally low prestige of Fiji Hindi, its use is very restricted.

This is in sharp contrast with the Fijian speech community, in which both Standard Fijian and the communalects enjoy widespread use and positive attitudes. One might expect that, in the biggest urban area in Fiji, which continues to attract people from many different regions and where marriage between speakers of different varieties of Fijian is more common, Standard Fijian may be displacing the other communalects. There are some indications in our data on language use at home that this may be the case, but this will have to be confirmed by further analysis.

As for Dravidian languages, the generally low figures for use, particularly at home, and the fact that these figures get increasingly lower with each generation, show overall a precipitous decline, positive attitudes notwithstanding. These apparently contradictory trends are a reminder of the common but often overlooked difference between stated attitude and behaviour, which can be very striking.

A series of semi-structured interviews conducted among the older generation of respondents after the survey confirms that descendants of South Indians still value the Dravidian language of their ancestors as a symbol of a separate cultural identity within the Indo-Fijian community, think that it is very important to maintain that language, and want their children to know it. Unfortunately many of these same people report not speaking the language to their children or not speaking it at home at all (Mugler, this volume). Some parents go as far as blaming the demise of Dravidian languages on the TISI Sangam, a South Indian organisation and the only body engaged in language maintenance efforts, through the teaching of Tamil and Telugu in some of the schools they run. In turn, the Sangam's increasingly desperate pleas to parents indicate a dawning realisation of the limited impact of these efforts and of the crucial role of what Fishman (1989, 1991) calls "intergenerational transmission".

The loss of Dravidian languages has been a gain for Fiji Hindi, a trend which started as soon as the first indentured labourers recruited in South India arrived in Fiji. Fiji Hindi, already established on the plantations and the majority language of the Indo-Fijian community ever since, has fulfilled the same vernacular functions of informal communication at work and, increasingly, in the home, now the last—and crumbling—bastion of Dravidian languages. The now near-total replacement of Dravidian languages by Fiji Hindi, a language with very low prestige, is another sign of the health of Fiji Hindi. Shudh Hindi, which has very high prestige among other Indo-Fijians in spite of its limited domains, is even less used and much less prized by descendants of South Indians who historically have no cultural link with it. This distinct attitude towards Shudh Hindi, along with the lesser stigma they attach to Fiji Hindi, can be interpreted as a marker of a distinct cultural identity. As for English, potentially another prestigious competitor, descendants of South Indians do not use it to any greater degree than do other Indo-Fijians, in spite of the much greater linguistic dislocation of the history of their arrival in Fiji and the consequent greater vulnerability of Dravidian languages to language shift.

If the future of Dravidian languages in Fiji looks grim, that of the other vernaculars seems bright. Although people consider English very important, that does not mean they are giving up their first language. And if urbanisation brings greater use of English among a mixed population, English is not the only lingua franca. Preliminary results on a series of questions on the language used at work with a member of the other major ethnic group indicate a sizeable number of people communicating in the vernaculars: thus an Indo-Fijian and a Fijian may use Fiji Hindi or Fijian with each other, rather than English. On some questions, figures for vernacular use are as high as 20% of the sample. These figures are quite high, but they would very likely be much higher in rural areas such as the sugarcane belt, where many Fijians and Indo-Fijians work side by side, or in small islands where the few Indo-Fijians have learned the majority language, Fijian.

Finally, it is ironic—if perhaps inevitable—that the more 'educated' one is, the less one thinks of one's first language. This situation is not likely to change in the near future, given the current education system and the increasing perceived importance of English worldwide,

regardless of whether knowing English is really important to everyone in the country. One might hope that such initiatives as the development of programs in Fijian Studies and Hindi Studies at the University of the South Pacific may counteract this trend.

There has recently been concern worldwide about language endangerment and the often lethal impact of the spread of English and other 'world languages' on languages spoken by small numbers of speakers (see, for example, Hale et al. 1992). At the same time, there is a renewed and growing interest in the use of vernaculars, particularly in education, and many consider the maintenance of these languages a priority for a group's cultural identity and a matter of cultural rights. Any decision or even discussion about language planning and policy issues needs to be based on as clear a picture as possible of the actual use of these languages, the beliefs that their speakers hold about them, the trends in language shift and maintenance, and even the possibility of language death (Gal 1979; Dorian 1981). We hope that this preliminary study is a step in that direction.

REFERENCES

Adam, R.S., 1958, Social factors in second language learning. PhD dissertation, University of London.

Dorian, Nancy C., 1981, *Language death: the life cycle of a Scottish Gaelic dialect*. Philadelphia: University of Pennsylvania Press.

Fishman, Joshua A., 1989, *Language and ethnicity in minority sociolinguistic perspective*. Clevedon: Multilingual Matters.

1991, *Reversing language shift: theoretical and empirical foundations of assistance to threatened languages*. Clevedon: Multilingual Matters.

Gal, Susan, 1979, *Language shift: social determinants of linguistic change in bilingual Austria*. New York: Academic Press.

Geraghty, Paul A., 1983, *The history of the Fijian languages*. Honolulu: University of Hawai'i Press.

Hale, Ken, Michael Krauss, Lucille J. Watahomigie, Akira Y. Yamamoto, Colette Craig, LaVerne Masayesva Jeanne and Nora C. England, 1992, Endangered languages. *Language* 68/1:1–42.

Hobbs, Susan, 1985, *Fiji Hindi-English English-Fiji Hindi dictionary*. Suva: Ministry of Education.

Lal, Brij, 1983, *Girmitiyas: the origins of the Fiji Indians*. Canberra: Journal of Pacific History.

May, Thor, 1990, Language in Suva: language use and literacy in an urban Pacific community. Paper presented to the Australian Linguistics Society Conference, Macquarie University, Sydney.

Siegel, Jeff, 1973, A survey of language use in the Indian speech community in Fiji. Unpublished field study project for the Culture Learning Institute, East-West Centre, Honolulu, Hawaii.

1987, *Language contact in a plantation environment: a sociolinguistic history of Fiji*. Cambridge: Cambridge University Press.

White, Ronald V., 1971, Language use in a South Pacific urban community. *Anthropological Linguistics* 13/8: 361–384.

Mugler, F. and Tent, J. "Some aspects of language use and attitudes in Fiji". In Tent, J. and Mugler, F. editors, *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.*
C-141:109-134. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.109*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.

CONCORDANCE STUDIES OF THE LANGUAGE OF ADOLESCENT FIRST LANGUAGE TOK PISIN SPEAKERS

GEOFF P. SMITH

1. INTRODUCTION

Tok Pisin is one of the most extensively studied of all pidgin languages, and numerous detailed descriptions of its lexicon, grammar, historical development and sociolinguistic status are available. The raw material on which these analyses are based include written records of speech samples dating back to the mid-nineteenth century (Keesing 1988; Mühlhäusler 1978); corpora of recorded material based on research in a variety of regions (Romaine 1992; Mühlhäusler 1979; Sankoff 1980; Goulden 1990; Dutton 1973; Laycock 1970; Wurm 1971); written standards by expatriates based on close consultation with native speakers (Bible Society 1966; Bergmann 1982; Mihalic 1971, 1982); and native speaker insights (Thomas 1990; Nekitel 1994). Some of the data collected for specific studies have been subjected to further analysis by other writers (Smeall 1975; Woolford 1979).

While a wealth of data is available, one area which has not been adequately addressed is the degree of variation which currently exists and the extent to which such variation is 'acceptable' to the members of the language community. Some attempts have been made to establish norms of acceptability, for example the written standards referred to above, but detailed research into the parameters of acceptability is yet to be carried out. Similarly, in academic studies describing Tok Pisin grammatical structures, the criteria by which the correct is distinguished from the incorrect, or the acceptable from the unacceptable are not always clearly spelled out. A number of dimensions of variation can be identified: regional varieties (Laycock 1970; Wurm 1971; Mühlhäusler 1977, 1985), sociolects (Mühlhäusler 1975), distinctions based on age (Lynch 1979; Romaine 1984, 1990a; Sankoff & Laberge 1973), degrees of formality (Brash 1971; Mühlhäusler 1991; Smith 1990b), medium (Siegel 1983; Romaine 1990b) and relationships with other languages (Mosel 1980; Smith 1994).

This paper looks at evidence for regional variation between two roughly comparable groups of adolescent first-language Tok Pisin speakers from widely separated geographical areas. The evidence is based on an analysis of a small corpus of recorded speech, using the concordance programs Oxford *Microconcord* (Scott & Johns 1993) and Longman *Mini-concordancer* (Chandler 1989). Some comments on the usefulness of these concordance programs for this type of analysis will also be offered.

Jan Tent and France Mugler, eds *SICOL: proceedings of the Second International Conference on Oceanic Linguistics:* vol.1, *Language Contact,* 135–146.
Pacific Linguistics, C–141, 1998.
© Geoff P. Smith

Smith, GP. "Concordance studies of the language of adolescent first language Tok Pisin speakers". In Tent, J. and Mugler, F. editors, *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.* **C-141:135-146**. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.135*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.

2. THE STUDY SAMPLE

The texts used for this study were transcripts of recorded speech of adolescents in the Western Highlands (WH) and in New Ireland (NI) Provinces who spoke Tok Pisin as their first language. All the respondents were in secondary or upper primary education. The WH samples were recorded at Mt Hagen High School and Hagen Park High School in December 1990, while the NI samples were recorded at Utu Provincial High School and Carteret Community School, Kavieng, in August 1990. The respondents were asked to talk about personal experiences or tell a story based on either a modern or a traditional theme. The WH informants consisted of 12 males and 10 females ranging in age from 14 to 19 and the sample consists of just over 11,000 words. The NI sample is slightly larger, at almost 14,000 words, and was collected from 16 males and 8 females ranging in age from 12 to 16. These are very small samples by the standards of present-day corpora, but large enough to allow us to gain some interesting insights into some differences between the two.

3. CONCORDANCE PROCEDURE

Tape-recorded speech was transcribed using a broad phonetic transcription and a DOS text file was prepared for each of the WH and NI samples. These files were examined using the Oxford *Microconcord* and Longman *Mini-concordancer* (version 1.01) programs. Both these programs basically produce concordances of specified 'search words', that is, they identify all occurrences of the search word and display the line of text in which these words occur to show the context. Each line can be expanded to a larger context sample if desired. The Oxford program is useful for analysing extremely large amounts of text, and can simultaneously search through up to 500 different files from over 900 directories. However, it is limited to displaying a maximum of about 1,600 concordance lines, the exact number depending on other computer software used.

The concordance lines can be arranged according to alphabetical listing of the word immediately to the right or left of the search word, or two or three places from it. Searches can be conducted for groups of alternate words, which is useful when dealing with transcripts which contain various renderings of the same word, such as *dispela* or *disla.* Searches can also be limited to collocations with other context words at a specified distance or 'horizon' from the search word. Wild card characters can also be specified, for example **sa*** would include *sa, save* and any other word beginning with 'sa'. Appropriate use of wild cards can make data analysis somewhat less arduous.

The Longman *Mini-concordancer* can in addition produce word frequency lists, which can not be produced by *Microconcord.* It can also produce larger numbers of samples than the 1,600 limit of *Microconcord*, which is occasionally useful in examining the contexts of frequently occurring items (e.g. *i*). However, the size of the files that the Longman *Mini-concordancer* can handle is limited to about 50,000 words. The following word frequency lists were produced using the Longman *Mini-concordancer.*

4. ANALYSIS

4.1 WORD FREQUENCY

In the sample from WH, 879 different words occurred, whereas the NI sample produced 918 different words. The 25 most frequently occurring words, including variations such as

phonological reductions, are shown in Table 1. The words are arranged in decreasing order of frequency, and the number of occurrences is recorded.

There were altogether 6,466 occurrences of the first 25 words in the WH sample, constituting 57.3% of the total. In the NI sample, the figures were 8358 and 60.2% respectively. Of the commonest words, function words such as *i*, *long*, *na*, *ia*, *olsem*, *bin*, and *sa* are well represented, as are the commoner pronouns (*mi*, *mipela*, *yu*, *em*, *ol* and *tupela*). The commonest verbs are *go*, *kam*, *tok*, *kisim* and *stap*. The only word used exclusively as a noun is *man*, and there are two adjectives, *dispela* and *wanpela*, although these function primarily as demonstrative pronoun and article respectively.

TABLE 1: THE 25 MOST FREQUENTLY OCCURRING WORDS IN WH AND NI

WH (11,292 words)					NI (13,888 words)				
Word	No.	Freq. (1/n)	Rank	NI Rank	Word	No.	Freq. (1/n)	Rank	WH Rank
*na**	747	15	1	3	*i*	1157	12	1	6
em	652	17	2	4	*long**	715	19	2	3
*long**	617	18	3	2	*na**	645	22	3	1
go	466	24	4	5	*em*	635	22	4	2
ol	463	24	5	7	*go*	623	22	5	4
i	311	36	6	1	*nau*	596	23	6	9
mi	311	36	7	8	*ol*	541	26	7	5
*bilong**	310	36	8	9	*mi*	406	34	8	7
nau	245	46	9	6	*bilong**	348	40	9	8
*tupela**	230	49	10	19	*ia*	312	45	10	12
*stap**	228	50	11	26	*bin*	229	61	11	19
ia	220	51	12	10	*kam*	225	62	12	13
kam	215	53	13	12	*wanpela**	200	69	13	22
*mipela**	199	57	14	16	*kirap**	195	71	14	31
*dispela**	189	60	15	18	*tok*	188	74	15	17
*olsem**	168	67	16	24	*mipela**	182	76	16	14
*tok**	151	75	17	15	*taim*	163	85	17	20
man	150	75	18	21	*dispela**	156	89	18	15
bin	124	91	19	11	*tupela**	149	93	19	10
taim	97	116	20	17	*save/sa*	146	95	20	21
save/sa	92	123	21	20	*man*	132	105	21	18
*wanpela**	90	125	22	13	*kisim*	121	115	22	24
*tasol**	85	133	23	24	*yu*	116	120	23	26
*nogat**	80	141	24	65	*olsem**	113	123	24	16
*kisim**	74	153	25	22	*tasol**	113	123	24	23

* denotes the inclusion of phonological variants or reduced forms (e.g. *bilong* includes *bilong, blong, blo*, and *bl'*).

The first observation to be made from this list is an apparently remarkable uniformity between the two samples. The nine commonest words are identical in both samples, although the order of frequency is slightly different. The 25 commonest words in WH contain only two which are outside the commonest 25 in NI and vice versa.

In an earlier study (Smith 1994), the extensive borrowing or calquing of English expressions by young first language speakers of Tok Pisin in Manus was noted. Typically, English semantics were kept more or less intact, while the borrowings were fully integrated syntactically into Tok Pisin. This tendency is also seen to some extent in the present study, with integrated lexical items such as *spredaut* 'spread out' and *kolektim* 'to collect' (WH), and *inteviu* 'interview' and *ektivitis* 'activities' (NI) well represented in the data.

4.2 DIFFERENCES IN WORD FREQUENCY

In spite of the apparent uniformity of the frequency lists, one or two differences are evident. In some cases, it may be that the subject matter of the discourse in a small sample gives a misleading impression of the relative frequency of words. For example, the word *keiv* 'cave' occurs seven times in the WH sample, not at all in the NI sample. This does not mean that the word does not occur in NI, but that the topic of one of the stories from WH happened to involve caves. To get a good idea of the frequency of such words with any degree of accuracy, a very large corpus would be needed. (Even extremely large corpora, such as the Bank of English's 200 million words used in the COBUILD English dictionary can give misleading impressions of the frequency of words as used in everyday speech.) However, in other cases it does appear that words are more commonly used in one location than the other. For example, the word *gel* 'girl' or *gels* occurs 30 times in the WH sample but only once in the NI text. This is not due to the nature of the topic, as the alternative *meri* is used with approximately equal frequency in the two samples. *Nogat* 'no' has a frequency of 1/113 in WH (78 occurrences), but only 1/350 (30 occurrences) in NI. As far as less frequent items are concerned, the word *gatim* does not occur in the NI sample, although one occurrence can be found in the WH data, which does indicate that this form may coexist with *gat* in some areas. *Stap* occurs more than twice as frequently in WH than NI and *i* more than three times as frequently in NI than WH. The significance of these differences will be discussed more fully below.

In the latter cases, and other examples, it appears that the difference in frequency reflects a fundamental difference in the function of the word in question. This calls for a more detailed look at contexts and collocations as revealed by concordances, which may reveal differences in the way the words are used.

4.3 ILLUSTRATIVE EXAMPLES OF COLLOCATIONS AND WORDS IN CONTEXT

4.3.1 *nabaut*

As an example, the occurrences of the word *nabaut* (including the phonological variant *nambaut*) will be analysed. *Nabaut*, generally glossed as the adverbs 'around' or 'about', occurs 40 times in the WH sample, giving it a frequency of 1/221. In the NI sample, the frequency is less than one third of this (1/657). However, the concordances show that the difference is not merely a question of number of occurrences, but that a differentiation of function is involved. A concordance highlighting the words preceding the target word was

examined to see how often it qualified nouns or noun phrases and how often verbs or verb phrases. In both the NI and WH samples, the use of *nabaut* as an adverb following verbs occurred with roughly comparable frequency. The usage is illustrated by the example *ol i go daiv nabaut* 'they went and dived around'. However, the WH data shows that the majority of the occurrences of *nabaut* followed nouns or noun phrases (e.g. *nogat ol kostiums nabaut* 'there were no costumes and things'). The overall frequency of this distinctive use in NI was only about one sixth of its frequency in WH.

TABLE 2: USE OF *nabaut* IN NI AND WH

	WH	NI
Total	40	16
following noun phrase	28 (70%)	6 (37%)
following verb phrase	12 (30%)	10 (63%)

Thus, the use of *nabaut* following nouns in phrases like *ol sotpela man nabaut*, meaning something like 'and that sort of thing, et cetera' appears to be well established in the WH sample, but far less so in the NI sample. However, it is not always clear if the *nabaut* refers to the verb or a noun in the verb phrase, for example in *ol go painim kaikai nambaut*, the *nambaut* could qualify either the verb *painim* or the noun *kaikai*. This gives a clue as to how the usage could have arisen, from an adverbial expression qualifying verbs reinterpreted as a qualification of nouns in verb phrases and then of nouns alone.

4.3.2 *kirap*

The frequency tables also indicate that the word *kirap* and its phonologically reduced variants *kira* and *kra* occur considerably more frequently in NI (1/71) than in WH (1/213). The word *kirap* normally means 'get up' or 'start up'. Again, it can be shown that the difference in frequency reflects different usage. The use of *kirap* in context shows a very frequent collocation with the word *tok*, illustrating its use in reporting speech. Standard phrases such as *em i kirap na tok olsem*, usually considerably reduced (e.g. *em kra to se*), are widely used when quotations are introduced. A collocation of *kirap* and its phonological variants with *tok* 'say', using a (0, 3) horizon, was examined (i.e. occurrences of *tok* within three words to the right of *kirap*). This picks up phrases such as *kra tok*, *kirap na tok* and *kirap na i tok*. The concordance showed 78 occurrences of this collocation in NI (1/134), but only 28 in WH (1/316), less than half the frequency in the total speech sample. However, collocation of *kirap* and *tok* accounted for a greater percentage of the uses of *kirap* in WH than in NI. The phrase usually indicates the initiation of an opening move in a dialogue as opposed to a response to what someone else has said.

TABLE 3: USE OF *kirap* WITH *tok* IN NI AND WH

	WH	NI
total occurrences of *kirap*	53 (1/166)	195 (1/71)
collocation with *tok* (0, 3)	28 (53%)	78 (40%)

Thus, the use of *kirap* in these samples is considerably more common in NI than in WH, but of those occurrences in WH a greater proportion are used to qualify *tok* to indicate an opening move in a dialogue.

4.3.3 ADVERBIAL CLAUSES WITH *taim*

Dutton and Bourke (1990) first brought to our attention a use of *taim* 'time' or 'when' from the Southern Highlands in clause-final position when used as a conjunction in 'when'-clauses. It has since been confirmed that this pattern occurs not only in the Southern Highlands, but elsewhere in the Highlands region and it has even been reported from NI (1990:259). A concordance of the word *taim* can quickly determine how common this syntactic pattern is in the WH and NI samples. In the NI samples, the word *taim* occurred 163 times, 55 of which were as conjunctions in 'when'-clause constructions such as *taim mi kam long hia* 'when I came here'. The remaining uses of *taim* were mainly as nouns (e.g. *dispela taim, wanem taim?*). None of the instances of *taim* used as a conjunction was in clause-final position. In the WH sample, *taim* occurred 97 times. Of the 31 uses as a conjunction, *taim* occurred in clause-initial position in 25 or 81% of instances, but in a significant minority of cases (6 or 11%) it was in clause-final position as in *mi stap long greid nain taim, mi fil olsem breiv liklik* 'when I was in grade 9 I felt a bit braver'.

TABLE 4: POSITION OF *taim* IN WHEN-CLAUSES

	WH		NI	
	No.	%	No.	%
taim – total	97	100	163	100
taim – when-clauses	31	32	55	34
clause-initial	25	81	55	100
clause-final	6	11	0	0

Thus, it can be seen that *taim* is used mostly as a noun both in WH and NI, but of the 30% or so of occurrences as a conjunction, it appeared clause-finally in 11% of cases in WH, not at all in NI.

4.3.4 THE PARTICLE *i*

Debate about the exact function of the particle *i*, variously described as a predicate marker, subject referencing pronoun or resumptive pronoun (Keesing 1988) continues unabated (e.g. Franklin 1980; Keesing 1988; Mühlhäusler 1987, 1990; Romaine 1993; Verhaar 1991; Woolford 1979). What is remarkable here is the occurrence of *i* as the commonest word in NI and the sixth commonest in WH, especially considering accounts of the decline of the use of the particle (Sankoff 1980:269) and even predictions of its possible eventual extinction

(Romaine 1993:259). Although the particle is in the six most frequently occurring words in both lists, there is a significant difference between the two samples. To begin with, *i* occurs over three times more frequently in the NI sample with 8.3% of the total number of words compared with 2.8% in WH. But this is not merely an overall greater frequency in all contexts, and the pattern of use needs to be examined in more detail.

Concordances of the use of *i* in context were examined, displaying both the preceding word and following words listed alphabetically. A number of distinctive patterns became obvious from a cursory examination of these displays. When concordances were arranged with the preceding word sorted alphabetically, the number of occurrences of *ia i* was conspicuous in NI, 78 in all. The comparable figure for WH was only one. Similarly, there was a huge discrepancy in the frequency of occurrences of *na i*: 114 in NI compared with nine in WH. Comparison of occurrences of *nau i* was even more dramatic, with 104 in NI compared with zero in WH. When the word following *i* was sorted alphabetically, these concordances highlighted differences in the co-occurrence of *i* with *go* and *kam*, with 30% of occurrences in NI, but 42% in WH. More significantly, the collocation *i stap* accounted for 18% of the WH occurrences, but only 3% in NI.

TABLE 5: USE OF *i* IN WH AND NI

	WH		NI	
	No.	%	No.	%
i total	311 (1/36)	100	1,157 (1/12)	100
ia i	1	0.3	78	7
na i	9	3	114	10
nau i	0	0	104	9
i sa	2	0.6	25	2
i go/i kam	131	42	346	30
i stap	56	18	40	3
pronoun + *i*	102	33	373	32

These figures point to a striking difference in the use of *i* in discourse in the two samples. The preference for clause-initial use after *na* and *nau* represents a discourse style very rarely found in the WH sample. The very limited use of *i* in sequence with aspect and other particles such as *ia* is also noticeable. Nearly half of the WH use of *i* is in the constructions *i go* and *i kam*, including both use as directionals and also to indicate continuing action. In fact 60% of all occurrences of *i* in WH preceded only three words—*go*, *kam* and *stap*. A significant proportion of approximately one third of both samples consisted of *i* in its well-known role as marking predicates following subject pronouns *em*, *mipela*, *ol* and *tupela*.

4.3.5 RELATIVE CLAUSES WITH *longen*

In transcribing the texts, *longen* and its variants (*logen*, *loen*, *len*) were written as a single word, as *-en* does not have an independent existence. The usual meaning of *longen* is 'of it' or 'its' or 'to it' or 'to him' etc., as for example in the extract from NI: *mi kirap, mi lap longen nau...* 'I got up and laughed at him now...'. This is quite similar to the meaning of *long em*. Dutton (1973:39) draws the distinction between *em* and *-en* in terms of emphasis or focus. However, in the present data, *longen* appears to have an additional role in delimiting relative clauses in a significant number of cases. For example, in *wanpela ples i nogat man silip longen* 'a place where there was nobody sleeping', *longen* has anaphoric reference to *peles*, meaning 'there' or 'in it', but also has a role in delimiting the relative clause. This tendency is carried further in some instances, where the relative clause delimiting function appears primary, and *longen* does not appear to have any anaphoric reference:

(1) *I man Wabag ia em poisinim em longen ia wokobaut kam i go.*
It was the Wabag man who poisoned him approaching.

In the above example, *ia* bracketing as described by Sankoff and Brown (1976) also contributes to the definition of the relative clause, although this is not always the case:

(2) *...bihainim dispela ol liklik ston tupela bin tromei long rot longen.*
...followed these little stones which they had thrown on the road.

In this example, again there appears to be no anaphoric referent for *longen*. This clause delimiting function is rarely seen with *long em*, and it is much more common in WH than NI, as shown by the figures in Table 6.

TABLE 6: *longen* AND *long em* AS RELATIVE CLAUSE MARKERS IN NI AND WH

	WH		NI	
	No.	%	No.	%
total occurrences – *longen*	34 (1/332)	100	20 (1/694)	100
longen as relative clause boundary	23	68	5	25
total occurences – *long em*	14 (1/806)	100	13 (1/1,068)	100 100
long em as relative clause boundary	0	0	2	14

4.3.6 *stap* AND OTHER FEATURES

Stap is used to mean 'is situated', as in *diwai stap hapsaid long rot* 'the tree was on the other side of the road', and more generally 'to live' or 'to be'. It also follows verbs to indicate continuing action as in *tupela wok long kaikai i stap* 'they two were eating'. *Stap* occurs with a frequency about three times greater in the WH sample than in NI. Again a number of collocations show a difference in co-occurrence of *stap* with other words.

Table 7 shows that roughly equal percentages of occurrences of *stap* follow *i*. There is a much greater tendency for *stap* to follow pronouns directly in the WH sample, approximately three times the percentage of occurrences in NI. The collocation *stap + go* or *i go* is again much commoner in the WH sample. *Stap* preceding *na* or *nau* at boundaries in discourse

occurs more frequently in WH, whereas the collocation *sa stap* or *save stap* is much more frequent in the NI sample.

TABLE 7: *stap* IN NI AND WH

	WH		NI	
	No.	%	No.	%
stap – all occurrences	209 (1/54)	100	98 (1/142)	100
i stap	47	22	20	20
pronoun + *stap*	59	28	9	9
stap (i) go	24	11	4	4
stap na/nau	52	25	17	17
sa/save stap	11	5	16	16

As pointed out by Lynch (pers. comm.), there may be substrate motivation for the greater frequency of *stap* in WH, as a variety of existential verbs have been described in languages of the Highlands region (Piau 1981).

Similar analyses can be carried out with other grammatical features such as the marker *bai* and other auxiliaries. *Bai* and its phonological reductions *ba* and *b'* (the form *baimbai* does not occur in either sample) occur with approximately equal frequency in both samples. There has been considerable discussion of the tendency for *bai* to shift to preverbal position in creolised forms (Romaine 1990a). In the present samples the particle occurs pre-verbally in WH in over half of the occurrences (52%), while preverbal use accounts for only 21% of occurrences in the NI sample. Combinations of *bai* with other modal, aspectual or negative particles such as *bai mas*, *bai no* and *bai sa* do occur, albeit infrequently, in the NI sample but are not found at all in the WH sample.

Preliminary analysis also shows interesting differences in the use of *bin* with other verbs, such as *gat*, and in collocations with other aspect markers, such as *sa*. For example, *bin gat* occurs nine times in the NI sample, *sa gat* twice and *bin sa* eight times. None of these collocations is found in the WH sample. The combination and ordering of modal and aspect particles could be one area where regional differences may become apparent in the future.

5. DISCUSSION: THE ORIGINS OF SYNTAX IN DISCOURSE

Sankoff and Brown (1976), in an important paper on the development of syntactic structures from discourse features, show how the particle *ia* became grammaticalised in the development of relative clause delimitation in the Tok Pisin spoken by young people in Buang. Some of the data presented here support the existence of an analogous route for the origin or elaboration of syntactic structures.

In some cases it appears that features are open to various interpretations. In the case of *nabaut*, described above, there may be some ambiguity about whether the verb or the noun in a verb phrase is being qualified, which could be the origin of innovative syntactic properties. Similarly with *longen*, the original role in anaphoric reference could be reinterpreted where the word occurred to indicate relative clause marking where the word occurred at the end of a clause. The role of *i*, in addition to marking the boundary between subject and predicate and other functions in indicating directionals and continuous action, appears to include discourse

marking, at least in the NI sample. And *kirap* in discourse appears to have a special role with the word *tok*, namely in indicating the opening move of a dialogue. Could *kirap* develop further as an auxiliary verb or even an aspect marker in future generations?

6. CONCLUSION

A comparison of the samples of speech from WH and NI shows both a remarkable uniformity and some interesting differences. Although the size of the samples is small, it appears that the extent of dialect differentiation is strictly limited, and manifested statistically rather than as a set of distinct phonological, lexical and syntactic features. There seems to be a differentiation of the function of some words such as *nabaut*, *i*, *stap*, *taim* and *longen*, which suggests significant regional variation. It remains to be seen whether such differentiation will become more marked with time or be held in check by social and regional mobility and the influence of more conservative second language speakers (Smith 1990a). Nevertheless, regional differences must be taken into account before making generalisations about the direction of change in creolised Tok Pisin in Papua New Guinea as a whole.

Concordance programs such as Oxford *Microconcord* and Longman *Mini-concordancer* have proved to be useful tools for an initial analysis of lexical and syntactic features. Patterns of use and differences in the frequency of occurrence can be observed much more easily than by examining text or recorded data alone. However, the value of these programs lies mainly in identifying features of potential interest, which can then be subjected to a fuller analysis.

The importance of large corpora of written and spoken language is increasingly being realised in the English-speaking world (Aarts et al. 1993). Dictionaries and grammars based on the evidence of use in the real world of the language community are taking the place of more idealised systems derived from pedagogical tradition or the intuitions of the writer. In Papua New Guinea, a large national corpus based on the spoken language of people throughout the country would be a most useful resource for language planning purposes.

REFERENCES

Aarts, Jan, P. De Haan and N. Oostdijk, eds, 1993, *English language corpora: design, analysis and exploitation*. Amsterdam: Rodopi.

Bergmann, Ulrich, ed., 1982, *Save na mekim: buk bilong kain kain wok na kain kain samting bilong helpim sindaun bilong yumi*. Lae: Liklik Buk Information Centre.

Bible Society, 1966, *Niupela Testamen*. Port Moresby.

Brash, Elton, 1971, Tok Pilai, Tok Piksa na Tok Bokis: imaginative dimensions in Melanesian Pidgin. *Kivung* 4/1:12–20.

Chandler, Brian, 1989, *Longman Mini-concordancer version 1.01*. London: Longman.

Dutton, T.E., 1973, *Conversational New Guinea Pidgin*. *PL*, D–12.

Dutton Tom and R. Michael Bourke, 1990, *Taim* in Tok Pisin: an interesting variation in use from the Southern Highlands of Papua New Guinea. In Verhaar, ed. 1990:251–262.

Franklin, Karl J., 1980, The particles *i* and *na* in Tok Pisin. *Kivung* 12/2:134–144.

Goulden, Rick J., 1990, *The Melanesian content in Tok Pisin*. *PL*, B–104.

Keesing, Roger M., 1988, *Melanesian Pidgin and the Oceanic substrate*. Stanford: Stanford University Press.

Laycock, Don, 1970, *Materials in New Guinea Pidgin (Coastal and Lowlands)*. *PL*, D–5.

Lynch, John, 1979, Changes in Tok Pisin morphology. Paper presented at the Thirteenth Papua New Guinea Linguistic Society Congress, Port Moresby.

Mihalic, Francis, 1971, *The Jacaranda dictionary and grammar of Melanesian Pidgin*. Milton, Qld: Jacaranda Press.

1982, *Stail buk bilong Wantok Niuspepa*. Port Moresby: Word Publishing.

Mosel, Ulrike, 1980, *Tolai and Tok Pisin: the influence of the substratum on the development of New Guinea Pidgin*. *PL*, B–73.

Mühlhäusler, Peter, 1975, Sociolects in New Guinea Pidgin. In K.A. McElhanon, ed. *Tok Pisin i go we?*, 59–75. *Kivung* Special Publication No. 1. Port Moresby: Linguistic Society of Papua New Guinea.

1977, On regional dialects in New Guinea Pidgin. In S.A. Wurm, ed. *New Guinea area languages and language study*, vol. 3: *Language, culture, society, and the modern* world, 533–537. *PL*, C–40.

1978, Samoan Plantation Pidgin English and the origin of New Guinea Pidgin. *Papers in pidgin and creole linguistics* No. 1, 67–119. *PL*, A–54.

1979, *Growth and structure of the lexicon of New Guinea Pidgin*. *PL*, C–52.

1985, Variation in Tok Pisin. In S.A. Wurm and P. Mühlhäusler, eds *Handbook of Tok Pisin (New Guinea Pidgin)*, 233–273. *PL*, C–70.

1987, Tracing the predicate marker in the Pacific. *English World-Wide* 8/1:97–121.

1990, On the origins of the predicate marker in Tok Pisin. In Verhaar, ed. 1990:235–249.

1991, Watching girls pass by in Tok Pisin. In Jenny Cheshire, ed. *English around the world: sociolinguistic* perspectives, 637–646. Cambridge: Cambridge University Press.

Nekitel, Otto M., 1994, Review of Francis Byrne and John Holm, eds *Atlantic meets Pacific: a global view of pidginization and creolization. Language and Linguistics in Melanesia* 25/2:194–196.

Piau, Julie Ann, 1981, Kuman classificatory verbs. *Language and Linguistics in Melanesia* 13/1-2:3–31.

Romaine, Suzanne, 1984, Relative clauses in child language, pidgins and creoles. *Australian Journal of Linguistics* 4/2:257–281.

1990a, Change and variation in the use of *bai* in young children's creolized Tok Pisin in Morobe Province. In Verhaar, ed. 1990:187–203.

1990b, Pidgin English advertising. In Christopher Ricks and Leonard Michaels, eds *The state of the language*, 195–203. Berkeley: University of California Press.

1992, *Language, education and development: urban and rural Tok Pisin in Papua New Guinea*. Oxford: Clarendon Press.

1993, The decline of predicate marking in Tok Pisin. In Francis Byrne and John Holm, eds *Atlantic meets Pacific: a global view of pidginization and creolization*, 251–260. Amsterdam: Benjamins.

Sankoff, Gillian, 1980, *The social life of language*. Philadelphia: University of Pennsylvania Press.

Sankoff, Gillian and Penelope Brown, 1976, The origins of syntax in discourse: a case study of Tok Pisin relatives. *Language* 52/3:631–666.

Sankoff, Gillian and Suzanne Laberge, 1973, On the acquisition of native speakers by a language. *Kivung* 6/1:32–47.

Scott, Mike and Tim Johns, 1993, *Microconcord*. Oxford: Oxford University Press.

Siegel, Jeff, 1983, Media Tok Pisin. *Papers in pidgin and creole linguistics* No. 3, 81–92. *PL*, A–65.

Smeall, Christopher, 1975, A quantitative analysis of variation: i in Tok Pisin. In Cogen, Cathey et al., eds *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society*. Berkeley: Berkeley Linguistics Society.

Smith, Geoff P., 1990a, Creolised Tok Pisin: uniformity and variation. In R.B. Baldauf Jr and Allan Luke, eds *Language planning and education in Australasia and the South Pacific*, 197–209. Clevedon: Multilingual Matters.

1990b, Idiomatic Tok Pisin and referential adequacy. In Verhaar, ed. 1990:275–287.

1994, Husat i bin Othoraizim?: new verbs in Manus Tok Pisin. Paper presented to the Seventh International Conference of the Society for Pidgin and Creole Linguistics, Georgetown, Guyana, August 1994.

Thomas, Dicks R., 1990, A course in practical Tok Pisin. In Verhaar, ed. 1990:399–409.

Verharr, John W.M., ed., 1990, *Melanesian Pidgin and Tok Pisin.* Amsterdam: Benjamins.

1991, The function of i in Tok Pisin. *Journal of Pidgin and Creole Languages* 6/2:231–266.

Woolford, Ellen, 1979, Variation and change in the i 'predicate marker' of New Guinea Tok Pisin. *Papers in pidgin and creole linguistics* No. 2, 37–49. *PL*, A–57.

Wurm, S.A., ed., 1971, *New Guinea Highlands Pidgin: course materials. PL*, D–3.

Smith, GP. "Concordance studies of the language of adolescent first language Tok Pisin speakers". In Tent, J. and Mugler, F. editors, *SICOL: Proceedings of the Second International Conference on Oceanic Linguistics Vol.1, Language contact.*
C-141:135-146. Pacific Linguistics, The Australian National University, 1998. DOI:*10.15144/PL-C141.135*
©1998 Pacific Linguistics and/or the author(s). Online edition licensed 2015 **CC BY-SA 4.0**, with permission of PL. A **sealang.net/CRCL** initiative.