Creole/Superstrate Code-Switching : Analysing the Dynamic Relationship between Tok Pisin and English in Papua New Guinea

Adam Blaxter Paliwala

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Department of Linguistics

University of Sydney

August, 2012

Creole/Superstrate Code-Switching : Analysing the Dynamic Relationship between Tok Pisin and English in Papua New Guinea

Adam Blaxter Paliwala

Thesis submitted in fulfilment of the requirements for the degree of

Doctor of Philosophy

Department of Linguistics

University of Sydney

August, 2012

Declaration

I certify that this thesis is my original work towards the Degree of Doctor of Philosophy.

To the best of my knowledge, due reference has been made to all information sources and literature used in this thesis, as well as to any material previously published or written by another person.

The field research that allowed the collection of data discussed in this dissertation was approved by the National Research Institute of Papua New Guinea under an approved multiple entry Researcher/Academic VISA 10300447724.

This thesis does not incorporate without appropriate acknowledgments any material previously submitted for another degree.

Abstract

Investigating variation in Tok Pisin a generation after Papua New Guinean Independence has found bilingualism with English the most significant factor in the development of this pidgin/creole (Romaine 1992; Smith, G 2002).

Two predictions for the development of Tok Pisin can be traced to the pre-Independence debates of the 1970s: that of decreolization and the formation of a creole continuum on the one hand (Bickerton 1975); and that of sociolectal variation and code-switching (CS) with English on the other (Mühlhäusler 1975, 1985).

This study establishes and discusses reference-points for expanded (creole) Tok Pisin and for Papua New Guinea (PNG) English. Primary data collected during fieldwork from a range of urban, anglicised spoken registers are presented as containing three types of multi-lingual behaviour. The findings of different types of CS show: deliberate separation of languages for particular effects (Poplack 1980; Auer 1999); highly structured unmarked combinations featuring a matrix-language (Myers-Scotton 1993, 2002); and also intermediate language characterised by unmarked CS based on an intermediate, congruent, grammatical system (Muysken 1997, 2000).

Relating samples of the primary data to native bilingual speakers shows that tolerance of mixed-language is high, and that grammatical changes in 'Tok Pisin' involved in CS may not be salient for some speakers. This indicates a connecting route between bilingualism and language change.

This thesis shows that 'Urban Tok Pisin' can best be understood as 'Language Mixing' involving a full range of CS types. It is shown that there are indications of an intermediate, congruent grammar underlying some variable behaviour that has been elsewhere categorised as creating a 'continuum' with English. Both Tok Pisin and PNG English are involved in this relationship.

It is argued here that congruency-based CS underlies examples of 'decreolized' Tok Pisin, and shown that well-developed CS-analysis including PNG English is vital to any further study of variation in Tok Pisin.

Key words: Tok Pisin, Papua New Guinea, Code Switching, Language Contact, Pidgin & Creole, Sociolinguistics

Acknowledgements:

The research and production of this dissertation has been a long and, at times, painful process spanning three decades and bridging two millennia. I have many reasons to be profoundly grateful to my many friends in Papua New Guinea, Australia, America, Europe, and Britain for the love and support they have given my through all the years it has taken to complete this work. Life only goes in one direction, and while I may have made the journey unforgettable you all have made it unmissable. I would name you all, but simply say: if you are reading this, and think I mean you, I do.

Particularly worthy of thanks for their unwavering support, through thick and thin, are two outstanding supervisors, Jane Simpson and Michael Walsh. I would also like to thank my administrative supervisor Linda Barwick for her priceless assistance and detailed notes on later drafts, and the Department of Linguistics, the Administration of the Faculty of Arts, and the International Office at the University of Sydney for their support of my candidature.

I would also like to thank the Northcote Trust for the scholarship that made the PhD possible, the Britain-Australia Society for providing this opportunity to British and Australian students to work between these two countries, and Jim Robbins and the staff of the Papua New Guinea National Research Institute for their advice, assistance, and access to their archives.

Professionally, I would like to thank Jeff Siegel for his care and support as well as many valuable conversations, critical discussions, and comments on various drafts and conference papers over the years, Geoff Smith for always staying in touch and keeping it real, and the late Otto Nekitel, Ken Sumbuk, Dicks Thomas, and the staff at UPNG Language and Literature for allowing me to access their expertise and their students.

For sheer inspiration I thank Nic Gisbourne for introducing me to new ways to think about the language I use every day, James Simpson for encouraging me to think in old ways about the languages it used to be, Mike Dowman for convincing me that we are all incredible language-puzzling machines, and Kimie Takahashi for reminding me that we are always people too.

My work in PNG would have not been possible without the love and support of my second family: I thank you always for keeping a place at York Road for me. I would also like to give special thanks to the Sydney-PNG Wantoks Club for many fine afternoon BBQs and a list of family to stay with during my early wide-ranging field trips.

To my colleagues in the SPCL and the international community of creolists, and in particular Kofi Yakpo, Peter Bakker and Magnus Hüber for organizing some truly inspirational gatherings; to FRIAS and Peter Auer for supporting some great work on code-switching; to the MRS Committee of Lissant Boulton, Michael Scott, and Melissa Demian and all those who waited on the steps at the British Museum and tolerated a linguist in their midst; and in particular to Bob Borges, Rachel Hendery, Guillame Fon Sing, and Sally J Delgado for enjoying the pastries: thank you all for being my community while I put the finishing touches on this work.

From my years with Sydney Linguistics I would like to thank John Gibbons for his steadying tone, Barbara Horvarth, Scott Keesing, and Ingrid Piller for conversations on sociolinguistics, and my initial supervisor, Bill Foley, for pointing me to literature on code-switching and the MLF model. Thanks also to my fellow postgrad students at Sydney: my roommate Mami Hiraike Okawara, Nicoletta Romeo, Hilàrio de Sousa, Joanne Page, Carmel O-Shannessy, and St.John Skilton – thank you for the community and inspiration. Thanks too to Andrew Ingram for showing me the shelf in the library where the Tok Pisin books lived, and especially to Brett Baker for giving me the best advice I never took.

I also thank *nem*, Adam Philipson, for the 2Ad tour of the Highlands, for an unforgettable, impossible trek from Simbu to Madang, and for the thousands of hours of tunes.

For meeting me in both urban and rural jungles, Lila San Roche, Melissa Crowther, and Mark Donohue, and for visiting me in jungles of concrete and steel, Malcolm Rammage, Ian McClintock, and Gerrard O'Donnell, Rob & Sam McGahan, Mark & Emma Derrick Read, Phoebe Birks, Alissa Roberson-Masters, Graham & Webster Greenleaf, Jodie Gillespie, Bridget Egan, Ben Patterson, Will West, Agnes Lucy Ash-Bell, Javier Mitra-Valdes, Oliver & Gillian & David Victor-Smith, Moiz & Shakira Ginary, Nona Harvey, Ben & Gabi Duigu, and Mani Blaxter Paliwala.

I would also like to thank Emma Beard for her invaluable comments on the statistical analyses in Chapter 9.

Above all others, however, I would like to thank my mother and father for being the beginning, the middle, and the end of this project, and of so much of my life over this time. Words cannot describe the love and devotion you have shown me, and I hope you feel it worthwhile, after all.

Na mi laikim tok tenk yu lo ol lain manmeri na pikinini bilong Papua Niugini husat i bin givim liklik toktok na toksave bilong ol i go long disela liklik samting bilong mi.

Table of Contents:

Abstract i
Acknowledgements: ii
Table of Contents: v
List of Figures: x
List of Maps: x
List of Tables: xi
List of Abbreviations: xii

1 Introduction 1

2 Papua New Guinea Language Ecology and Tok Pisin 9

2.1 Geography and People 9
2.2 Colonial History 10
2.3 Social Change 12
2.4 Indigenous Language Diversity 15
2.5 Contact Languages 16
2.6 Language and Identity: diglossia in PNG? 20
2.7 Summary 24

3 Literature Review: 25

3.1 Part I : Pidgin and Creole Languages 26
3.1.1 Defining 'Pidgin' and 'Creole' Languages 27
3.1.2 The Creole 'Life-cycle' and Decreolization models 34
3.1.3 Forming Creole Continua 37
3.1.4 Development History of Tok Pisin 41
3.1.5 New Varieties in Tok Pisin 50
3.1.6 New Varieties in English : the PNG Dialect 54

3.2 Part II : Code-switching Research 57
3.2.1 Terminology and distinctions 58
3.2.1.1 Definitions 60
3.2.2 Code-switching vs. Variation approaches: 62
3.2.2.1 Linguistic Variables, Multilingualism, Creoles and Continua 65
3.2.3 Code-switching : 'Juxtaposition' and Borrowing 67
3.2.3.1 Bilingual behaviour, 'Interference', and Borrowing 67
3.2.3.2 Poplack : Constraints and Nonce Loans 71
3.2.3.3 Myers-Scotton : Matrix Languages and Integration 72
3.2.3.4 'Mixing' Languages : the Range of Bilingual behaviour 73
3.2.4 Code-switching and Social Meanings 77
3.2.4.1 Diglossia and Meaning 78
3.2.4.2 Meanings in CS : CA and the Markedness Model 79
3.2.4.3 Acts of Identity 81

3.3 Part III : An analytic Framework: CS typologies 83
3.3.1 Diglossia and Alternation 84
3.3.2 Dominant Languages: the MLF 87
3.3.3 Mixed Languages: Congruent Lexicalization 88
3.3.4 Conclusions : Different 'Types' of CS 90

3.3.5 A 'Continuum' of CS? 92
3.4 Part IV: Analytic Framework - examples 94
3.4.1 Alternation: Situational and Metaphorical 95
3.4.2 Insertion: Matrix Language 97
3.4.3 Grammatical Mixing : Congruent Lexicalization? 100
3.5 Summary and Questions 104

4 Project Design 108
4.1 Research Questions 110
4.2 Research Method 113
4.2.1 Research Design and Field Trips 114
4.2.2 Location and Selection of Speakers 118
4.2.3 Elicitation and Recording of Eata 122
4.2.4 The Dataset 125
4.2.4.1 Significance & Limitations of the Dataset 128
4.2.5 Data Processing Techniques 130
4.2.5.1 Transcription 131
4.2.5.2 CLAN Programs and coding 133
4.3 Presentation of Examples 136
4.4 Summary 136

5 Findings 1: Tok Pisin and English in Papua New Guinea 138
5.1 Phonology 141
5.1.1 Vowels 144
5.1.2 Consonants 156
5.1.2.1 Consonant Clusters 164
5.1.3 Phonological Reduction 168
5.1.4 Summary 170
5.2 Morphology 171
5.2.1 Verbs : marking Transitivity *-im* 172
5.2.2 Adjectives : *-pela* 175
5.2.3 Pronouns 178
5.2.4 Plurals 183
5.2.5 Conjunctions 186
5.2.6 Other Affixes in Corpora : expansion vs. code-switching 190
5.2.7 Summary 194
5.3 Morpho-syntax 195
5.3.1 Determiners 196
5.3.2 Possession 200
5.3.3 Copula 201
5.3.4 DO 202
5.3.5 Interrogatives 202
5.3.6 Negation 206
5.3.7 Predicate Marking 207
5.3.8 The Verb Phrase : Increasing Congruence 211
5.3.9 Prepositions 221
5.4 Lexis 226

5.4.1 Discourse Markers 227
5.5 Conclusions 228

6 Findings 2: Alternation between Languages 229
6.1 Examples of Alternation 231
6.2 Code Choice and Markedness 235
6.2.1 Code Choice in the Houses of Parliament 237
6.2.2 Discourse-patterning: Tag Switching and Alternation 243
6.2.3 A Place for Un-Marked Switching 245
6.3 Alternation: Expressive Functions of CS 247
6.3.1 Pragmatic Functions 247
6.3.1.1 Reiteration 248
6.3.1.2 Addressee/Topic 249
6.3.1.3 Interjections and Interrogation 250
6.3.1.4 Qualification 251
6.3.1.5 Demonstrative/Explicative 251
6.3.1.6 Declarative 253
6.3.1.7 Focus 254
6.3.1.8 New Information 256
6.3.2 Metaphorical Switching 258
6.3.2.1 Quotation 258
6.3.2.2 Topic-Comment 259
6.4 Alternation, Diglossia, and Community 261

7 Findings 3 : Insertion 263
7.1 Examples of Insertion 266
7.2 Borrowing and Expansion 269
7.2.1 Cultural Loans 272
7.2.2 Core Loans 277
7.2.3 Integration 282
7.2.3.1 Syntactical Integration 284
7.2.3.2 Morphological Integration 284
7.2.3.3 Phonological Integration 287
7.3 Insertion Analysis of 'Anglicised' Tok Pisin 294
7.3.1 Single Words 297
7.3.2 Adjective + Nouns 299
7.3.3 Two-word Verbs 302
7.3.4 Mixed Constituents vs. Nonce Loans 304
7.3.4.1 Double Marking Plurals 304
7.3.4.2 Double Marking Aspect 307
7.3.4.3 4M Model and L2 morphology 308
7.3.5 Larger Constituents 316
7.3.5.1 ML + EL Constituents 317
7.3.5.2 EL Islands 318
7.4 Conclusions 322

8 Findings 4: Complex Behaviour – the Prospect of Congruent Lexicalization ..323

8.1 Congruent Lexicalization and Variation 324
8.2 Examples of Tok Pisin/English Congruent Lexicalization 329
8.2.1 Homophonous Diamorphs 330
8.2.2 Morphological Integration 336
8.2.3 Multi-Constituent Code Mixing 339
8.2.4 Non-Constituent and Bi-Directional Switching 340
8.2.5 Back and Forth Switching 342
8.2.6 'Function Words': System Morphemes and a Congruent System 344
8.2.7 Violation of some Constraints on Code-switching 347
8.3 Congruent Lexicalization in a Creole/Superstrate pair 350

9 Findings 5 : Speaker's Perceptions – the UPNG Responses 354

9.1 Diglossia and Language Choice 355
9.2 Establishing Support for a Hierarchy of Salience 359
9.2.1 Method 360
9.2.1.1 Selection of speech samples 364
9.2.1.2 Informant Groups 371
9.2.1.3 Administering the Exercise 373
9.2.1.4 Data Analysis Techniques 374
9.2.2 Responses to the Listening Exercise 375
9.2.2.1 Quantitative Responses to the Listening Exercise: Frequency 376
9.2.2.2 Quantitative responses to the Listening Exercise: comparing the two student Groups 381
9.2.2.3 Quantitative responses to the listening exercise: scaling the samples384
9.2.2.4 Summary of Responses and Analysis 396
9.3 Perceptions of CS: Interpretations from the Listening Exercise 400
9.3.1 Extending the Listening Exercise 401
9.3.2 Conclusions from the Pilot Exercise 403

10 Conclusions 404

10.1 Findings on Previous Research 404
10.2 Project Findings : Conclusions 408
10.2.1 CS and Creole Continuum 410
10.2.2 CS and Borrowing 411
10.2.3 Types of CS in Tok Pisin 412
10.2.4 Additional Findings 412
10.2.4.1 Hierarchy of Salience 412
10.2.4.2 Mixed Language and Identity 413
10.3 Summary of the Research 414
10.4 Suggestions for Further Research 415

REFERENCES: 418

APPENDIX I : Audio Recordings 438

APPENDIX II : 'Core Corpus' Files and Length 443

APPENDIX III : CLAN Coding 444

Readme.txt....444
Special transcription and coding practices:....446
Transcription conventions:....448
CODING scheme: [cf. LIDES s4.2]....450
OTHER conventions:....451

APPENDIX IV : Example 'Core Corpus' Text – CLAN format....465

Filename: TiT#08_208 – Mnmuddy.cha....465

APPENDIX V : Software Used in this research....467

APPENDIX VI : Listening Exercise Introductory text....468

APPENDIX VII : Listening Exercise Samples....469

Sample 1 : Clan format example : single Alternation....470
Sample 2 (11) : Insertion : single words, Adj+N, V - integrated....471
Sample 3 : Insertions : single words and Adj+N....472
Sample 4 : Insertion : N, Adj+N, Discourse markers, numbers....473
Sample 5 : Alternation : back and forth....475
Sample 6 : Alternation : Back and forth; Insertion : Conj, Discourse....477
Sample 7 : Insertion : Islands, Pronouns, N - unintegrated....478
Sample 8 : Insertion : multiple : Adj+N, two word V, Islands, N - integrated....479
Sample 9 : Insertion : Adj+N, N, Islands, V - integrated....480
Sample 10 : CLx? : Alternation & Insertion : Prep, Det+N, Discourse....481
Sample 11 (2) : Insertion : single words, Adj+N, V - integrated....481
Sample 12 : CLx? : Insertions & Alternations : multiple : Adj+N, V, N, Conj, Pron, Adv, Islands....482
Sample 13 : Insertion : N, Conj, Discourse (Pron + Adv)....483
Sample 14 : Insertion : V....484
Sample 15 : Insertion : N, Discourse (Pron+V), V, Adv - Intergrated....484
Sample 16 : CLx? : Insertion : V, Adj+N, mixed VP, Conj, V – unintegrated, N....485
Sample 17 : CLx? : Alternation : mixed, Pron, Question; Insertion? Pronouns....485
Sample 18 : Insertion : Island (NP, predicate), Adj+N....486
Sample 19 : Insertion : N, Island (idiom)....486
Sample 20 : PNGE....487
Sample 21 : CLx? : Alternations : back and forth : flagged....487
Sample 22 : CLx? : Insertions & Alternations : back and forth - idioms....488
Sample 23 : CLx : back and forth...489
Sample 24 : Insertion : N, Island (idiom) - integrated....490
Sample 25 : CLx? : Insertion : N – unintegrated, N, Conj, Adv, Prep, Conj, Discourse 491
Sample 26 : Alternation : repetition and back....492

List of Figures:

Figure 2.1 : PNG Language Triangle 21

Figure 3.1 : National Census 2000 - Bi-literacy 47

Figure 5.1 : Tok Pisin Vowel Space (adapted from Laycock 1985: 296 (table 1)) 149

Figure 5.2 : PNG English Vowel Space (from A.-M.Smith 1986: 116) 150

Figure 5.3 : English Syntax - Verb Phrase Structure 211

Figure 5.4 : Tok Pisin syntax - Verb Phrase Structure (ref. Sankoff, G 1991 : 64) 212

Figure 5.5 : Tok Pisin Syntax - Verb Phrase Structure (ref. Romaine, 1992 : 277) 212

Figure 6.1 : Alternation - Structure (ref Muysken 2000: 7) 232

Figure 7.1 : Insertion - Structure (ref. Muysken 2000: 7) 266

Figure 8.1 : Congruent Lexicalization - Structure (ref. Muysken 2000: 8) 324

Figure 9.1 : Listening Exercise Responses : Bar Charts : 'most English' samples 377

Figure 9.2 : Listening Exercise Responses : Bar Charts : 'most Tok Pisin' samples 378

Figure 9.3 : Listening Exercise Responses : Bar Charts : (313) Insertion = Sample 9 378

Figure 9.4 : Listening Exercise Responses : Bar Charts : (312) Alternation = Sample 1 379

Figure 9.5 : Listening Responses Bar Charts : (318) back-and-forth Alternation = Sample 6 and (314) Congruent Lexicalization = Sample 22 379

Figure 9.6 : Listening Exercise Responses : Bar Charts : Insertions with ragged Alternation switching = Sample 21 380

Figure 9.7 : Samples Ordered by Median Score, showing IQR whiskers and Mode 386

Figure 9.8 : Samples Ordered by Median Score with Proportion Tok Pisin Words indicated 398

List of Maps:

Map 4.1: Papua New Guinea Field Sites (map adapted from Wikipedia) 119

List of Tables:

Table 3.1 : Types of Functionally Meaningful Alternation (ref. Gumperz (1982: 59) and Auer (1995: 120)) 86

Table 4.1 : Primary Data and the 'core corpus' 126

Table 5.1: Core Tok Pisin Phonology - five vowel set 145

Table 5.2 : Core Tok Pisin Phonology - three basic diphthongs 145

Table 5.3 : Tok Pisin Phonology - maximum vowel set illustrated by contrasts (ref. Laycock 1985: 302 & Mihalic 1971: 4) 146

Table 5.4 : Tok Pisin Phonology - vowels and vocalic nuclei (ref. Laycock 1985: table 1). 148

Table 5.5: Tok Pisin Phonology - sample non-standard diphthongs (ref. G.Smith 2002: 48)153

Table 5.6 : Core Tok Pisin Phonology - consonants (adapted from Mihalic 1971) 157

Table 5.7 : Core Tok Pisin Phonology - consonants (adapted from Laycock 1985 : table 1) 159

Table 5.8 : Tok Pisin Phonology - phonemic consonant clusters (adapted from Laycock 1985) 166

Table 5.9 : Tok Pisin Phonology – reduced forms of *dispela* in the core corpus 168

Table 5.10 : Tok Pisin Morphology - cardinal numbers 177

Table 5.11 : Tok Pisin Morphology - standard pronouns 179

Table 5.12 : Tok Pisin Morphology - conjunctions (compiled from Mihalic (1971), Verhaar (1995) and G.Smith (2002)) 187

Table 5.13 : Tok Pisin Syntax - preverbal markers in Tok Pisin and their meanings (following Romaine, 1992 : 277) 214

Table 7.1 : Insertion : examples of V-TR morphological integration in the core corpus 286

Table 9.1 : Table of numbered Listening Exercise samples with 'average' scores, language, and CS types 370

Table 9.2 : Table of numbered samples - ordered by Median score and quartiles 388

List of Abbreviations:

Abbreviation:	Meaning:
1,2,3	Pronoun: first, second, third person
Adj	Adjective
AdjP	Adjective Phrase
Adv	Adverb
Aux	Auxilliary
B	Borrowing
CV	Consonant-Vowel syllable structure
CM	Code-mixing
COMP	Complementizer
CONJ	Conjunction
COP	Copula
CS	Variously defined, used here as Cover term for 'code-switching', 'codeswitching', 'code switching', or 'code-mixing' : multilingual behaviour
DET	Determiner
DP	Determiner Phrase
EL	Embedded Language
EMP	Emphatic Particle
Eng	English
EXCL	Exclusive pronoun (for first person)
FEM	Feminine gender
FL	Fused Lect
INCL	Inclusive pronoun (for first person)
IQR	Interquartile Range: statistical measure of range: difference between 1st and 3rd Quartile figures
L1	First language
L2	Second language
LM	Language Mixing
LOC	Locative (preposition)
MASC	Masculine gender
ML	Matrix Language
Mod	Modal
MOMASE	Morobe, Madang, and Sepik
N	Noun
NEG	Negative/negator
NEU	Neutral gender
NP	Noun Phrase
NUM	Numeral
Obj	Object
PL	Plural marker
PNG	Papua New Guinea
PNGE	Papua New Guinea English
PM	Predicate marker
pose	Possessee
posr	Possessor
Poss	Possessive

PP	Preposition Phrase
PREP	Preposition
PROG	Progressive
Pron	Pronoun
Q	Question word
Qtag	Question tag or particle
REL	Relativizer
Q1	1st Quartile (calculated with median average)
Q3	3rd Quartile (calculated with median average)
sg/pl/dl/tl	Pronouns: singular/plural/dual/trial number
Subj	Subject
Tag	Tag
Temp	Temporal marker
TP	Tok Pisin
TR	Transitive (suffix)
V	Verb
VP	Verb Phrase
xxx	(in example transcripts : indicates an unclear section or deletion)
#	(in example transcripts : indicates a pause)

'Ektin Fensi'

Ektin fensi mi no bisi

Taim yupela kam na Ingilis lo mi

Olsem yu bilong narapela kantri

Na yu mekim olsem yu no bilo Pienji

Ektin fensi yu no gat sem

Yu giamanim usat turru nau yu daunim nem

Ingilis i go i go na het i pen

Mi paul turu taim yu mekim olsem

(Lyrics by O-Shen, featuring Ugly B. 2005)

1 Introduction

This study started from an interest in the relationship between social change and language change and a personal connection with Papua New Guinea (PNG). It was conceived as a sociolinguistic study of one of the world's 'new' creole and pidgin languages, as a survey of variation in the language of the post-Independence generation of Tok Pisin (TP) speakers and its relationship to a developing sense of national linguistic and cultural identity. I set out to investigate 'the state of Tok Pisin' a generation after Independence.

As a child born in Port Moresby in 1975, I was excited to conduct fieldwork in Papua New Guinea and collect new primary data on the way in which Tok Pisin was spoken in the early 21st Century. The inspiration for this study was the great social changes that had taken place during my lifetime, and the possibility that these were paralleled by changes in the 'National' language.

Building on existing studies on Tok Pisin, on pidgin/creole language development, and on sociolectal variation in Papua New Guinea the initial plan was to provide a contribution that combined information on the use and scope of Tok Pisin in different domains with a variationist analysis of the language system, with particular focus on the differences between 'Urban' and 'Rural' varieties and on their relationship to English.

However, throughout fieldwork I saw that, 25 years after Independence, Tok Pisin had a currency in almost every situation, from discussions of village affairs in council meetings, to political debates in the National Parliament, to discussions between pilots and air traffic controllers and water-cooler conversations in government offices and corporate law firms, and even in the interview of a Chief of Police on an English-language radio station. While there was general agreement that English was the appropriate language for international conversation, that it was the language of big business, education, and official government activity, time and again the evidence showed language mixing: with Tok Pisin and English

used together expressively, sometimes markedly, sometimes entirely unmarked and unremarked upon by speakers and listeners.

It was clear that any pre-existing notions of restricted domains/topics of use had to be revised, as did previous characterisations of a 'diglossic' relationship between these languages. Furthermore, only a detailed consideration of the interaction between Tok Pisin and English in PNG could attempt to address the variation in these languages.

Different positions on the unfolding dynamic between Tok Pisin and English have been established. Bickerton (1975a) predicted this dynamic would result in a creole continuum: a range of variation in speech connecting the two languages, as had been observed in the Caribbean creoles. In the post-Independence period, however, the current relationship between English and Tok Pisin has been characterised as 'sociolectal' and involving 'code-switching' (CS) by Mühlhäusler and others (Mühlhäusler 1985d, 1985h; Siegel 1997; Smith, G. 2000, 2002; Siegel 2008).

This dissertation advances an understanding of the development of Tok Pisin through its interaction with the second largest single language spoken in Papua New Guinea: English. While on-going contact between a pidgin/creole language (Tok Pisin) and its lexifier (English) is of interest to creole language studies, the 'decreolization' model (s3.1.2) remains questionable, and the processes driving differentiation between language varieties and language change remain unexplored (s3.1.3). In considering the effects of Tok Pisin/English bilingualism on Tok Pisin as it is used in in-group conversations, media broadcasts, and public oratory and debate this research enables us to look in detail at variation in Tok Pisin in terms of its relationship with English.

A generation after Independence, it is clear that diglossic separation between these languages is not community-wide (s2.6), and that what has been described as code-switching (CS) behaviour is a significant feature of post-pidgin, 'Urban', and 'anglicised' varieties of Tok

Pisin (Mühlhäusler, Dutton and Romaine 2003; De Groots 2008). Consideration of current models of bilingual behaviour (s3.2-s3.3) suggests the view that code-switching amounts to more than just "random mutual interference" (Bickerton 1973: 641). Consequently, we may revisit Bailey's (1971: 342) suggestion that language interaction creates creole continua.

Instead of assuming a continuum of variation, I decided to investigate whether 'Urban' Tok Pisin could be characterised in terms of a mixture between the stable 'Rural' Tok Pisin grammar and English by applying code-switching theory to natural language behaviour. If a continuum was being formed through language interaction, bilingual Tok Pisin/PNG English speakers would be the most likely candidates to produce a mesolectal variety.

Continuing data collection and many interviews in the field, I found that throughout PNG code-swtching takes place both as a default, as in many of the urban families I worked with, and as negotiation of the language of communication. Often conversations begun in one language could be heard to shift to another language in markets, at holiday resorts, in bars, and in many other situations where strangers came into contact with each other. In villages too, different languages were in constant, fluid use.

Speaking with the younger generation of English-educated Tok Pisin speakers it became clear that there was little feeling that language choice was particularly dependent on context: though certain situations such as job interviews, lectures, and supervisions were understood as requiring the use of standard English, it was the person they were speaking with who most determined the language they would use. Significantly, young adults were aware that education in English was the most significant factor influencing the way a person might speak Tok Pisin. English remained, for all the young people I spoke with, conceptually the language of the elite, the language that can be seen to be structuring divisions in rural, as well as urban, communities.

At the heart of this dissertation are detailed analyses of the linguistic features observed in a set of 22 core texts drawn from my field recordings. In the speech of Papua New Guineans, I heard a range of mixing behaviours from clear alternation between languages, to widespread borrowings, to inserted and integrated words and phrases. In addition there were occasions where the morphosyntactic 'frame' of people's speech sounded as though it was drawing heavily on English. The recordings of this naturally occurring speech in conversation, political debate, and radio broadcasts provided me with rich texts: a resource for the investigation of CS in anglicised Tok Pisin and its significant role in the development of the language.

Initially I had two intentions: firstly, to investigate whether different varieties of Tok Pisin could be described in terms of CS involving the Insertion or borrowing of English into a Tok Pisin morphosyntactic Matrix Language Frame (MLF) (Myers-Scotton 1993a); secondly, to assess whether variation in the morphosyntactic Tense-Mood-Aspect system of different Tok Pisin speakers could be shown to form a continuum, as in Caribbean creoles (DeCamp 1971; Bickerton 1973; Rickford 1979, 1987; Winford 1997). The data collected focused my attention on the first of these, and my analysis found the MLF model of CS inadequate to explain all the language I had recorded. Providing a detailed linguistic analysis of bilingual Tok Pisin/English behaviour became the primary objective of this project.

I found that no single formal model of CS could account for the range of behaviour exhibited by Tok Pisin/English bilinguals, which includes different types of code-switching found in bilingual communities around the world. I found that both Alternations under syntactic equivalence (Poplack 1980) and Insertions based on morphosyntactic integration (Myers-Scotton 1993a) were taking place. Furthermore, in addition to cases indicating significant maintenance of syntactic structures in language mixing, I heard behaviour best understood as emerging from a mixed or shared grammar (Congruent Lexicalization) (Muysken 1995, 2000)

arising in a community where language mixing is so commonplace and frequent that code-switching has become the 'unmarked code' (Myers-Scotton 1993b).

Combined, these models provide a description of the fabric of language against which social and expressive dynamics of bilingualism play out, and suggest the orderly rules underlying speakers' sense of the well-formedness, or otherwise, of multilingual production.

In an experimental assessment of the different types of CS I had identified, I found that native-speakers were not necessarily aware of the significant impacts language mixing had on the grammatical structure of Tok Pisin. This was fascinating because the wealth of research into the range of CS behaviour suggests that language mixing can lead to language change, that code-switching is connected to other kinds of language variation, and that it can, in fact, lead to new stable varieties of the languages involved.

Consequently, my initial desire to engage with the range of variation in Tok Pisin, to describe the state of the language, synchronically, today, first required an account of this crucial, additional, dimension: the impact of bilingual behaviours on what people called 'Tok Pisin'. This dissertation develops such an account, one that may inform discussions of the diachronic 'development' of Tok Pisin (Mühlhäusler 1985d), with new varieties characterised by unmarked CS potentially providing a bridge between the grammars of English and Tok Pisin and indicating that a continuum may yet emerge between them.

This dissertation explores these issues with reference to my original data and observations collected during four periods of field work in Papua New Guinea between 2000 and 2003. It finds that behind the National Census figures on multilingualism (National Statistical Office 2002a; Paliwala 2012) lies the full range of language contact behaviour observed in other communities.

In the process of addressing the question of code-switching as an account of new varieties of urban, mixed language in Papua New Guinea I developed a versatile, replicable method for

archiving, analysing, and comparing Tok Pisin/English data from PNG. The examples throughout this dissertation are drawn from a structured multi-layered database of analysed multi-lingual 'Urban' Tok Pisin speech. Documentation on the coding of this database is provided in the appendices.

Following an introduction to the language ecology of Papua New Guinea, with particular reference to the social, political, and linguistic changes consequent on European contact in the 19th century (Chapter 2), I review the existing research literature in four parts (Chapter 3): Part I considers research and theories on Pidgin and Creole languages, including work on Tok Pisin; Part II discusses approaches to bilingual behaviour and definitions and understandings of 'code-switching'; Part III focuses on the specific models and theories contributing to the analytic framework applied in this dissertation; Part IV illustrates the implementation of the analytic framework with reference to examples of CS behaviour from other bilingual communities and from previous research in Papua New Guinea.

Chapter 4 outlines the research questions, methods of field work, primary data of the 'core corpus', and data processing techniques, and is followed by five chapters presenting my findings.

Chapter 5 shows that the Papua New Guinean dialect of English (PNG English or PNGE) bears phonological resemblance to Tok Pisin, and that variability in structural requirements in different varieties of Tok Pisin and PNG English extends this congruence.

Chapter 6 shows that an analysis of code-switching using both these languages indicates expressive, marked, functional and contrastive inter-sentential Alternation.

Chapter 7 shows the more intimate integrated Insertional switching associated with borrowing and the incorporation of foreign language material in the morphosyntactic matrix of Tok Pisin.

Chapter 8 shows that there is evidence in the data of language containing the features associated with Muysken's Congruent Lexicalization, a type of code-switching that depends on a shared underlying grammar. This third type of code-switching behaviour indicates the existence of mesolectal grammars, and is strongly suggestive of the development of a possible continuum of language behaviour that connects Tok Pisin and English in Papua New Guinea.

Finally, in Chapter 9 I provide a general discussion of qualitative impressions by speakers and a detailed analysis of the salience of these different types of code-switching as assessed in an pilot exercise designed to gauge speakers' responses to naturally-occurring mixed language.

The analysis of this exercise addresses the question of whether highly-educated bilingual speakers judged mixed language more on the basis of the prominence of CS and the relative quantities of material from each language than on the basis of the type of material involved, regardless of consequent changes to the grammars of the languages. It also considers whether Alternation is more salient than Insertion, and whether Congruent Lexicalization featuring small amounts of grammatically active English morphology is less salient than the well-integrated Insertion of verbs.

The concluding chapter presents a summary of the findings, conclusions, and suggestions for further research (Chapter 10). It discusses a characterisation of 'Urban Pidgin' as LM (language mixing) (Auer 1999) featuring Congruent Lexicalization (Muysken 1997, 2000).

The appendices provide material supplementary to the dissertation. Appendix I provides a list of the primary data: audio recordings collected during fieldwork. Appendix II gives details of the 22 'core corpus' audio texts from which the examples in this dissertation are drawn. Appendix III reproduces the documentation associated with the database, and includes the transcription conventions used in the examples throughout this dissertation. Appendix IV provides an example of a fully-coded 'core corpus' text. Appendix V lists the software used in preparing the database. Appendix VI and Appendix VII relate to the listening exercise

described in Chapter 9: providing the text of the introductory briefing and of the 26 samples selected from the primary data.

This dissertation addresses whether different types of code-switching (CS) can be 'pinpointed' (Mühlhäusler 1985d: 48) in Tok Pisin today through linguistic analysis, and provides a preliminary investigation of whether particular mixtures are always salient for speakers. Where CS is identified, it discusses whether this means that Tok Pisin and English remain separate languages, whether new varieties of mixed language are becoming normative in PNG, and whether CS provides a mechanism for the development of a creole continuum.

The high amount of mixed language found in the corpus raises questions about whether CS actually explains the previously observed signs of 'decreolization' in Tok Pisin, and has consequences for future attempts to define variation in Tok Pisin and PNG English.

2 Papua New Guinea Language Ecology and Tok Pisin

This chapter presents information on the social history of Papua New Guinea (PNG), and in particular (s2.2-s2.3) the developments following contact with European powers in the 19th century. It describes the linguistic ecology (Mühlhäusler 1996) as it has developed in PNG (s2.4-s2.5), and the elements of the social, political, and linguistic space (s2.6) of the national community.

2.1 Geography and People

Papua New Guinea (PNG) is the Independent nation occupying the Eastern half of the island of New Guinea, part of the Australia-New Guinea continental plate currently separated from the Australian mainland by the narrow 150km wide Torres Strait. Inhabited by plants, marsupials, and birds from the Asia-Pacific region, including Australia (to which the southern part was connected by a land bridge during the last Ice Age) current archaeological evidence suggests human habitation for 40-50,000 years (Gillespie 2002; O'Connell, James F. and Allen 2004; 2007). People who landed, or migrated by land bridge, on the North and West coasts moved along the coast to the Huon Peninsula and inland to the extremely pleasant Highland Valleys by c15,000 years ago (Dorney 1990: 19). Some people settled on the smaller islands in the region, including what are now known as New Britain and New Ireland, The Solomon Islands, Vanuatu, and, further east, the islands out to Fiji. These Melanesian people were supplemented by additional waves of migrations, which included Asian people along the southern strait between New Guinea and Australia 6-8,000 years ago who are the genetic ancestors of Polynesian people (Soares, Rito, Trejaut *et al.* 2011). Traces of these waves of migration can be found in the broad split between the Austronesian and non-Austronesian (Papuan) language families spoken in the region. Surveys in the 20th Century found 830 indigenous languages spoken in the Eastern part of the island and 276 in the West (SIL).

2.2 Colonial History

Sean Dorney (1990: 26-31) provides a comprehensive history of contact between foreign nations and the peoples of Papua New Guinea. The Spanish explorer Ynigo Ortiz de Retez recorded the island of New Guinea on a map of 1545, and this may have been as much as thirty-five years after the first European contact. The last five hundred years have seen contact with Malays and Dutch explorers, who claimed the western part of the island as part of the Netherlands Indies colony in 1660. German plantation and missionary activity commenced in the 19th century, and the north-east was officially annexed by Germany in 1884, where it became part of a managed network of plantations with German Samoa. Shortly afterwards, Anglo-Australians (initially for the British Empire) claimed the south-east 'Papuan' region.

In political settlements following World War I, Australia became the sole administrator of the east, with German New Guinea being mandated by the League of Nations as an Australian administrative region, and British New Guinea, known as Papua since 1904, an External Territory of the Australian Commonwealth. Following the Second World War, in 1945 a single Australian administration was established to cover both the Australian Territory of Papua and the United Nations Trust Territory of New Guinea. The west became an Indonesian province in 1969. In 1975 the Independent State of Papua New Guinea was created from the two Australian-governed regions in the east of Papua and New Guinea. Dorney (1990) provides an excellent account of the political and social process to Independence, and beyond.

With the intervention of foreign nation states the tribal groups on the accessible exterior of the island found themselves in contact with cultures with very different kinds of social, political, material, and linguistic characteristics. One of the significant effects of the new technologies and ideas brought by these contacts of the last 150 years was the creation of new connections between regional groups of people already living on the island. In particular these occurred as

consequences of the plantation labour system, missions, and the introduction of new infrastructure under the German occupation (Mühlhäusler 1985a: 44-54). Reed presents an image of language as central to this process:

> Once the lads of any village are recruited, they mingle in the intimate proximity of labor compounds with other youths from all parts of the Territory. Here again pidgin acts as a socializing agency in breaking down the older patterns of tribal animosity. Friendships may arise between boys whose villages are hardly a generation removed from a state of perpetual feud. The possession of this common speech, limited though it be, can hardly be overestimated as a creative force in cultural change.
>
> (Reed 1943: 287)

With the Australian administration English became the administrative language of New Guinea for the first time. The New Guinea Campaign of the Second World War gave people across both Territories unprecedented exposure to foreign soldiers, attitudes and cultures, and the post-war Australian administration surveyed the eastern mainland and islands and experienced urban migration (Mühlhäusler 1985a: 62). Reed (1943: 285-287) highlights the role of young people in this process, and as this generation aged, so the nation grew up. In 1975 Papua New Guinea declared Independence from Australian rule.

Today, over thirty-five years later, region of origin, whether Islands or Highlands, Coastal, Papuan, or Sepik, remains, for many Papua New Guineans, the most important marker of personal identity within their country. As in many countries (Anderson 1983), a Papua New Guinean 'national identity' remains something that must be 'created' and 'sustained' (Narakobi 1983), and a subject of debate (Tavurvur 2008). Yet, in the sense of being a part of a nation with incredible cultural diversity, and in the 'national language' of Tok Pisin, many people in Papua New Guinea accept that it is possible to act, think, walk, and to talk 'Papua New Guinean'.

The following sections of this chapter explore the history and dimensions of this identity and its relationship to language. Issues of urban migration, National political culture and discourse, and language choices for education and within the National Constitution frame the specific social field of my research.

2.3 Social Change

The growth in a Papua New Guinean cultural consciousness has enlarged with political independence, but the new sense of 'National' identity (Gewertz and Errington 1999) is something that emerged from the contact period post-WWII, under the Australian Administration, and ultimately enshrined in the PNG National Identity Act 1971. Where plantation workers had gained an initial sense of regional diversity (Mühlhäusler 1985a, 51-2), and Mission stations had provided centres for education and cultural change, as well as promoting regional lingua francas (Mühlhäusler 1985a: 52), it was the outreach projects of government administrators and the education of Papua New Guineans in English (Mühlhäusler 1985a: 62) that established a generation of political pioneers who took the country to Independence in 1975 (Dorney 1990: 33-38).

Alongside this process of political integration and change came widespread migration of many people from rural to urban areas, resulting in significant expansion in the population of towns and cities. Where urban housing did not have sufficient capacity for the migrant population, settlements were established on the periphery of urban areas. As the capital Port Moresby grew in population, the local Motu-speaking Papuan people found their villages subsumed as enclaves, and a similar pattern was repeated elsewhere. Today, in every province, villages located close to urban centres (peri-urban villages) have experienced social and economic changes in parallel with towns and cities. While the population of the country remains overwhelmingly rural today (Laycock 1985c; National Statistical Office 1994, 2002a, 2002b) the distinctions between town life, particularly in the settlements, and village life, especially for those close to urban areas, are becoming increasingly blurred.

Political independence and economic prosperity in the late 1970s and early 1980s combined with active English-medium education programs to produce a new cultural elite (Romaine 1992: 86-88). By the 1990s, anthropologists Gewurtz and Errington (1999) observed that the

characteristic egalitarianism of traditional Melanesian communities was replaced in some areas by social hierarchies supported and propagated by the elitist attitudes of the English-educated and affluent middle-class participants in the money-economy.

Linguistic researchers have associated this cultural shift with a shift in linguistic attitudes from an unstratified diversity to one characterised by strata based around English as an elite language and the Papua New Guinean variety of Melanesian Pidgin English, Tok Pisin, as a language of national solidarity (Sankoff, G. 1980c: 10, 24; Romaine 1992: 20). Furthermore, different language-ecologies formed in urban and rural communities. Tok Pisin was introduced into village life by plantation workers and government officials as a contact language of authority and social mobility. In the urban areas, with English more present as the language of authority, power, and status Tok Pisin took on the role of solidarity between people who might not have another language in common (Sankoff, G. 1980c: 25-26).

Today, Tok Pisin operates throughout the country as the language of national solidarity, and is widely spoken by the younger generations of both sexes, where in the past it was characteristically a male language, used with outsiders while women would be multilingual in the languages of their community (Laycock 1966: 45). As Romaine notes:

> Part of the reason for Tok Pisin's success lies in the two opposing social identities it may index. On the one hand, it is a language of colonialism and oppression, while on the other, it signals solidarity forged among Melanesians sharing a common fate beyond the village on plantations and in towns. Its allegiances transcend local village ethnicities and linguistic groupings. Vernacular languages are associated with the village and with the traditional values of village life. While English is to some extent now replacing the function that Tok Pisin once had as a language of wider communication, Tok Pisin is still seen as a local language of solidarity. While originally acquired for practical purposes, it has now come to serve integrative functions.
>
> (Romaine 1992: 54)

In his study of Official use of Tok Pisin 1964-1973 Noel anticipated this process of change, in particular noting the significance of Tok Pisin for Papua New Guineans and PNG-based expatriates in marking their identity when overseas (Noel 1975: 80).

Wurm and Mühlhäusler observed this changed role of Tok Pisin in social mobility:

> In contrast to other pidgins, Tok Pisin has not remained a language of inequality, or one restricted to domains involving inequality, but has become a language for upward social mobility and participation in political and economic power. In recent years it has become a symbol of cultural and national identity for many.
>
> (Wurm and Mühlhäusler 1985b: 8)

Elsewhere Wurm contrasts the value of Tok Pisin in the National community with that ascribed to English by the élite and by the less-educated alike who he describes viewing English "as the sole key to wealth and economic success" (Wurm 1985: 71-2). Wurm asserts an alternative reality:

> In present-day Papua New Guinea, knowledge of English is almost totally irrelevant for many school leavers.
>
> (Wurm 1985: 72)

One explanation for this is provided by Romaine, who suggests:

> [The] educational system was originally designed to train an élite corps of administrators rather than to develop resources or skills at the community level.
>
> (Romaine 1992: 337)

Romaine's position is that this accounts for the cultural differentiation of school-leavers observed by Ahai (1984), and underlying the difficulties faced by a 20 year old man from the Finisterre Range:

> But I had after all finished school and therefore it was very difficult for me to just make a garden (to live as a planter), and if I were to go to the market in front of other people's eyes (to sell vegetables), then I would be ashamed, and this is how it is, and I and the young people in town of whom many have a school education, we wandered through town, we picked pockets, mugged people, things like that.
>
> (Keck 1993)

In Papua New Guinea today English remains the primary language of Education, though the Tok Ples Pre School program (Secretariat 1995) allows for traditional languages, or Tok Pisin, to be used in grades 1 & 2. While it may not be used on a daily basis by much of the population, English remains a significant part of life in most urban areas, and has a powerful role in shaping social distinctions within the country as a whole. It has also developed in

Papua New Guinea into a new communicative variety of English, with features in common with other World Englishes (s3.1.6) and also with Tok Pisin (Chapter 5).

While English remains the official language of education, of popular religion, government records, two of the national newspapers, the national television station, and many of the radio stations around the country, Tok Pisin has gained a post-Independence currency in all these areas. Pre-school and early years of education can be taught in Tok Pisin, and it remains at the periphery of school life for many children: used outside of formal classwork and in the playground. Political speeches, debates in the National Parliament, and public discourse in both rural and urban areas prominently features Tok Pisin. Many government publications are translated into this national language. The Bible has been carefully translated into Tok Pisin, and it remains a significant language for pastors and church services around the country. In the media, the weekly national Tok Pisin *Wantok* newspaper remains ever-popular, the national television station *Em TV* has regular news bulletins in Tok Pisin, and Tok Pisin is the language of popular music and comment on the radio. And, of course, Tok Pisin is the language in the mouths of daily interactions for millions of Papua New Guinean people: it is the language of the streets, the language of the marketplace, the language of the water cooler, and the language of parties, meetings, and celebrations for most urban people. It is the default language in multilingual settings, and it is celebrated as *the* National Language.

2.4 Indigenous Language Diversity

Much important work is done in Papua New Guinea (PNG) on the ancient indigenous ways of life inherited by its people from their ancestors, from before European contact. These cultures and languages are officially recognised in the Constitution of the country (at section 4(3)). The linguistic diversity of PNG has been variously commented on (for some discussions, see Sankoff, G. (1980d: 8-9); Foley (1986: 8-9); Mühlhäusler (1996: 22-41)). Time depth, geographic barriers, and cultural attitudes towards language as a "badge of identification"

(Romaine 1992: 26) have been offered as plausible explanations for the remarkable diversity. Most importantly, perhaps, and contrasting with the European experience of linguistic convergence essential factors such as standardised forms, centralised administration and control, education through schooling, and mass media culture were not present for any one language prior to the post-WWII Australian Administration (Romaine 1992: 26). Indeed, these factors have only really emerged on the National scale post-Independence.

Against this backdrop of diversity, there are accounts showing multilingualism was commonplace (Sankoff, G. 1980b; Romaine 1992: 56), though Laycock (1966: 4) does suggest that bilingualism increased after European contact, and that pre-contact multilingualism was an individual skill for inter-group communication, rather than a community norm. In particular, he identifies interlinguistic borrowing of words and culture concepts as a consequence of the new patterns of bilingualism with Tok Pisin (Laycock 1966: 45).

As cultural changes (Wurm and Mühlhäusler 1985b: 9) and social pressures distance people from their old ways of life many of these languages are under threat. Some excellent work has been done, by for example G.Sankoff (1980d), and by Kulick (1992), on the way in which pressure from one new language affects the language behaviour of traditional Papua New Guinean communities, and Mühlhäusler (1996) and Nettle & Romaine (2000) have drawn on ecological metaphors to describe this interaction.

2.5 Contact Languages

With the birth of the new nation came official recognition of two new languages developed from regional lingua francas, Tok Pisin and Hiri Motu, alongside an imported language, English. In the Constitution literacy in one of these three languages along with an indigenous language is an explicit goal (section 2 (11)), and competence in Tok Pisin or Hiri Motu was ranked alongside knowledge of an indigenous language *tok ples* or *ita eda tano gado* in

citizenship assessments. Today all three languages may be used in the National Parliament[1], in broadcasting, and in official publications. In terms of popular consciousness, however, Tok Pisin is the language of the nation, and it remains the preferred language of parliament (Brash 1975; Noel 1975; Wurm 1985: 71; Romaine 1992: 12; Nekitel 1998: 150-168; Paliwala 2004).

Characteristic of the new language culture is the personal history of Papua New Guinea's first National Prime Minister, Sir Michael Somare, who learned Tok Pisin at home with his family in New Ireland, then the community language of Kuanua, and only learned his 'own' language Murik at the age of 6 when his father went on leave to Korau (Romaine 1992: 7). Kulick (1992) revealed an even more-common pattern of language shift, where children growing up in their own village communities are primarily speakers of Tok Pisin and only passively bilingual in their village language. For those growing up in urban environments, lack of access to a community of traditional-language speakers further limits bilingualism with a *tok ples*.

The national political, social, and media landscape is dominated by English and Tok Pisin: the languages of the new culture that has emerged from international contact and the many aspects of social and material change.

However, in the early colonial period, Tok Pisin was not associated with an indigenous identity at all. Across the Pacific, three languages (or, perhaps more properly, groups of language varieties) are considered to be different varieties of Melanesian Pidgin English

---

[1] It is worth noting, however, that despite the official status of Hiri Motu, its use on the floor of the house is often met with vociferous facetious criticism by MPs from areas outside of the southern, Papuan, region on the basis that it is a *tok ples*, a local (village) language (see p.252, below).

(Melanesian Pidgin): Pijin in the Solomon Islands, Bislama in Vanuatu, and Tok Pisin in Papua New Guinea. If we take the general account of Tok Pisin having developed from a Melanesian Pidgin English used on plantations in the Pacific in the 19th century provided by Mühlhäusler (1985a), Keesing (1988) and others, then Melanesian Pidgin was a foreign language associated with the plantation setting. The Papua New Guinean language was isolated with the Samoan plantation language (Mühlhäusler 1978) from other regional varieties following the annexation of New Guinea by Germany in the 1880s. Tok Pisin developed as a distinctive variety in the plantation setting in German New Guinea, and was not widely spoken in the southern, Papuan region under Australian control. Part of this distinctiveness can been seen in the German component of the conservative Tok Pisin lexicon (Mühlhäusler 1979a). In New Guinea, German missionaries adopted a number of different local languages as regional languages in order to spread the Christian faith among multilingual Papua New Guineans. Dramatic social changes have taken place in PNG that have produced the new 'national' cultural and linguistic space, and this has become a space occupied predominantly by Tok Pisin.

Continuing in use as a contact language between German plantation owners, workers, and missionaries the stable plantation language was taken home by local people and taught to successive generations as a foreign language (Reed 1943: 285-286) vital for access to the new society. G.Sankoff was able to observe a crucial change in this attitude occurring in the 1970s:

> [F]or the past thirty years or so, Buang parents have seen the acquisition of Tok Pisin as perhaps the chief resource which will gain their children access to the riches of colonial society, and have been anxious that they learn it. … The last four or five years have shown a growing awareness that, on the contrary, Tok Pisin does not get you far enough in the new society, that it is not the white man's language after all. It is English that must be known.
>
> (Sankoff, G. 1980c: 23)

Tok Pisin had developed a life and a currency outside of the plantation context as language of wider communication for indigenous people, largely free of contact with English-speakers

until the end of World War I. The use of Tok Pisin by both the Allies and the Japanese in World War II reinforced its currency, yet the politics of its use by the Australian Administration after the war were questioned, with a suspicion that this pidginized English was being maintained in order to 'keep the natives in their place' (Mühlhäusler 1996: 90). By the 1970s, however, with access to English-language education, a shift in attitudes took place, and the ground was laid for Tok Pisin to become the dominant language of the national community, initially among the urban population, and then, post-Independence, spreading from the towns and cities, and the New Guinea region in particular, across the nation as a whole.

Detailed linguistic work, e.g. (Hall 1943, 1955; Laycock 1966; Mihalic 1971; Sankoff, G. 1977; Wurm 1977; Mühlhäusler 1979a; Sankoff, G. 1980d; Wurm 1980), firmly established Tok Pisin's credentials as an expanded language of full and fluent expressive capacity, and Mühlhäusler's (1975) description of different sociolects, from rudimentary 'bush pidgin' to an 'anglicised' 'urban pidgin' was particularly significant in establishing that 'Tok Masta', some native-English speakers' version of Tok Pisin, should not be considered a legitimate variety of the language, but rather an approximation of it. The papers collected in McElhanon (1975) decisively establish that by the time of Independence, Tok Pisin was a thoroughly nativized language which it was quite possible to speak 'badly', rather than a variety of English, and Verhaar's detailed analysis of the written language of a range of Tok Pisin publications (Verhaar 1995) has shown that a 'virtual standard' for written Tok Pisin has been effectively established.

The weight of Tok Pisin scholarship agrees that the language exists in a range of forms, styles, and registers across the full expanse of communicative functions[2]. It is a language of solidarity and national identity, as well as one stratified in a range of social, and potentially other, dimensions.

2.6 Language and Identity: diglossia in PNG?

Wurm & Mühlhäusler suggested a separation between the roles of different languages in PNG:

> It seems clear that Tok Pisin, English and the local vernaculars are in complementary distribution from the point of view of their communicative functions in present-day Papua New Guinea.
>
> (Wurm and Mühlhäusler 1985b: 8)

Wurm described the relationship between Tok Pisin and English in the developing urban National culture as distinct and diglossic while noting that Tok Pisin is taking over from some traditional languages:

> The rapid spreading of this subculture in recent years, predominantly in urban environments, but also to some considerable extent in some rural areas, has resulted in a dramatic increase of the functional role and geographical importance of Tok Pisin which is now in a diglossic relationship with English in urban setting, i.e. the two languages exist side by side, with Tok Pisin and English fulfilling mutually exclusive, specific social roles and functions. In several areas Tok Pisin has replaced, or is beginning to replace, the indigenous vernaculars on its way to creolisation
>
> (Wurm 1985: 67)

The description of pre-Independence relations between languages in PNG as 'diglossic' is commonplace in accounts from the 1970s and 1980s (Adler 1975; Bickerton 1975a; Laycock

---

[2] Though Walczyński was not able to find evidence of the 'metalinguistic' function in his corpus from *Wantok* newspaper (Walczyński 2012), where Tok Pisin is used in pre-school programs appropriate terminology has been developed for the teaching and learning of additional languages.

1985a; Mühlhäusler 1985a; Piau and Holzknecht 1985; Wurm 1985). Recently, however, Nick de Groots has made explicit the erosion of this separation; while English has a high-register use, overlap with Tok Pisin is now the norm:

> English is the principal language of education, administration and commerce. But English is not the most comfortable for the people – it can be strange and academic. For young and old, when dealing with village, family, or daily life matters, Tok Pisin is the language of choice. In tertiary education, for example, even though a lecture may be conducted in English, when the teacher and students discuss issues together, the language used is often Tok Pisin; or in a problem-solving or counselling situation, Tok Pisin is invariably the preferred language of communication.
>
> (De Groots 2008)

In her work during the early period, G.Sankoff indicated the conditional nature of diglossia when she explicitly equated a rural *tok ples*/Tok Pisin dichotomy with an emerging Tok Pisin/English pattern in urban communities, with English "as a symbolic marker of status in the urban society" (Sankoff, G. 1980c: 26). Mühlhäusler, Dutton & Romaine observed that the influence of English post-Independence made distinctions between urban and rural varieties of Tok Pisin less clear than they were around Independence, though de-centralised Provincial administration has led to *tok ples* pushing back Tok Pisin in some communities (Mühlhäusler, Dutton and Romaine 2003: 7). While functional separation between languages may occur in individual communities, these vary from place to place, and it is no longer possible to talk about strict diglossic separation between languages at the level of the national language community.

I characterise the previously separate functional spaces of Tok Pisin, English, and *tok ples* in PNG as a triangle of relations between languages, as in

Figure 2.1.

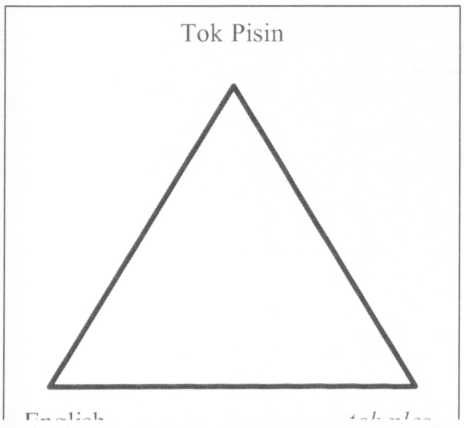


Figure 2.1 : PNG Language Triangle

Depending on the social and geographic setting, different dichotomies may be in place. For example, in the Milne-bay region, in the absence of a historical role

for Tok Pisin, the functional axis would be between English and *tok ples*. In rural communities in PNG, especially on the North Coast, the relevant axis would be between *tok ples* and Tok Pisin, while in an urban setting, the Tok Pisin/English axis comes into play. This can be seen in the use of Tok Pisin words in the village setting (Laycock 1966; Sankoff, G. 1980c, 1980a; Kulick 1992) and, in contrast, of English words as a marker of urban identity (Mühlhäusler 1985d: 140; Mühlhäusler, Dutton and Romaine 2003: 7).

However, evidence of changing social and functional roles for these languages, and the spread of Tok Pisin throughout the post-Independence nation, make the identification of any particular axis entirely context-specific.

This breaking up of distinctions between languages is directly referred to by Bickerton (1975a: 24), who predicted that social mobility and the growth of English-speaking elites would lead to the erosion of diglossia. This, coupled with imperfect acquisition of English and processes of social linguistic targeting leads, Bickerton hypothesised, to the creation of a linguistic continuum, a 'creole continuum', as found in the Caribbean:

> If that group habitually uses English in its everyday life, then he will attempt to acquire English. But the probability is that he will have had a much less extensive training in English than those he seeks to emulate; inevitably, therefore, what he speaks will be a kind of English heavily flavoured with Tok Pisin. This in turn will become the target of speakers a little lower down the social ladder, who will speak a variety of English yet more admixed with Tok Pisin. By this time, the accumulation of features which are neither truly English nor truly Tok Pisin will be so considerable as to form a sizeable part of the input to the language acquisition device of any child growing up in an urban area; and since the input mix of English, Tok Pisin and the hybrid varieties will vary proportionately for every child, it will not be long before we have an urban spectrum containing all linguistically possible varieties intermediate between Tok Pisin and English. The process will then filter back into the country along the lorry routes, and airlines, and the continuum situation will be complete.
>
> (Bickerton 1975a: 24)

However, describing Tok Pisin as a 'pidgin' with 'post-pidgin' characteristics, Mühlhäusler presented a contrasting hypothesis: explicitly pointing to the fact that we can identify 'code-switching' (CS) as evidence that a continuum was not developing:

> That code-switching can be pinpointed is an indication that one is not dealing with a post-pidgin continuum.
>
> (Mühlhäusler 1985d: 148)

In research on Papua New Guinea there is evidence of CS between Tok Pisin and indigenous vernaculars (Laycock 1966; Sankoff, G. 1980c; 1990; Kulick 1992; Stroud 1998), and between Tok Pisin and English (Mühlhäusler 1985d; Romaine 1992; Siegel 1997; Smith, G. 2002; Paliwala 2003), indicating that diglossia there is 'User-Orientated diglossia' under Britto's (1986) schema (s3.2.4.1), that is 'Fishman's Diglossia': a general, but not exclusive, association of particular languages with particular functions, domains, and social values. This goes along with features identified by Britto such as the variability of functional norms, with elites in PNG using English for conversation as well as formal purposes and with Tok Pisin being used in conversation by children in villages exclusive of the local language, and of acquisition through education, with both English and Tok Pisin commonly being learned in the school environment.

G.Smith (2002), describing Tok Pisin 'creole' varieties (see s3.1.1 for the issues around definitions of 'pidgin' and 'creole'), indicates that there are two potential directions for Tok Pisin and English in PNG, and potentially for *tok ples* where those languages are in consistent contact with Tok Pisin or English:

> The future direction of Tok Pisin depends on a number of variables, such as official language policies and education policies. Generally, there are two main possibilities, those of a post-creole continuum and English/Tok Pisin diglossia. If a post-creole continuum develops, there is likely to be more and more borrowing from English, including idiomatic expressions. If a diglossic situation becomes established, one would expect the development of a greater number of home-grown Tok Pisin idioms arising to fill gaps in new or existing semantic fields, or to increase expressive flexibility and range. Development in both directions simultaneously is also a distinct possibility.
>
> (Smith, G. 2002: 285)

With Tok Pisin being used extensively alongside English in schools and universities, in political debates (Nekitel 1998; Paliwala 2004) and in everyday exchanges, and with the substantial self-reported community of Tok Pisin/English multi-literates (National Statistical

Office 2002a; Paliwala 2011) (and see s3.1.4 below) the foundations for the erosion of diglossia are clear. However, under 'Fishman's diglossia', as identified by Britto (1986), and especially bearing in mind the different roles of Tok Pisin, English, and *tok ples* in different communities within PNG, G.Smith's proposal of "[d]evelopment in both directions simultaneously" (ibid.) appears likely.

Consequently, Mühlhäusler's (1985d: 148) proposition that the identification of CS indicates that the graduated variation of a continuum in Tok Pisin is not developing requires some further investigation. This provided a starting point for my research project.

This dissertation describes research into the relationship between Tok Pisin and English in terms of types of CS and the open question of a continuum between Tok Pisin and English.

2.7 Summary

This chapter has described the linguistic space of Papua New Guinea as well as the national context of language behaviour and linguistic debates concerning the on-going relations between Tok Pisin and English in Papua New Guinea.

Three key ideas have been identified as relevant: creolization as a social and linguistic process, diglossia, and the relationship between code-switching and a pidgin/creole to superstrate continuum. Chapter 3 examines the application of these concepts in linguistic research.

3 Literature Review:

This chapter is in four parts.

Part I of this chapter (s3.1) considers in depth the literature on pidgin and creole languages and Tok Pisin in particular. It highlights issues in the categorisation of languages as 'pidgin' or 'creole' and the position of Tok Pisin in this regard (s3.1.1), the concept of the creole 'life-cycle' model and its application to PNG (s3.1.2), and the crucial debate on the development of creole continua (s3.1.3). In the discussion of the history of the development of Tok Pisin (s3.1.4) and of new varieties (s3.1.5) I focus on the different positions taken by Bickerton (1975a) and Mühlhäusler (1985d: 148) (s2.6) on the developmental future of the language, and the continuing characterisations of the relationship between Tok Pisin and English as involving code-switching (CS) rather than a continuum by Siegel (1997, 2008), G.Smith (2000, 2002), and in an MA dissertation by Wakizaka (2008).

Part II (s3.2) addresses research and models of code-switching in order to provide a framework for the investigation of CS between Tok Pisin and English. It begins (s3.2.1) with a discussion of the terminological distinctions and categorisations in the field of research on bilingual behaviour, before discussing the relationship between CS and variationist analyses (s3.2.2). I consider the different traditions in CS research and their contributions to the understanding of different behaviour (s3.2.3) before discussing the literature on CS as an expressively meaningful technique for expressing identity relations in the language community (s3.2.4). I move beyond a single-model account of CS to frame here a study of different types of CS, one that throws new light on issues of decreolization and continua in Tok Pisin.

Part III collects the different strands of CS research to present an analytic framework, suitable to appreciating the greatest possible range of behaviour that may be considered as 'code-switching' (s3.3.1-s3.3.3). In addition to distinguishing three structural 'types' of behaviour

(s3.3.4), I discuss accounts of a hierarchy, or 'continuum' of types of CS that characterise the behaviour of different multilingual communities (s3.3.5).

Part IV operationalizes this framework with reference to examples drawn from previous research.

In this way, I introduce from the current literature on CS behaviour a set of categorisations which have not previously been applied to Tok Pisin and English in Papua New Guinea and which, in the light of the positions taken by Bickerton and Mühlhäusler, may provide new insights into the development of Tok Pisin.

3.1 Part I : Pidgin and Creole Languages

The study of pidgin and creole languages has developed since the 1950s from a marginal area of study associated with sociolinguistics (Bickerton 2010: 39) to become, by the late 1980s, a significant field in its own right. This reflected and encouraged changing social and political attitudes to the communities and contexts of these languages. Due to the crucial role played by Tok Pisin as a language of communication in Papua New Guinea during World War II, early negative attitudes to pidgin and creole languages (Mühlhäusler 1996: 90) as inferior varieties of colonial languages, or as languages without structure or value were called into question by Hall through his early grammar and later papers (Hall 1943, 1955, 1962). Studies of Caribbean creoles revealed systematic variation and presented evidence for behavioural 'creole' continua (see s1 and s3.1.3) within those language communities (Le Page 1960; Bickerton 1975b; Edwards 1980; Bickerton 1981; Edwards 1983; Rickford 1986, 1987; Winford 1992, 1993; Chaudenson 2001).

Following a number of conferences on pidgin and creole languages, by the late 1950s creole linguistics became established as a subfield of linguistic study (Cohen and Toninato 2010: 9). The 1968 Conference on the Pidginization and Creolization of Languages (Hymes 1968)

marked a watershed in the field of pidgin and creole studies, which has come to be considered relevant to issues of language development and change in historical linguistics, language contact and variation in sociolinguistics, language learning and first and second language acquisition, and to associated areas of psycholinguistics as well as to discussions of language universals. Bickerton's account of common features in creole grammars was similarly significant in opening up the field through a specific focus on tense, mood, and aspect systems (Bickerton 1981, 1984).

Today the field of pidgin and creole linguistics benefits from the insights provided by numerous research projects which have addressed and reframed germinal concepts and definitions of 'pidgin' and 'creole' languages, their 'life cycle', the impact of social and political factors on their historic and on-going developments, and the nature of creole continua. In this process of intellectual development the contributions of new generations of researchers who come from pidgin and creole language speaking communities have been vital.

I review the research literature below and relate (s3.1.4) the key concepts to Tok Pisin.

3.1.1 Defining 'Pidgin' and 'Creole' Languages

'Pidgin' languages typically arise in situations where people come into sustained contact with each other and have a communicative requirement without either a shared language or the opportunity to learn each other's language. In very brief contact situations, ad hoc jargons and sign language may be adequate. Even with longer contact, the languages that are developed tend to be structurally and functionally focussed on particular modes and situations: the specific requirements of the contact situation in which they are developed (Hymes 1971; Sankoff, G. and Laberge 1980 (1973): 196-197). Pidgin languages have been characterised as being developed when speakers use the words of one or more key languages with a grammatical system adapted from different sources (Ferguson 1971: 147-148; Hymes 1971:

81; Rickford 1977: 196; Traugott 1977: 73). Conceptually this is often seen to reflect the different power-relations in contact situations: speakers of the less powerful group adopting the lexicon of the more powerful group but basing the structure of their production on a form of their own native languages.

However, the process by which this takes place remains a matter of debate: the 'superstratist' position is that pidgins and creoles represent modified forms of their lexifier ('simplified', though that is a highly contested term (Siegel 2004: 140), or restructured); various 'substratist' positions consider the role of other languages spoken in the communities that develop pidgins and creoles in influencing the pidgin/creole structures: to a greater or lesser degree, they provide the cultural and linguistic background to the lexicon adapted from elsewhere. Whether either of these strong positions are appropriate to a particular contact situation, the concepts of 'superstrate' and 'substrate' inputs into the pidgin lexicon and grammar are dependent on the view of different inputs: from outsiders imposing a variety from the 'top down', to insiders adapting to new varieties from the 'bottom up'.

Bickerton's 'Bioprogram Hypothesis' (Bickerton 1984) represents a third, 'universalist', position. Rather than positing pidgin language development as a process of second language acquisition, he suggests that pidgin genesis (whether from the initial contact situation, as in earlier formulations of his theory, or from existing Original Contact Languages (OCL) (Bickerton 2010) as in more recent modifications) is shaped by the forces of first language acquisition: pidgin and then creole languages emerge rapidly from contact situations where a break in language acquisition leads to a reliance on a universal, innate, human language "bioprogram" to develop the structure of communication.

N.Smith (1994) gives an annotated list of pidgin, creole, and mixed languages from around the world. Commonly, pidgin languages were developed for use in situations of trade, and so information-sharing functions predominate, and such indigenous trade pidgins have been

described occurring throughout Papua New Guinea (Dutton 1983; Foley 1988; Williams 1993).[3]

In colonial contexts, pidgin languages have been significant languages of inter-cultural communication, and extended interaction between the different language groups using a pidgin is necessary for them to develop stable structures. However, without native speakers pidgin languages are dependent on the continuing social context in which they arose.

In Papua New Guinea the social history of contact combined with the great diversity of indigenous languages allowed Tok Pisin to survive and flourish outside of its original context of European cultural contact (s2). The result was an extended use of Tok Pisin as a language of inter-group communication and, eventually, its structural and expressive expansion and development to cover a full range of communicative functions (Mühlhäusler 1985a).

Commonly, such changes are associated with another social process: that of nativization, or the use of a pidgin as the first language of children growing up within the contact situation. While a 'pidgin' is, by definition, not speakers' first language, the term 'creole' is used to describe a language that is spoken natively by people from birth. As a consequence of a language becoming the primary resource for children to express themselves, creole languages have structures which support the fullest possible range of expressive and communicative functions, including emotional, social, and imaginative dimensions which may be

---

[3] However, see Winford on the formal classification of trade Motu and especially Pidgin Yimas as "somewhat simplified forms of Motu and Yimas respectively, only partially reduced so as to facilitate their use by non-native speakers in trading and other contacts with native speakers. Their degree of reduction is not nearly as extensive as that found in, say, Russenorsk. Hence, they should arguably be referred to as simplified languages, rather than pidgins." (Winford 2003: 21)

undeveloped by pidgin speakers who have other languages to fall back on (Hall 1962; Bickerton 1981; Thomason and Kaufman 1988).

Following from Schuchardt (1909), Hall (1962) drew support for the 'life-cycle model' of creole development from Bloomfield's definition of "creolized languages" as pidgins which have become "the only language of the subject group" (Bloomfield 1935). The 'life-cycle' view of pidgin/creole development proposed a pidgin precursor to every creole language, and placed the linguistic innovations of children growing up with a pidgin as their only language at the centre of theories about the development of the expanded, creole grammar. It was on this basis that Bickerton's 'Bioprogram Hypothesis' (Bickerton 1984) argued that similarities between creoles around the world were due to children's reliance on an innate human 'Universal Grammar' to develop structures that would support the additional communicative functions they needed to express: whereas a pidgin language focussed primarily on information-exchange, with adults able to fall back into their native languages for complex social and expressive functions, when the community shifted to using a pidgin language as its only means of communication the creole-speaking children had to develop clear ways to talk about the future, the past, their emotional states, and to streamline their language for social interaction. Bickerton's hypothesis was that the immediate requirements of monolingual communication led people to rely on the universal 'bioprogram' grammar.

Bickerton never included Tok Pisin in his hypothesis due to the fact that it did not develop abruptly. Subsequent research established that 'abrupt' creolization may be a rarity (Mufwene 2001: 130-140), and probably did not take place even in Bickerton's core case: Hawaiian Creole (Singler 1992, 1996; Siegel 2000; Roberts 2004; Singler 2006a, 2006b). Largely due to a disconfirmation of the posited facts of broken transmission, to the weight of empirical and theoretical observations that show connections between structures in pidgins and creoles and those in both superstrates and substrates, and to research into the social history of

communities which have developed these languages around the world, Bickerton's 'universalist' position is given very little credence in current Pidgin and Creole studies.

Thomason and Kaufman (1988) present evidence from various types of non-pidgin language contact situations that show more extreme restructuring than found in those creoles supposedly the result of Bioprogram effects. They provide a 'Borrowing Scale' that characterises the progression of language contact phenomena. Various authors have written on the debate over whether creolization has usually been a slow process or an abrupt one, and the significance of vernacularization in developing a stable, 'creolized', grammar (Goodman 1985; Lefebvre 1986; Mühlhäusler 1986; Mufwene 1988; Bickerton 1991, 1992; Mufwene 1994, 2000, 2001; Hudson, C.L. and Eigsti 2003).

Today, key debates on the origins of Pidgin and Creole languages revolve around investigating the substratist and superstratist positions, discussions of a proposed original, greatly reduced, pidgin precursor, and questions as to whether 'creole' languages can be considered a typological class. For a single summary of these concerns, see DeGraff (2003).

Working concurrently with Bickerton, G.Sankoff presented a view of Tok Pisin development which took the extended length of time it took to develop as central. G.Sankoff's work in Papua New Guinea presented an alternative view to Bickerton's proposal of 'abrupt' language sophistication by children working with a limited 'pidgin' input and a 'universal grammar' by demonstrating that in the extended process of adult-to-adult transmission of Tok Pisin the grammatical innovations associated with 'creole' development had been made by 'pidgin' speakers (Sankoff, G. 1977; Sankoff, G. and Laberge 1980 (1973): 193, 208-209). Reed (1943: 285-286) and Salisbury (1962) had also provided early accounts of Tok Pisin being taught and learned by adults in village communities. Following Slobin (1977), Degraff (1999a), Singler (2006a), and Siegel (2008) all discuss the different roles of adults as language innovators and children as language regulators. If the "bioprogram" effects were real

at all, they were taking place for adults as well as children, though Degraff, Singler, and Siegel each focussed on the substrate-influenced innovations of adults. Explicitly, Singler (2006a: 158) rejected the idea that creolization was identical to nativization, arguing with reference to Jourdan and Keesing's (1997: 411-416) work on Solomon Pijin (another variety of Melanesian Pidgin) that it is urban adults moving to use the second-language 'pidgin' as their primary language who introduce the structural and functional expansions considered as 'creolization'.

Consequently, anything other than a sociological distinction between first language (L1) = 'creole' and second language (L2) = 'pidgin' languages is insecure. In his discussion of the Life-Cycle theory, Rickford provides definitions of 'pidgin' and 'creole' languages that are both straightforward and practical: a pidgin is "a lingua franca arising from interlingual contact and showing reduction and admixture relative to its "input" languages", while a creole is "an expanded version of the pidgin that develops with its use as native or primary language" (Rickford 1987: 32). Rickford removes the first/second language creole/pidgin distinction from his definition, paralleling G.Sankoff's (1980d: 193) observations on Tok Pisin by acknowledging "use as ... primary language" alongside nativization as a source of expansion. Urbanisation, alongside the extended periods of plantation life and nativization, may produce this primary use, distinguishing jargons and stable pidgins from creoles and "pidgincreoles" (Bakker 2006).

For this reason McWhorter regards describing Tok Pisin as a pidgin language "vacuous" (McWhorter 1998). However, loose definitions of 'pidgin' and 'creole' languages as second- and first-languages respectively remain popular, as does the definition of Tok Pisin as an 'expanded pidgin': a 'pidgin' language (in that it is a second or additional language for many of its speakers) that has many or all of the features of a 'creole', first-language variety (Mühlhäusler 1979b, 1985h). Describing Tok Pisin as an 'expanded pidgin' acknowledges Rickford's observation (above), and that it is the primary language for many of its second-

language speakers, particularly in urban areas, while retaining a descriptive distinction between Tok Pisin and 'creole' languages as the first-languages of most of their speakers.

The developmental connection between pidgin and creole languages so defined, however, remains a subject of debate. Alleyne (1971) and Chaudenson (1992, 2001) highlight the absence of a 'pidgin' predecessor to Haitian Creole, and Mufwene (2000, 2009) maintains a socio-historical definition of 'creole' languages that distinguishes them from 'pidgin' languages and similarly excludes a pidginization phase in the development of creoles. These 'superstratist' accounts associate creoles with the languages of their lexicon, and view their development as language change away from a standard form, in contrast with the claim of the 'life-cycle' model for the discontinuity of pidginization as a key stage in creole development. The work collected by DeGraff (1999b: 3) expressly focuses on the different accounts of Bickerton and Singler of the relationship between pidgin and creole languages in terms of the role of nativization. McWhorter addresses the question of whether 'creole' languages can be defined typographically (McWhorter 1998) objecting to Mufwene's persistent contention that creolization is no different from other forms of language change (Mufwene 2001), and that creoles should be thought of as varieties of their lexifiers subject to the same contact-induced change as many other varieties (see Thomason and Kaufman (1988); Mufwene (1994)).

Myers-Scotton, in her account of the relationship between code-switching and the processes of pidgin and creole formation, explicitly treats the two 'types' of language indiscriminately.[4] She lists the creole-like features of expanded pidgins; the concurrent existence of pidgin,

---

[4] While her argument on typology is compelling, the application of the MLF model to pidgin/creole genesis is contentious in Pidgin and Creole studies. For a discussion of the principal concerns, see Siegel (2008: 142-147).

second language, and creole, first language, speakers, in many communities using 'pidgin' languages; the absence of evidence for pidgin predecessors for all creoles; and the gradual processes of creolization as reasons why distinguishing the languages on developmental, as opposed to sociological, bases is unproductive (Myers-Scotton 2002: 272).

Consequently, while Tok Pisin may formally be described as an 'expanded pidgin' with some 'creole' speakers, it should be equated with other, creole, languages of primary communication.[5]

3.1.2 The Creole 'Life-cycle' and Decreolization models

DeCamp (1971) expanded on Hall's 'life-cycle' model to include a 'post creole' phase: proposing that where an expanded, creole language remained in contact with the language that was its 'superstrate', speakers would continue to adapt and change the language to make it structurally resemble the more prestigious or powerful language. The notion of 'decreolization' strongly associates the process of pidginization with the 'imperfect acquisition' of a language by speakers from 'substrate' linguistic communities, and assumes a continuing desire to speak the 'superstrate' will lead speakers to target that language for imitation, and thereby continue to change their creole into a variety of that language.

The concept of decreolization provided a teleological elegance to the 'life-cycle' model: spontaneous jargons of convenience become established pidgins over time, and then, as they

---

[5] In this dissertation I use 'pidgin-creole' as a term to denote the entire range of varieties of Tok Pisin in terms of, for example, a continuum. I use 'expanded pidgin' to describe the conventional variety. Today, however, this contains many features elsewhere described primarily in 'creole' varieties. 'Pidgin' remains the PNG English term for Tok Pisin.

are adopted as the native language of new generations in the community, they expand to become fully-expressive 'creole' languages. Eventually, with the benefit of education and access to the 'superstrate' they become 'true' dialects of the more powerful language. In essence, this image of language birth and growth mirrored models of social and political development more generally: through extended contact with 'modern' cultures, indigenous communities would transition to join the ranks of 'Developed Nations'. As such it has been subject to subsequent revision as the contributions and voices of pidgin and creole speaking communities have been heard.

A distinction was drawn between 'basilects', varieties of pidgins and creoles that seemed most dissimilar to the superstrate, and 'acrolects', varieties of those same languages that seemed to resemble dialects of the superstrate. The term 'mesolect' was used to describe any varieties that appeared to exist between the most fundamental pidgin and creole language, the basilect, and, acrolectal, speech which was similar to the standard, prestige, language, usually a language of colonisation or other external authority.

The 'life-cycle' model had support in both diachronic accounts of the historical developments of pidgins in Africa and the Pacific (as consequences of indigenous interaction as well as colonial endeavour) joining with the conception of Atlantic creoles as nativized versions of earlier West African pidgins, and in the synchronic observation of creole continua (DeCamp 1971; Bickerton 1975b; Rickford 1987).

For Jamaica, Thomason and Kaufman (1988) interpret the observed creole continuum (see s2.6 for some key concepts around this term in PNG) as arising through a process by which urban creole speakers, experiencing greater exposure to English as a result of social and political changes, developed new varieties of creole which, compared to the varieties of creole spoken by their rural counterparts, have closer resemblance to English in certain areas of pronunciation, lexicon, and syntax.

LePage and Tabouret-Keller (1985) propose a mechanism by which such differential social associations of English and Patwa, the local name for Jamaican creole English, in Jamaica encourage flexible performance by locals. Such variable performance, and the nuances of identity it can be used to indicate as it develops norms within the community ultimately generates 'decreolized' varieties of the creole.

In the case of Tok Pisin we have a fundamental English-lexified pidgin which, during the period of German rule in New Guinea, was isolated from the lexifier, and so expanded its lexicon by borrowing from languages other than English (Mühlhäusler 1979a), and developed its grammar independently of influence from English norms (Mühlhäusler 1985d). It was only following the First World War that English became the administrative language of the Territory of New Guinea, as it had always been in Papua. Consequently, the role of English as target or superstrate is not clear, and significant internal developments, through expansion and creolization, have been observed (Sankoff, G. and Brown 1976; Sankoff, G. 1980d; Sankoff, G. and Laberge 1980 (1973); Smith, G. 1990; Romaine 1992; Smith, G. 2002).

Mühlhäusler (1980: 22) introduced the terms "post-pidgin" and "post-pidgin continuum" to describe situations specifically relating to pidgin languages, making a distinction O'Donnell and Todd (1980) did not. Their description of the developmental process for continua applies equally to both groups of languages:

> When a creole or expanded pidgin exists in a community where its lexical source language is the language of education and politics ... the two linguistic systems inter-influence each other with the result that one finds, not two distinct systems, but an unbroken spectrum between the pidgin or creole on the one hand and the prestigious standard language on the other.
>
> (O'Donnell and Todd 1980: 52)

Mühlhüsler (1985d) gives an account of post-pidgin varieties in Tok Pisin, and described such varieties as "involving the restructuring and/or replacement of earlier lexicon and grammar in favour of patterns from the superimposed 'target' language" (Mühlhäusler 1997: 211-12).

3.1.3 Forming Creole Continua

De Camp acknowledges that there may be discrete dialects for individuals (DeCamp 1971: 368), despite maintaining that the Jamaican language system as a whole cannot be characterised in terms of distinct social dialects (DeCamp 1971: 354). This argument rests on the production of property-item matrices that show a graded scale of the various intermediate grammatical forms of expression found in creole communities. The continuum of variation makes dividing creoles into distinct dialects into an "arbitrary" decision (Labov 1973: 343).

The existence of a linguistic continuum of variation between 'basilectal', more creole, varieties, and 'acrolectal', more standard, varieties, via intermediate 'mesolectal' varieties in creole language communities has been demonstrated through observation (DeCamp 1971; Rickford 1987; Ho and Platt 1993). However, the superstratist theories and the 'life-cycle' model essentially differ in their account of the diachronic development of the intermediate 'mesolectal' varieties. Under the 'life-cycle' model, the basilectal creole is spoken first as a development of pidginized language, and the mesolect and acrolect being created as progressive developments over time. Other scholars have argued that change in creole languages does not necessarily take place in the direction of the lexifier (Le Page 1977; Bailey, G. and Maynor 1998). In Tok Pisin language internal developments have been shown to take place through the expanded pidgin and creolization phases (Sankoff, G. 1980d; Sankoff, G. and Laberge 1980 (1973); Mühlhäusler 1985d; Smith, G. 1990; Siegel 1997; Smith, G. 2000, 2002) (see s3.1.4). Superstratist models reverse the direction of this development, with the acrolect being spoken in mesolectal varieties during periods of close cultural contact, and basilectal forms only being spoken once the balance of communities shifts away from native-speakers of the superstrate varieties (Rickford 1987: 32; Chaudenson 2001; Mufwene 2001, 2008). As a consequence, the term 'creole continuum' (or 'pidgin

continuum') is now preferred as a non-teleological, purely descriptive, option (Siegel 2008: 237).

Providing crucial evidence in this debate, socio-historical and historical research into Atlantic creole communities has shown that different varieties of creole have been present since the earliest history of creole language (Le Page 1960; Alleyne 1971; Edwards 1984; Lalla and D'Costa 1990; Valdman 1991; Aceto 1999). Furthermore, sociolinguistic research has shown that creole continua consist of distinct varieties, which are associated with different social and cultural values (Edwards 1980; Le Page and Tabouret-Keller 1985; Rickford 1986; Winford 1992; Winer 1993; Winford 1993; Winer 1995).

Other researchers have concluded that in some situations pidgin/creole continua do not, in fact, exist: finding that the observed variation is due to the interaction of discrete systems (Lawton 1980; Edwards 1983; Siegel 1997; Winford 1997). This interaction is framed as code-switching (CS) or as style-shifting.

The notion of creole continua arising from the interaction of discrete systems was proposed by Bailey (1971: 342), but rejected by Bickerton on the basis that it suggested that: "the continuum is simply produced by the random mutual interference of two discrete and self-consistent grammars" (Bickerton 1973: 641). However, subsequent research into the interaction of grammatical systems in mixed language behaviour (CS) (see s3.2, below) has provided systematic models of multi-lingual and multi-dialectal performance that can inform consideration of the continuum as the result of non-random behaviour.

Some research into creole continua has found resemblances to bilingual settings, suggesting that these are sociolinguistic rather than linguistic continua and are created through the performed interaction of different varieties of creole (Winford 1997; Migge 2007). Elsewhere, Aceto maintains that Basiamentos creole has developed a continuum through processes of internal variation that are considered normal in established languages but rarely considered as

motivators of language change in creoles (Aceto 1999). English is not the metropolitan prestige language in Basiamentos, and social variation is slight, yet there exists radical variation within the verb phrase which parallels Bickerton's observations on continua (Bickerton 1975b).

Aceto's position may be compared with Mühlhäusler's accounts of sociolects in Tok Pisin (Mühlhäusler 1975; Mühlhäusler 1985g). While Aceto is engaged in a project of hypothetical historical reconstruction the distinction made between 'Urban' and 'Rural' Tok Pisin has developed in living memory, with Mühlhäusler, and other commentators, accounting for the differences between these varieties as being the result of post-war exposure to English as the administrative language of the expatriate and governmental settlements that became Papua New Guinea's towns and cities.

Fundamentally, mixing with English is claimed to be the source of the sociolectal variation in Tok Pisin, just as it is in cases of decreolization elsewhere in the world. However, in PNG mixing is described as taking place between separate systems: as code-switching (Mühlhäusler 1985d; Siegel 1997; Smith, G. 2000, 2002; Siegel 2008).

Observations of the similarities between behaviour in creole communities, bilingual behaviour elsewhere, and processes of monolingual language development provide us with a picture of creole language variation without the developmental direction of 'decreolization'. Aceto (1999) maintains that the social use of a range of distinct language varieties in a community can produce mixed forms which structure a continuum between them without necessarily moving the whole community away from one system and towards another. Winford describes style shifting between different systems (which patterns the variable behaviour of observed continua) as providing insight into processes of language change, but notes "[w]e cannot, of course, confuse style shifting with decreolization, or assume that all variation is a reflection of change" (Winford 1997: 274).

The key issue is one of targeting. The premise that the English language, functioning as superstrate in the national community, is the target language for changes in Tok Pisin in Papua New Guinea is fundamental. A concern that English as a municipal and world language will be the default target underlies arguments presented by some commentators (Healy 1975; Mühlhäusler 1985d; Wurm 1985) against borrowing from English as a developmental strategy for Tok Pisin. These arguments originate with concerns that if the integrity of Tok Pisin is compromised its value as a lingua franca will be diminished.

Only the attitudes of people using a creole language are significant to any discussion of 'targeting' languages. Associations of prestige encourage speakers to emulate a particular variety of speech, either as a dialect of a language or as a separate language system. This is why decreolization theories rely on the notion of a 'superstrate' language in the community. However, issues with the identity-associations of a particular variety (see s2.6 and also s3.2.4.2, below) may encourage people not to emulate a variety: some people may choose to resist using a 'prestige' variety as it conflicts with their own identity. In a discussion of the common-core account of the origin of pidgin grammars, Mühlhäusler (1985d: 78-80) makes a suggestive account of language mixing where he emphasises that as pidgin grammars expand and develop systematic structures of their own they become less permeable to structures from other systems. New borrowings may too easily conflict with, or upset, existing systems. As a consequence, even when Tok Pisin is mixed with English in the urban context, it may not necessarily resemble the English language any more than conservative Rural Tok Pisin. The mixing of system A with system B produces a third system, C, which is determined by, Mühlhäusler suggests, such factors as historical and typological distance, degrees of systemacity in structures in each language, and by the degree of social tolerance for "ill-integrated and unstable systems" (Mühlhäusler 1985d: 79).

With contemporary models of code-switching and other varieties of language mixing, however, we are equipped with theoretical tools which were not well-developed when

Mühlhäusler provided his treatment of anglicisation in Tok Pisin. Applying the models of behaviour described by Poplack, Myers-Scotton, Muysken, Auer and others (see s3.2, below) to the mixed language of anglicised Tok Pisin may allow us to appreciate the regularities in what, for Mühlhäusler, was an "unstable" form of language.

Of these, G.Smith (2002: 202-209) in his discussion of what he describes as the "wholesale borrowing" (Smith, G. 2002: 99) from English into contemporary varieties of Tok Pisin explicitly identifies Myers-Scotton's Matrix Language Frame model (1993a) as of value in his analysis.

G.Smith's own corpus, however, while collected from first-language Tok Pisin speakers in the language contact environment of English-medium schools, is composed of monologues provided by schoolchildren who were explicitly directed to use Tok Pisin. As a consequence there are only minimal indications of code-switching in his corpus, and while he does find "a good deal of influence from English" (Smith, G. 2002: 202) in some sections of the community, he has no evidence of an established continuum. Siegel (1997, 2008) compares investigations of creole continua around Melanesia, and maintains that there is insufficient evidence to substantiate a claim for their existence.

Consequently, whether an association is made between decreolization and continua or not, questions about the on-going relationship between English and Tok Pisin, the mechanisms of their interaction, and the possible directions of change between each of them remain to be resolved.

3.1.4 Development History of Tok Pisin

Excellent accounts of the social and linguistic development of Tok Pisin are provided by Wurm and Mühlhäusler in their 1985 *Handbook of Tok Pisin* (Wurm and Mühlhäusler 1985a). Mühlhäusler (1974, 1985a, 1985d) provides an account of the development of Tok

Pisin from Pacific Jargon English, through to a stabilised variety of Melanesian Pidgin English particularly during the German plantation period when New Guinea plantations were exclusively associated with those is Samoa (Mühlhäusler 1978), to the nativized variety developed by WWI, with creolization through the birth of first-language speaking children taking place as a process concurrent with the expansion and diversification of Tok Pisin among the adult community, and a final stage of 'depidginization' and 'decreolization' through continuing contact with English. In the *Handbook* this is compared with Todd's (1974) model which has a nativized pidgin developing into a 'post creole continuum' due to "increased pressure from the lexical source language" (Mühlhäusler 1985a: 66) and presented as an improvement on the 'life-cycle model' as described by Hall (1962), primarily because it captures a sense of the continuing dynamics in Tok Pisin development, for which (first-language) creolization is not a significant phase. Vitally, Mühlhäusler notes that all the proposed models obscure a fundamental reality of Tok Pisin development: that different sections of the Papua New Guinean community are using the language at different 'phases' of its development at the same time; that 'diachronic' developments over time can be observed synchronically (in the same time) in PNG as one moves from urban centres to highly rural communities. Effectively: that Tok Pisin in Papua New Guinea exists concurrently in many different phases throughout the country (Mühlhäusler, Dutton and Romaine 2003: 5). In a sense, the name 'Tok Pisin', officially agreed-upon in 1981 (Romaine 1992: 19), is a cover term for a wide range of varieties within PNG (Mühlhäusler 1985c: 28).

The presence of different 'sociolects' was identified by Mühlhäusler (1975) and has been fundamental to characterisations of Tok Pisin. However, it is clear that post-Independence developments in the national community in terms of social mobility and access to English have weakened the proposed 'Urban'/'Rural' distinction in Tok Pisin. Romaine (1992) found 'Urban' features in the speech of children from rural communities, as did G.Smith (2002),

who concluded that access to English was the most significant factor distinguishing varieties of the language, rather than social context.

> [T]he sociolectal division of the language is mainly the result of differences in contact with the English language and Western culture: until shortly before the independence of the country only a small number of native inhabitants, mostly the town dwellers, had easy access to English. Since then the English language has increasingly penetrated into the interior through the school system, and consequently the differences between the native sociolects seem to be getting smaller, as found by Dutton in his 1985 field research in New Guinea.
>
> (Mühlhäusler, Dutton and Romaine 2003: 4)

In terms of geographic variation, little evidence since Smith (2002: e.g. Section 5.4) has been provided to show regional varieties exist, yet this may be due to a lack of large-scale data-collection projects: many Tok Pisin speakers claim to be able to recognise the regional origin of an individual by their pronunciation or syntax (Mühlhäusler 1985h: 256). Crowley has questioned the appropriateness of Mühlhäusler's developmental model as appropriate for other varieties of Melanesian Pidgin (s2.5) on exactly the basis of multi-variate developmental pathways (Crowley 1990a: 384-6). In PNG, age of acquisition, multilingualism, the social role of Tok Pisin, and in particular access to English all impact on the way the language is spoken. Nonetheless, it is the 'Rural' pidgin of village communities in the Morobe, Madang, Sepik (MOMASE) region that is considered to be the oldest, best, purest, and most ideal form of Tok Pisin (Mihalic 1971, 1975). This is the conservative standard by which other varieties are measured.

Mühlhäusler's account of the internal development of Tok Pisin (Mühlhäusler 1985d) incorporates the concept of the pidgin-creole life cycle. Following on from his discussion of the origins of the language in an unstable jargon he highlights the influence of socialisation processes in forcing stabilisation. Mühlhäusler favours Salisbury's (1967: 46) comparison of Siane in 1952 with the plantations of New Britain in 1881. He observes that as a community of speakers grows in size idiosyncratic deviations from developing community norms of speech may become the subject of overt criticism. This negative re-enforcement of

community norms becomes the significant factor in encouraging the stabilisation of the new language of communication. Such social pressure is convincingly held by Mühlhäusler to be the primary factor in the standardisation of communicative patterns that led to the development of Tok Pisin as a stable language.

Correction of aberrant pronunciations or grammatical constructions can only be a process of an established language community. As Tok Pisin came to be used by larger numbers of people, developing norms of pronunciation and grammar lead to its stabilisation, and ultimately to the expansion of its grammar. Mühlhäusler identified second-language speaking adults as the significant vector for this process (Mühlhäusler 1985g: 480), and G.Sankoff observed against this backdrop of adult-to-adult transmission that children's Tok Pisin was only significantly differentiated from adult language in the area of phonology (Sankoff, G. 1980e: 269), with impacts on their grammars through morphophonological deletion. Other 'creole' features in children's speech are also represented in second-language adult speech (Sankoff, G. 1977, 1980d; Sankoff, G. and Laberge 1980 (1973)).

Subsequent work by Romaine and by G.Smith (Romaine 1990; Smith, G. 1990; Romaine 1992; Smith, G. 2002) has illustrated the continuation of these changes in children's speech over subsequent generations, but whether adult norms are eventually effected by the innovations of creolized speech (Mühlhäusler 1985a: 41) has not yet been resolved.

Kulick (1992) observed that in Gapun, children were largely responsible for teaching other children Tok Pisin, and this accords with many observations on the role of the schoolyard in early language socialisation for Tok Pisin speakers. The growing existence of a nationally-mediated standard for Tok Pisin may lead to increasing distinctions between varieties in the language, yet as an, unofficial, national written standard has been established (Verhaar 1995), so too variation in spoken language has continued to develop.

Mühlhäusler's consistent observation is that Tok Pisin represents a different case to Atlantic creoles in that second-language varieties show the same characteristics as first-language (creolized) Tok Pisin (s3.1.1), and consequently establishes that creolization is not a necessary requirement for grammatical expansion: the pressure for efficient communication is stronger than that for linguistic and structural sophistication. He finds that true and rapid creolization, in the sense of the wholesale nativization of a jargon or basic pidgin as the 'mother tongue' of children, as a mechanism of structural change is only found in situations where whole populations have been permanently moved (Mühlhäusler 1985a: 41).

For Mühlhäusler, Tok Pisin represents a third 'type' of creolization (Mühlhäusler 1985g: 479), differing from Hawaiian and Atlantic creoles which, it was prominently argued at the time (Bickerton 1981, 1984), rapidly developed from a jargon to a creole, and from other creoles which first stabilised as a pidgin before creolizing. Tok Pisin exists as an expanded pidgin: a second language with the features of a creole. Making the distinction between 'pidgin' and 'creole' languages, as I have discussed above (s3.1.1), may, in this case, be redundant.

In their collection of Tok Pisin texts, Mühlhäusler, Romaine, Dutton et. al. observe that creole innovations such as phonological reduction and cliticisation, obligatory plural marking, relative clauses produced without a relativizing pronoun, multiple embedding in single sentences and an increased speech tempo are all present in varieties where Tok Pisin is not the first language of speakers ("non-creolised" varieties) (Mühlhäusler, Dutton and Romaine 2003: 11). As Mühlhäusler observes:

> The more useful Tok Pisin becomes, the earlier it is learned; the age of learning is a true continuum, and the distinction between first- and second-language learning is becoming blurred
>
> (Mühlhäusler 1985d: 149).

In their studies of schoolchildren's Tok Pisin, Romaine (1992) and G.Smith (1990, 2002) have built on Mühlhäusler's account and described the 'creole' language in considerable

detail. Vernacularization has been observed to be more significant than creolization (viewed as the use of the language by 'native', first-language speakers) in the development of Tok Pisin, with adult-to-adult transmission and early acquisition of the language apparently leading to many developments considered a consequence of first-language use in other pidgin/creole languages (Sankoff, G. 1980d; Sankoff, G. and Laberge 1980 (1973)) (see also s3.1.1 above).

In PNG today, first-language Tok Pisin speakers exist, and though exact numbers are unknown (Wakizaka 2008; Paliwala 2011, 2012) the Summer Institute of Linguistics (SIL) reports a figure of 122,000, including 50,000 monolinguals (Lewis 2009). The most directly verifiable data available to date are literacy figures from the 2000 National Census, which records c64,000 people mono-literate in Tok Pisin (National Statistical Office 2002a).

First language, 'creole' Tok Pisin has been the subject of research by Romaine (1992) and G.Smith (2002). However, many of the features identified as characteristics of 'creole' Tok Pisin are found also in the fluent speech of multilingual adults. Mühlhäusler observes:

> The differences between non-creolised and creolised Tok Pisin are slight rather than drastic.
>
> (Mühlhäusler 1985d: 151)

The main differences are in rate of speech and phonological reduction. Mühlhäusler's account of creolization in Tok Pisin (Mühlhäusler 1985d: 148-162) distinguishes 'Rural' and 'Urban' sociolects, though G.Smith (2002) concludes that geographical location is no longer as significant. In general, Mühlhäusler's distinction is based on the influence of English on the varieties he terms 'Urban':

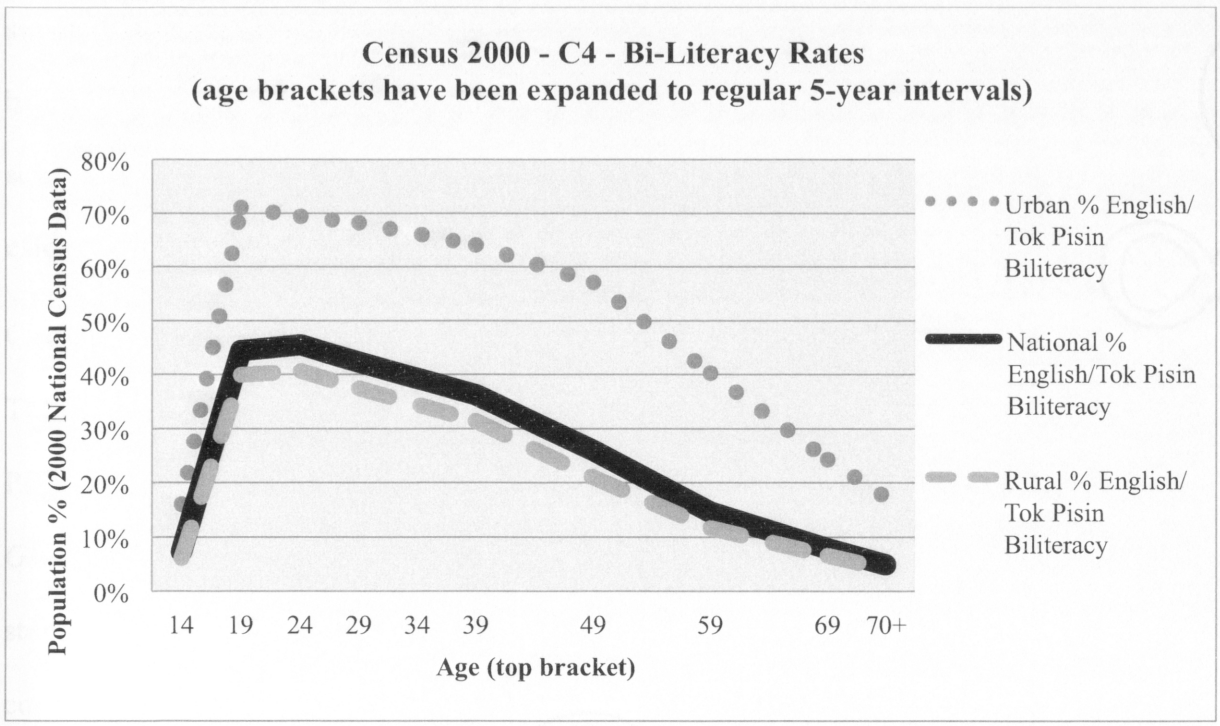
> The principal force shaping Urban Pidgin is the renewed contact with English resulting from English education and widespread English-Tok Pisin bilingualism.
>
> (Mühlhäusler 1985h: 248)

Considering Census data on literacy 15 years after this observation (National Statistical Office 2002b) reveals how G.Smith's position is borne out. Figures from the 2000 National Census

indicate that the Papua New Guinean national population is predominantly rural: 4.5 million out of a total of around 5.2 million people. The 2000 National Census asked people what languages they could read and write with understanding, and so aimed to collect data on literacy rather than spoken language. However, we can use the reported figures as an indicator: a baseline population of people acknowledging skills in different languages. It may be less reliable as an indicator of monolingualism (as people were only required to report the languages they could read or write) but 2000 National Census bi- (or multi-) literacy are considered here as underpinning actual bi- (or multi-) lingualism, though whether the multi-literate skills are actually in regular use remains a significant question (Siegel 2010).

Figure 3.1 presents a plot from data collected on literacy and shows that for the adult urban population between 15 and 49 around 60-70% claimed bi-literacy with English and Tok Pisin. This is significantly higher than the rural population, where 20-40% is the norm.

Figure 3.1 : National Census 2000 - Bi-literacy



Taking literacy figures from the 2000 Census as a baseline figure of language competence generally (Paliwala 2012), the national population of English-Tok Pisin bilinguals represents at least 24% of the total population, including young children and the over-60s.

It is important to note that this bi-literacy is a feature of the latest generation of Papua New Guineans: those who were young adults (c25 years old) at Independence in 1975 were around 50 years old in 2000, and those born since Independence were 25 years old in 2000. This can clearly be seen in Figure 3.1 with increased percentages for these age groups in both rural and urban communities compared to the over-50s. The post-Independence rural generation reported the same Tok Pisin/English biliteracy rates as the urban 50-59 year olds: those born between 1941 and 1950. Mühlhäusler's (1985h: 248) assertion on the vigorous contact with English that is a feature of 'Urban' Tok Pisin can be clearly seen in the 2000 Census figures as a characteristic of 40% of the young rural population also. This indicates that any influence on Urban Tok Pisin may also be a significant factor in Tok Pisin on the national level, and supports G.Smith's assertion that the rural/urban distinction is not as significant as bilingualism with English (Smith, G. 2000; 2002: 210-211). Clearly, the impact of English on schoolchildren's Tok Pisin is a factor that must be given careful consideration, as are the effects of English on adult speech.

Cutting through such impacts on spoken language, Verhaar's (1995: 3) "virtual standard" of Tok Pisin, derived from written sources, explicitly excludes reference to the national Tok Pisin newspaper Wantok due to an expectation of anglicisms. Verhaar's detailed *Reference Grammar* (Verhaar 1995) is developed with reference to texts based around an ideal 'Rural' standard, such as the *Buk Baibel*. However, Mundhenk (1990) illustrates the compromises and conflicts involved in composing such a standard. Indeed, understanding variation in Tok Pisin may require a focus on bilingualism with English, but, in line with later research into bilingual behaviour, code-switching, and the development of mixed language varieties, Mühlhäusler suggests:

> ... it would be an over simplification to assume that all borrowing from English will make Tok Pisin more like English. It is true that one of the tacit assumptions in many studies of post-pidgin or post-creole continua has been that the eventual outcome of contact between an English-derived pidgin or creole and standard English is English. However, it is not obvious why the mixing of two linguistic systems should lead to the replacement of one system by another.
>
> (Mühlhäusler 1985h: 248)

Certainly, studies of CS in communities have revealed that the development of unmarked mixed systems is a more common outcome, so the question may be asked whether bilingual behaviour in Papua New Guinea is, indeed, producing "a new third system" (Mühlhäusler 1985h: 248) (see also : Mühlhäusler 1985d: 80), and whether this can be seen as mesolectal between Tok Pisin and English:

> [I]n spite of heavy borrowing, Urban Pidgin does not appear to be more readily intelligible to a speaker of English than Rural Pidgin. At the same time, it is no longer easy to understand for speakers of Rural Pidgin.
>
> (Mühlhäusler 1985h: 248)

Alongside the characterisation of Tok Pisin as a fully expanded pidgin language is the notion of its 'post-pidgin' development. Mühlhäusler uses this term to define space analogous to the 'decreolization' conceived under the 'life-cycle' model, yet his distinction between 'post-pidgin' and the 'Urban' 'anglicised' "third system" he suggests is not entirely clear. In his discussion of 'post-pidgin' varieties of Tok Pisin Mühlhäusler (1985d) indicates the role mixing with English has in developing a mesolect, where the variation across the expanded range of Tok Pisin allows the possibility of borrowing, as in a creole (Mühlhäusler 1985d: 133).

In Papua New Guinea it is the case that English, the lexifier language for Tok Pisin, is largely the language of formal business, international trade, and social commerce. English is the official language of education, and is the dominant language of publication, and consequently of politics, government, and administration. In this respect PNG appears a typical case for decreolization: a country with widely-spoken expanded pidgin, learned as a first language by at least 6% of the population and spoken fluently by at least 43% of the population (National

Statistical Office 2002a; Paliwala 2011, 2012); where the national superstrate is also the lexifier language of the expanded pidgin; and where there is a healthy (at least 33% (Paliwala 2011, 2012)) population whose bilingualism with English is encouraged by English-medium education and English-language national media. However, the role of English as material superstrate, or as the target language for speakers of Tok Pisin, has not been thoroughly investigated.

Among researchers on Tok Pisin, Bickerton (1975a), G.Sankoff (1980c: 26), and Kale (1990: 190) commented on the possibility of a Tok Pisin-English continuum, while O'Donnell and Todd (1980: 52), Aitchison (1981: 212-14), and Romaine (1992: 323) maintained that decreolization and/or post-creole continua exist in PNG as in the Caribbean and in West Africa, and Harvey (2007) and Devette-Chee (2011) argue for this on the basis of evidence from specific registers of Tok Pisin.

In contrast, Mühlhäusler gives examples to show the post-pidgin situation in Tok Pisin, but ultimately argues against a unilinear process of "depidignisation" (Mühlhäusler 1985d; 1997: 211-21). Similarly, Wakizaka (2008) finds that 'anglicisation' does not equate to decreolization: like Devette-Chee they find that Tok Pisin remains distinct from, though in a potential continuum with, English.

3.1.5 New Varieties in Tok Pisin

Mühlhäusler's position (Mühlhäusler 1985d; 1997: 211-221) is that it is because of the variation between highly-structured, conservative Tok Pisin and flexible creolized or expanded Tok Pisin that new types of mixing with English may be possible, even while such mixing may erode the stable structure of 'pidgin' Tok Pisin and reduce Tok Pisin's value as a comprehensible lingua franca (Mühlhäusler 1985b).

Mühlhäusler observes: "messy languages are good mixers", though he notes also: "contact between Tok Pisin and English is unlikely to result in a smooth transition between the two systems" (Mühlhäusler 1985d: 134). He maintains that a continuum is unlikely, and, as I discuss below (s3.2), argues that the identification of code-switching indicates that a Tok Pisin form of a 'creole continuum' is not present.

Cases of code-switching of various types are commonplace (Mühlhäusler 1985d: 132-148; 1985h: 271-273), but can the breadth of contemporary models of bilingual behaviour give some insight into the development of "intermediate varieties" (Romaine 1992: 321) in Tok Pisin?

Mühlhäusler indicates the difficulty: A distinction must be made between a mechanical mixture and fusion. In the former case borrowed elements can be uniquely assigned to one or the other system, in the latter a new construction can be assigned to neither or both simultaneously. He describes the development of a new third system as a possible result of language contact in the post-pidgin stage, while noting that linguists have considerable difficulties in identifying and describing such new systems (Mühlhäusler 1985d: 133).

While Laycock (1985a), countering Bickerton (1975a), takes the view that diglossic separation between Tok Pisin and English will prevent the formation of a continuum, he also argues that "the heavily anglicised varieties of Urban Pidgin are, for most of Papua New Guinea, as dead as Tok Masta"[6] (Laycock 1985a: 667), and observed in a footnote:

---

[6] 'Tok Masta' is a term used to describe the attempts at Tok Pisin made by foreigners. It is described and discussed by Mühlhäusler (1979b, 1985h), and others, and discarded as a jargon that is not actually a legitimate variety of the language.

> Even radio broadcasters have largely given up the extremely anglicised Tok Pisin of a few years ago … it is only as a joke today that one is likely to hear such linguistic mixtures as *yumi mas tok long grass-roots level nau* or *yu no ken putim cart before the horse* – that is, as a relaxed form of speech among speakers at home in both English and Tok Pisin.
>
> (Laycock 1985a: 668) (Tok Pisin italicised here)

With the 2000 National Census data as an indication, we must conclude that the spread of both English and Tok Pisin in Papua New Guinea has resulted in this community of "speakers at home in both English and Tok Pisin" (ibid.) being the majority in urban settings, and a highly significant minority in rural areas.

Consequently, the conditions and behaviours Laycock claims were dying out in 1985 may be a significant factor in 21st Century Papua New Guinean Tok Pisin, alongside the development of first-language 'creole' varieties and the use of Tok Pisin by a great many multilingual adults of the post-Independence generation.

Bickerton's case for the eventual decreolization of Tok Pisin and the emergence of a continuum with English (Bickerton 1975a) (s2.6) was based on the social and linguistic circumstances of Papua New Guinea as it moved to Independence, which were, he argued, as described by DeCamp (1971: 351). Consequently, he presented the decreolization phase of the 'life cycle' process as inevitable. A decade later, Mühlhäusler (1985d: 148) explicitly argued that the pinpointing of 'code-switching' between Tok Pisin and English indicated that no continuum had formed between them. Crowley came to a similar conclusion in relation to Bislama in Vanuatu (Crowley 1990), and Jourdan (1989) has supported this.

In Papua New Guinea, while some researchers have made reference to a pidgin, or creole, continuum (O'Donnell and Todd 1980; Romaine 1992: 323; Mühlhäusler 1996), detailed study of data from across Melanesia by Siegel (1997), and from within PNG specifically (Smith, G. 1998, 2000; 2002: eg. 209-211) has not found the kind of scalable variation considered fundamental to the continuum model. Intermediate varieties, such as the 'anglicised' Urban Pidgin, exist, and form a continuum with the more conservative speech of

Rural Pidgin, but evidence of the presence of intermediate, mesolectal, varieties between English and Tok Pisin remains unavailable. The characterisation of the status of the relationship between Melanesian Pidgin English varieties and English in Melanesia as involving the interaction of two discrete systems, rather than as a range of varieties forming a single continuum from e.g. Tok Pisin to English has become a standard position.

A recent Masters study (Wakizaka 2008) of interview recordings of Tok Pisin speakers has considered language change and the erosion of Tok Pisin through 'anglicisation'. Wakizaka concluded that Tok Pisin remained a vibrant language which remained separate from English; and they found, in agreement with G.Smith (2002), that 'anglicised' Tok Pisin featured English elements embedded within a distinctively Tok Pisin Matrix Language Frame (MLF) (Myers-Scotton 1993a). While noting that the use of English in urban areas was being promoted by access to IT, and that this could lead to an increasing urban/rural linguistic divide, Wakizaka's analysis does not, however, take account of the distinctive characteristics of PNG varieties of English. Nor is their primary data based on naturally-occurring conversations of the type which may include the most casual speech.

Conversely, in a PhD dissertation discussing the written language of internet forums, Harvey (2007) applied the same concept of Matrix Language (ML) to data from the internet and from letters to *Wantok* newspaper to conclude that internet users' language showed signs of a ML turnover (Myers-Scotton 2002), associated with a greater influence from English and less access to Tok Pisin. Harvey bases this analysis on a notion of a 'standard' Tok Pisin grammar which is characterised by obligatory marking of the predicates with the particle *i*. However, while this is certainly a characteristic of the written 'virtual standard' described by Verhaar (1995) and followed by the editorial staff at *Wantok*, variability in the marking of predicates with *i* is a well-established feature of many spoken Tok Pisin datasets (Smeall 1975; Woolford 1979b; Mühlhäusler 1987, 1990; Smith, G. 1990; Verhaar 1991; Romaine 1993; Smith, G. 2002), and has not elsewhere been clearly associated with the impact of English. As

a consequence, the loss of the particle *i* that Harvey describes in Internet discussions compared with the written standard may, in fact, be a feature of the less formal register of Tok Pisin adopted by contributors online.

This is fascinating as Hinrich's work on Jamaican language in similar computer-mediated communication found that the "ordered and small transitions of the spoken continuum" between basilectal and mesolectal speech in Jamaican society have been replaced in online communication by two separate codes, with code-switching between them (Hinrichs 2006: 40).

Devette-Chee (2011) considers the lyrics of PNG musicians as evidence for code-switching forms that, she argues, represent decreolization (see s3.1.3 on problems with this position), but maintains that this has led to the establishment of new varieties of Tok Pisin that exist in a continuum with English.

The recent studies by Wakizaka, Harvey, and Devette-Chee support the on-going vibrancy of Tok Pisin as a language of solidarity and national identity, though they differ in their datasets and their conclusions about the relationship between Tok Pisin and English in code-switching, and the status of 'anglicised' Tok Pisin as language variety completely distinct from English.

My research project investigated these themes in greater depth through reference to diverse scholarship on both code-switching and on Tok Pisin and PNG English. In particular I focused on conversational data to investigate the on-going presence of anglicised varieties of Urban Pidgin and their relationship to code-switching behaviour between Tok Pisin and English in Papua New Guinea.

3.1.6 New Varieties in English : the PNG Dialect

Under the Australian Administration of Papua, English was the official language while 'Police' Motu, a pidginized variety of the Motuan language that is now known as Hiri Motu,

was developed and used as a lingua franca (Dutton 1978, 1982; Wurm and Mühlhäusler 1985b). As Tok Pisin developed and spread in German New Guinea in the north, so 'Police', or Hiri, Motu gained currency in the southern British New Guinea, later the Australian Mandated Territory of Papua (s2.2). Some Papuan people in contact with foreigners and those of the western, Trobriand Islands, became used to English as the colonial language, while people in New Guinea were isolated from English-speakers during German rule. Following World War I, with both areas under Australian administration, Papuans came into contact with Tok Pisin speakers, and New Guineans with speakers of both Hiri Motu and English.

Historically, some negative attitudes were recorded towards Tok Pisin by local people in the southern region, potentially a legacy of English-speaking, Australian, attitudes to the 'pidgin English' (Wurm 1985: 71-72). English, however, came to be seen as an aspirational language, to which people had limited access. From the 1950s primary and government school education was in English, and it became and remains the language of the educated Papua New Guinean élite, the official language of government records and communications, and the dominant language in the national media. Importantly, in the generations since Independence, the English language taught as a foreign language in schools and used in universities and in government business has been taught by Papua New Guinean teachers and used primarily as a language of communication between Papua New Guineans themselves.

As it has come to be learned, used, and taught by Papua New Guineans, so English in Papua New Guinea has developed its own characteristics, so that it has been considered alongside other 'World Englishes' of the "outer circle" (Bolton and Kachru 2003) of, largely post-colonial, 'New English' varieties (Platt, Weber and Ho 1984). It has been recognised as quite distinct from the 'Standard' or 'International' Englishes, usually dialects of British, American, or Australian English, of the "inner circle" (Bolton and Kachru 2003). While not consistently represented in discussions of World/New Englishes, which occasionally include Tok Pisin as a Papua New Guinean variety of English (Bailey, R.W. and Görlach 1984 [1982]; Cheshire

1991; Schneider 2011), a position Mufwene (Mufwene 1994) explicitly argues in support of, the existence of Papua New Guinea English and its characteristics are broadly acknowledged (see, for example Romaine's introduction to the section on Pacific Englishes in Cheshire's volume (Romaine 1991)). This indigenised variety of English has been researched and described, and included in comparative studies such as that of Platt, Weber, and Ho (Platt, Weber and Ho 1984).

The indigenised variety of PNG English emerged out of urban communities in the late 1970s and early 1980s: discussing Tok Pisin sociolects in 1979 Mühlhäusler considered the emergence of "a kind of Papua New Guinea English" (Mühlhäusler 1979b: 237) as a possible alternative to the crystallisation of a stable intermediate system between Rural Tok Pisin and English, and by 1982 he was able to describe the structure of Urban Tok Pisin as being the result of contact "with the type of English that is spoken in Papua New Guinea" (Mühlhäusler 1984 [1982]: 454). Characteristics of this Papua New Guinea English (PNG English, or PNGE) were recorded and discussed early in this period by A.-M.Smith (Smith, A.-M. 1978, 1986, 1988a, 1988b) and throughout the 1980s by other authors working around the country, largely with reference to written registers (Davani 1984; Barron 1986; Yarupawa 1986; Holzknecht 1989). Platt, Weber, and Ho (1984) provide examples of both spoken and written usage. Melchers and Shaw (2011) observed more recently:

> [I]n New Guinea English is used mainly with other local speakers for internal purposes, and consequently can become very nonstandard.
>
> (Melchers and Shaw 2011: 183)

Crucially, they note that considering Tok Pisin borrowings from English alongside the 'nonstandard' features of some PNG English suggests that "something like an English-Tok Pisin continuum is developing" (Melchers and Shaw 2011: ibid).

As has been the case in the contexts of Caribbean Creoles discussions of the relationship between Tok Pisin and English in PNG have frequently assumed an international, seemingly

"country-free" standard English (see Irvine (2004); Dueber (2009). In the case of PNG, this has often been an Australian variety, and key resources on Tok Pisin grammar refer to overseas norms. However, for Papua New Guinean speakers it is PNG English that is the source of borrowing in 'anglicised' Tok Pisin, and as such the characteristics of PNG English will be significant in the consideration of the Tok Pisin and English grammars in Chapter 5, and throughout this dissertation.

3.2 Part II : Code-switching Research

Like research on pidgin and creole languages, research into bilingual behaviours, including those broadly described as 'code-switching', has developed relatively recently within linguistics (s3.2.3). This section covers issues from bilingualism studies relevant to consideration of the relationship between Tok Pisin and English in PNG. Subsequent to Mühlhäusler's presentation of code-switching as evidence that a continuum has not come into being (Mühlhäusler 1985d: 148) (s2.6 & s3.1.5) studies of code-switching behaviour around the world have established that multilingual communities can combine languages in a number of different ways, and debate continues within the field of code-switching research about the development of novel language varieties, mixed codes, within such communities. This research is significant in that it highlights how some types of mixed language behaviour can produce new varieties of language, and therefore suggests a possible middle ground between the positions of Bickerton and Mühlhäusler on continua in Tok Pisin.

In the following section I discuss the terminological distinctions and definitions in 'code-switching' research (s3.2.1) before considering the status of CS as variation in language behaviour (s3.2.2). A review of the major trends and paradigms in CS research is presented (s.3.2.3) and the role of language in forming and expressing issues of identity are discussed and shown to be significant in theories concerning the social motivations and explanation for CS (s.3.2.4).

3.2.1 Terminology and distinctions

Fundamentally, and as the discussion of the various research traditions, models, and typologies below will show, 'code-switching' is a problematic term that covers a variety of different types of multi-lingual behaviour. A fundamental division in research traditions is between examining the linguistic constraints on code-switching (s3.2.3) and exploring the social motivations for code-switching (s3.2.4). Broadly speaking, the distinction is between the structure of multilingual behaviour, and the meaning of that behaviour. Significantly, switching languages in either different situations or with different speakers or for rhetorical effect is seen as motivated, and is most often clear Alternation (Chapter 6) between languages at clause or sentence boundaries. More intimate types of switching through Insertion (Chapter 7) or through frequent and varied combinations of different elements (Chapter 8) have often been observed to occur without specific meaning. In the findings chapters I will explore the relationship between socially marked and unmarked switching and different types of behaviour.

In this section I explore the different descriptive terms and typologies used in CS research. Basic distinctions are drawn between 'intra-sentential' switching as switching languages within a sentence, 'inter-sentential' switching as switching languages between sentences, from one sentence to another, and 'tag-switching' as switching between a particular utterance and the tag, interjection, or closing phrase attached to it.

The terms 'code-switching', 'code switching', 'codeswitching', 'code mixing', and 'code-mixing' are used variably to describe either all types, or specific types of mixed language behaviour. For example 'code mixing' is a term which was initially adopted to make a distinction between cases of intrasentential switching which involve the introduction of many single lexemes from a second language (see e.g. (Kachru 1978; DiSciullo, Muysken and Singh 1986; Swigart 1992)) and the outright borrowing of these lexemes into a language.

Bokamba (1988) offered a distinction between code-switching (CS) as including full sentences ('intersentential' CS) and code-mixing (CM) as including 'intrasentential' CS and also bound morphemes.

In a study of bilingual "Panjabi/English" 'language Alternation', Moffatt (1990) develops distinctions between 'code-mixing' 'code changing' 'code-switching' and 'language switching', depending upon the position of the switch point in terms of the speaker's turn, and whether a switch was motivated by considerations of addressee interaction. While such distinctions are consistently elaborated in Moffatt's thesis, the application of terminology there is specific, I would suggest, to that study, and indicates different behaviour than in other work reviewed.

Gumperz (1982) discussed "conversational code switching" and provided germinal descriptions of the types and functions of individually meaningful switches. Auer and others working within the Conversation Analysis framework have defined 'code Alternation' and described 'mixed languages' characterised by certain types of regularised language combinations (Auer 1984, 1998a). In a later summary Auer (1999) distinguished Alternation from Insertion as structurally different, but observed the presence of both of these in his definitions of "codeswitching" and "language mixing": types of behaviour he distinguished on the basis of the meaningfulness of individual switches. In this account he redefined 'mixed languages' as "fused lects".

Muysken (2000) uses 'code-mixing' as a general term for different types of bilingual behaviour: Insertion, Alternation, and Congruent Lexicalization, that cover a range of activity from contrastive, locally meaningful juxtapositions, to dialect-standard variation.

Attempts to establish terminological clarity have been made. Myers-Scotton rejected the term 'code-mixing' as creating "unnecessary confusion" (1993a : 24). While applying Muysken's typology in my analysis, I prefer to follow Myers-Scotton's conventionalised usage and apply

CS (code-switching) as a neutral cover-term for all the different language mixing behaviours, while making specific distinctions and definitions made in existing research clear in the discussion below (s3.2.1.1).

The issue of linguistic borrowing (B) and its relationship to code-switching is one that recurs throughout the relevant literature, and is related to questions about the role of CS in language change. Such questions, as we shall see, are at the heart of this discussion: creole/superstrate CS takes place between two languages that are already much more closely related than many other language pairs.

3.2.1.1 Definitions

As research in bilingual communities has progressed, various models and hypotheses have been developed and tested with reference to different datasets. Poplack and her associates, working in the variationist tradition and with, initially, North American Spanish/English or French/English communities engaged with different types of behaviour to Myers-Scotton and her associates working in Africa on Swahili/English and various African languages. Within Europe, Auer is central in a tradition which looks for meanings in bilingual behaviour as being constructed within the conversational acts themselves, while Muysken has been involved with structural analyses of bilingual production.

Definitions of CS revolve around two broad conceptions of language combination. Gumperz (1982: 52) defined "conversational code switching" as involving the "juxtaposition" of speech from different grammatical systems or subsystems. His inclusion of 'subsystems' is significant in that it allows CS to cover switching between dialectal grammars, or alternative patterns of pronunciation. Milroy and Muysken describe the "alternative language use by bilinguals of two or more languages in the same conversation" (Milroy, L. and Muysken 1995) as central to research on bilingual language use. Auer uses 'code Alternation' as a cover term for code-switching and transfer:

> [A] relationship of contiguous juxtaposition of semiotic systems, such that the appropriate recipients of the resulting complex sign are in a position to interpret this juxtaposition as such.
>
> (Auer 1995: 116)

These definitions focus on the expressive ways in which using more than one language, or variety of language, in communication can convey meanings through the combination and contrast of languages: through 'Alternation' and consequent 'juxtaposition'.

However, the notion that all cases of bilingual production carry, or more properly create opportunities for the construction of (Wei 1998), meaning has been called into question. In fact, the assumption that the meaning and motivation for individual code-switches must be found in bilingual speech underlies some of the deepest divisions between different approaches to mixed language behaviour. Work on "un-marked" (Myers-Scotton 1993b) code-switching, on the Insertion (Myers-Scotton 1993a, 1997a; Muysken 2000; Myers-Scotton 2002), rather than the Alternation, of languages, and its relationship to borrowing between languages (Romaine 1989a: 131-145; Myers-Scotton 2002:153-162), has encouraged an appreciation of the different types of CS behaviour, the different functional and social values it may express, and the different roles it has in bilingual interaction.

Some linguists resist the application of the term 'code-switching' to cases of Insertion (Boeschoten 1998), but Myers-Scotton has extensively argued for the dual-activation of both language systems in bilingual production, and that the grammar of the second language is active even when individual words are inserted into constituents formed in the first, 'matrix', language (Myers-Scotton 1993a, 2002). From this standpoint Insertions are also 'switches' from one language system to another.

In addition, studies of 'mixed codes' (Lefebvre and Muysken 1988; Gardener-Chloros 1995; Savic 1995; Auer 1998a), characterised by particularly intimate and regularised forms of code-switching reveal the extent to which mixed language behaviour is a phase in the

development of new language varieties. This research adds detail to the behaviour underlying Thomason and Kaufman's (1988) work on linguistic borrowing and language change.

The terminological issues discussed above (s3.2.1) have arisen largely through the interaction of these different research traditions and the differing ranges of phenomena they have encountered. Bringing together this work in the late 20th century, Auer (1999) and Muysken (2000) composed typologies of language behaviour with reference to the types of contact and resulting behaviour.

Muysken (2000: 264) makes explicit an association between his model and the processes of contact-induced language change through borrowing and shift discussed by Thomason and Kaufman (1988). While Auer identifies both 'code switching' (CS) and 'Language Mixing' (LM) as types of behaviour which may include either Insertions or Alternations, and discusses the existence of structured 'Fused Lects' (FLs) (mixed languages such as Michif (Bakker and Mous 1994; Bakker 1997)) Muysken's typology distinguishes Alternation, Insertion, and free lexical variation within a mixed, congruent, grammatical structure (as in style or dialect switching), as Congruent Lexicalization. For Muysken, 'Code-Mixing' (CM) is the cover term for these different types of behaviour, equivalent to Myers-Scotton's 'Code-switching' (CS), and covering the behaviours Auer describes as 'language alternation'.

3.2.2 Code-switching vs. Variation approaches:

Attempting to apply a code-switching analysis to language that contained a mixture of standard and vernacular forms proved unproductive for Labov (1970: 35) when he considered different varieties of urban language in the United States (also see Labov (1972: 188-189) and Muysken (2000: 125)) . His concept of CS could not explain why a short passage of English might contain 18 switches between standard and vernacular forms. Labov developed the variationist paradigm for considering language behaviour which showed different subsystems in interaction in such a setting. This paradigm has been specifically applied to studies of

creole communities (DeCamp 1971; Bickerton 1975b; Sankoff, G. 1980a: 32) as well as to studies of bilingual behaviour (Sankoff, D. and Poplack 1981; Poplack, Wheeler and Westwood 1987; Poplack 1993; Poplack and Meechan 1995). Labov proposed the validity of the variationist model in describing the continuum of "variants [...] in 'free variation' within the same system" (Labov 1970: 34).

This notion of a single system underlying creole language production was contrasted by Labov with the two discrete systems in multi-lingual behaviour. Citing an example of Spanish/English code-switching, Labov characterised this as non-rule governed variation (Labov 1971a). He was establishing a fundamental principle distinguishing code-switching from variation when he wrote "to demonstrate that we have a true case of code-switching, it is necessary to show that the speaker moves from one consistent set of co-occurring rules to another" (Labov 1970: 34).

These separate systems are to be found in the behaviour of the speech community as a whole, rather than in the particular idiolects of individuals, who may command only a particular range of variables. However, taking this into consideration along with Auer's (1995: 116) observation that it is through the 'juxtaposition' of systems that separation is demonstrated reveals, I would suggest, a certain circularity: Labov diagnoses code-switching by the observable separation of different sets of rules (codes), while Auer diagnoses separate sets of rules by the observation of code-switching (for a specific recognition of this see Auer (1999: 313)).

Crucially, this circularity brings into question the definition of a 'code' or 'system' in language. As G.Sankoff generally observes:

> (1) Multi-code situations often appear to be marked by extremely frequent and rapid switching that, to put it bluntly, defies explanation, if by explanation one means accounting for every switch [...] (2) In many cases, analysts have experienced difficulty in attributing segments to one code or the other. This is especially true in cases involving diglossia, e.g., a creole and the standard European language from which the bulk of its lexicon is drawn.
>
> (Sankoff, G. 1980d: 32)

Where separate systems can be established, whether they are dialectal, sociolectal, or linguistic, then the analysis of variation may be better considered as code-switching. As Nagy and Meyerhoff state, identifying linguistic variables that are semantically equivalent across separate systems may be a "will o' the wisp" (Nagy and Meyerhoff 2008: 5).

Within a single 'language' system, 'style-shifting' is the term often applied to behaviour which brings more than one subsystem into play. It is often considered in terms of changes in register, or between different dialects of a language. Crucially, Auer's definition of 'code-Alternation' as a cover-term for code-switching includes the criterion of 'juxtaposition', highlighting the role of contrast in the construction of meaning in mixed language production, and on this basis he excludes 'style-shifting' as "a gradual transition from dialect to standard" from his definition (Auer 1995: 117). Writing in the same edition, Giacalone-Ramat makes a similar distinction, and suggests that "from a linguistic point of view speaking of switching makes sense only if the two systems are independent" (Giacalone Ramat 1995: 46-7). This strict definition is clearly not applied by other authors, who make explicit the similarity between inter-language behaviour ('code-switching') and intra-language behaviour ('style-shifting') (Hymes 1967: 9; Gumperz 1968: 381; Sankoff, G. 1980d: 29-30). Romaine is similarly specific, stating:

> [m]ixing and switching for fluent bilinguals is thus, in principle, no different from style-shifting for the monolingual. Bilinguals just have a wider choice ...
>
> (Romaine 1989a: 132)

As with Labov's work, these differences depend upon whether the language systems, at whatever level, can be seen as part of a single whole system featuring 'variation', or as discrete systems, 'codes', interacting. Significantly, Giacalone-Ramat adopts her definition of

'code-switching' in discussion of her data of mixed Italian standard/dialect language on the basis that the standard and dialect are distinguishable systems, even while acknowledging that "a continuum with fuzzy boundaries" (Giacalone Ramat 1995: 46) forms from the many gradients between sets of varieties that are broken up into discrete points and structures only through the process of analysis. Her argument is that the notion of discrete systems in language behaviour is effectively a construct of analysis, rather than a reality of behaviour (Gardener-Chloros 1995: 86).

Giacalone-Ramat's distinction is based on the recognition that shifts in the formality of language, 'style-shifting', are possible within both standard Italian and Dialect 'systems', which exist in diglossic relations. The informal-formal (subsystems) axis is therefore independent of the standard-dialect (systems) axis, and so she applies separate definitions of code-switching and style-shifting to maintain this distinction in the multidimensional linguistic space of her community. In this her definition of CS overlaps with Auer's definition of style-shifting along a standard-dialect continuum.

Underlying the differing definitions of style-shifting is the reality of speaker behaviour: the specific expressive and structural behaviours used by speakers when they seem to move along each axis may, in fact, be similar. Winford points out in the context of creole continua that such behaviour may "provide insight on the processes of change and shift that have always operated in these situations" (Winford 1997: 274). The fundamental principle of discreteness in distinguishing CS from variation is characteristically problematic.

3.2.2.1 Linguistic Variables, Multilingualism, Creoles and Continua

The notion of the linguistic variable is a powerful tool for understanding and mapping how linguistic rules spread throughout a community (Sankoff, D. 1978: xviii). Probability analyses can be used to model patterns of variation against social factors (Paolillo 2002: 3), and this

can reveal patterns of relationships between social and linguistic factors. However, the implicational scales of variation that underpin continua analyses rely on a uni-dimensional axis of variation, and this has been called into question by Rickford's work on creole varieties in Guyana (Rickford 1987: 23). Furthermore, the variationist approach may be inadequate to describe the realities of situations where "the linguistic heterogeneity present is part of an on-going linguistic change" (Kay 1978: 82), and so inapplicable in contexts of on-going and vigorous language change.

Fundamentally, then, the application of variationist analyses to multilingual contexts has been called into question, and the role of code-switching analyses in creole/superstrate contexts has likewise been questioned. Consequently, it appears that the characterisation of the Melanesian situation as involving code-switching, and not variation within a single continuum-like situation, depends on the successful identification of separate systems for Melanesian Pidgin and the English standard, and on the absence of apparent intermediate varieties.

The variationist paradigm treats the occurrence of each variable as a single trial, and consequently neglects potential patterns of co-variation (Sankoff, D. and Laberge 1978: 119). Code-switching analyses are inherently sequential, however, and look at segments of language that may contain different variants in terms of their relationship to each other. Muysken (2000: 123-126) discusses how some forms of code-switching behaviour can actually create patterns of variation traditionally treated under the variationist paradigm precisely because they are dependent on a shared underlying system. As a result, the empirical question of whether evidence of shared grammatical structures is available becomes fundamental not to distinguishing single-system 'variation' from multi-system 'code-switching', but to distinguishing between different types of code-switching (see s3.3).

Various arguments for potential continua of code-switching behaviour (Auer 1998b: 16-20; 1999; Muysken 2000) (see s3.2.7) therefore provide a significant inroad into understanding

how multilingual behaviour may lead to the appearance of a single-system continuum in pidgins and creoles.

3.2.3 Code-switching : 'Juxtaposition' and Borrowing

This section discusses the origins of CS research, specifically, in observations of bilingual behaviour, and the consequences of a concept of the 'ideal bilingual' for definitions of CS (s3.2.3.1). The contributions of Poplack (s3.2.3.2) and Myers-Scotton (s3.2.3.3) are particularly considered as introducing an appreciation of the regularities in bilingual speech, as well as expanding an appreciation of the common features of actual bilingual behaviour (s3.2.3.4).

3.2.3.1 Bilingual behaviour, 'Interference', and Borrowing

The earliest work on CS emerged from studies of bilingualism, such as Leopold's record of his daughter's acquisition of English and German (Leopold 1939-49) and Weinreich's study of Swiss bilinguals (Weinreich 1953). These highlighted what was seen as the 'problem' of interference. Focussed around issues in the acquisition and use of 'correct', fluent, standard forms of language, Weinreich observed:

> [A] bilingual's speech may suffer from the interference of another vocabulary through mere OVERSIGHT; that is, the limitations on the distribution of certain words to utterances belonging to one language are violated. In affective speech, when the speaker's attention is almost completely diverted from the form of the message to its topic, the transfer of words is particularly common.
>
> (Weinreich 1953: 60)

He provided a definition of the "ideal bilingual" as someone who clearly distinguishes between languages, and maintains a separation between them:

> [T]he ideal bilingual switches from one language to the other according to appropriate changes in the speech situation (interlocutors, topic, etc.), but not in an unchanged situation, and certainly not within a single sentence.
>
> (Weinreich 1953: 73)

Haugen (1950) too distinguished interference from code-switching, which he conceptualised as maintaining the distinct character of separate languages (Gardener-Chloros 1995: 74). In his study of Norwegian-American language use and acquisition of English he established an influential trope in studies of bilingualism:

> Except in abnormal cases speakers have not been observed to draw freely from two languages at once. They may switch rapidly from one to the other, but at any given moment they are speaking only one, even when they resort to the other for assistance. The introduction of elements from one language into the other means merely an alteration of the second language, not a mixture of the two.
>
> (Haugen 1950: 211)

This statement contains three key assumptions that have been investigated over the past 60 years. Firstly, that the 'free' use of two languages simultaneously is abnormal. Secondly, that one language is primary in any speech act. Thirdly, that language mixing does not take place: languages are changed through bilingual behaviour, but they are not combined into a third variety. Subsequent research has investigated constraints on multilingual behaviour, proposed models for the maintenance of a primary language in multilingual behaviour, and the possible emergence of mixed languages.

Such early descriptions of the "ideal bilingual" have indeed proved to be idealistic, and the broad category of 'interference' has been shown to include a range of behaviours, the diversity of which has led to different definitions, models, and approaches to bilingual behaviour being developed. Furthermore, recent theoretical revisions of the concept of 'language' encourage us to acknowledge that the notion of a monolingual 'standard' form for bilinguals to learn is essentially fictitious: 'languages' are collections of styles, registers, dialects, genres, and practices (Blommaert 2010).

Work by Poplack, D.Sankoff, and their associates (Poplack 1980, 1981, 1987; Sankoff, D., Poplack and Vanniarajan 1990b, 1990a) took the distinction between CS and interference as the basis for their structural and sociolinguistic descriptions:

> What is important is that this phenomenon [i.e. code-switching] be clearly distinguished, first conceptually, and then operationally as much as possible, from all other consequences of bilingualism which involve not alternate use, but the truly simultaneous use of elements from both codes... Not least important, all of these phenomena should be distinguished from speech errors which involve elements of both languages, and which may be properly considered 'interference'.
>
> (Poplack 1987: 72)

Here, making clear distinction between inter-language borrowing (B) and CS as the "alternate use ... of elements from both codes" was vital. Where a bilingual interspersed their speech with words from a second language (L2), as commonly occurs, Poplack established phonological, morphological, and syntactic integration as the criteria to distinguish material which had been borrowed into the first language (L1) and could conceivably be part of monolingual competence within that speech community (Poplack 1980, 1981). This permitted the exclusion of fully-integrated L2 material as B forms, outside the purview of CS theory, and was independent of any requirement to establish community-wide acceptance of the L2 material as part of the L1 system. Under this definition, Poplack maintained the separation of CS as 'ideal' bilingual behaviour from 'interference', including the inclusion of 'borrowed' vocabulary (Weinreich 1953: 60).

However, such simple diagnostic criteria to distinguish code-switching from borrowing are not adequate. Romaine observes the common assumption that bilinguals borrow material or switch in order to fill lexical gaps in a language is, in fact, incorrect in the case of fluent bilinguals (Romaine 1989a: 132). She argues that the most common function of CS is to repeat the same thing in both languages. In Papua New Guinea functional repetition in discourse has been frequently observed between English and Tok Pisin, especially in broadcast and print media, in terms of the establishment of synonym pairs (Mühlhäusler 1985h: 250-1; Mihalic 1990: 270-271). Romaine highlights that a distinction between CS and B forms on the basis of integration has its roots in Haugen's proposed continuum of bilingual phenomena in terms of code-distinctiveness, with borrowing representing minimal distinction, code-switching maximal distinction, and 'interference' reflecting something intermediate

between the two (Haugen 1956). Myers-Scotton (1992) specifically identifies the problems with the simple exclusion of B forms from CS analyses, especially those that assume that single-occurring L2 words are automatically borrowings (Reyes 1976), and points to Pfaff (1976, 1979) and Poplack (1980) as raising the problem of distinguishing the two phenomena.

Poplack's solution, 'integration' as a diagnostic, is expressly questioned by Myers-Scotton (1992: 33), who highlights that though the majority of single-word forms in most datasets do show higher levels of phonological integration than multiple-word L2 forms, "one must conclude that far from *all* B forms can be distinguished from *single* CS forms on the basis of their phonological integration into the ML" (Myers-Scotton 1992: 31).

Recourse to community-standards is not a simple solution either: Gardener-Chloros (1991: 162) established that assessments of the status of L2 material as B or CS forms may vary widely, and while monolingual acceptance may be an ideal criterion, both G.Smith (2002: 205) and Franklin (1990: 335) in regard to Papua New Guinea and Gardner-Chloros (1995: 74) in more general terms highlight how this may be inapplicable in some communities.

Myers-Scotton does suggest that a distinction between B forms and CS forms may be statistically measurable (Myers-Scotton 1992: 31, 33, 35), and Poplack and Meechan provide an indication of how such patterns may be observed (Poplack and Meechan 1995), however, establishing such distinctions has proved overly problematic for most researchers. Commonly, a frequency-diagnostic is applied, with recurrent forms judged to be borrowings, while infrequent forms considered CS. Myers-Scotton (1993a) applies this criterion as does Romaine in her consideration of children's Tok Pisin (Romaine 1992: 232). However, as G.Smith (2002: 205) observes in his work on Tok Pisin, very large corpora are required to make frequency-distinctions work.

3.2.3.2 Poplack : Constraints and Nonce Loans

Developing from the closed definition of CS as separate from many bilingual behaviours, Poplack's work proved germinal to the development of structural theories of CS, and with D.Sankoff she proposed two constraints on bilingual language in their 'Formal Grammar of Code-switching' (Sankoff, D. and Poplack 1981): the Equivalence Constraint, and the Free Morpheme Constraint (see Chapter 6). However, the frequent presence of L2 forms that showed switching between morphemes in violation of the Free Morpheme constraint, forms where morphemes from the L1 were used on material from the L2 without a complete adaptation of the pronunciation of the L2 material to match L1 phonology, compelled her to adopt the concept of the 'nonce loan' from Haugen (1950): as a L2 'borrowing' that is integrated syntactically and morphologically with the L1, but is unintegrated phonologically (Sankoff, D., Poplack and Vanniarajan 1990a).

This category of 'nonce loan' retains currency in considerations of language mixing, though it remains problematic in various datasets (see also Chapter 7). The criterion of integration is not always sufficient to distinguish a nonce CS form from a borrowing. Poplack et. al. suggest that Finnish/English bilinguals implement nonce borrowing for all English nouns (Poplack, Wheeler and Westwood 1987). In Spanish/English switching from the United States Sobin (1976) indicates that some phonologically integrated English items are switches and, inversely, Elias-Olivarse (1976) indicates that unintegrated forms may be borrowings. In their study of Tamil, D.Sankoff, Poplack and Vanniarajan (1990a) rely heavily on the morphological and syntactic framework to define any lexical material that is not native to that framework as a borrowing.

Romaine (1989a) distinguishes borrowing as a phenomenon that forms a continuum from a restricted list of established, integrated, widespread and accepted new words to an unrestricted list of 'nonce' loans. This contrasts with code-switching behaviours that do not form a

continuum: smooth intra-sentential switching and functionally marked, flagged, switching are clearly different.

Yet this distinction is not necessarily supported. Pfaff showed that the morphological integration of verbs is quite common for bilinguals (Pfaff 1976), and Myers-Scotton (1993a) reached a fundamentally different conclusion to D.Sankoff et al (1990a) and Romaine (1989a).

3.2.3.3 Myers-Scotton : Matrix Languages and Integration

Myers-Scotton used the morphosyntactic structure of an utterance to determine the 'matrix' language (ML), but considers all non-native, 'embedded language' (EL) material integrated within it, regardless of the compatibility of its phonological form, as CS (see also Chapter 7). In this, she picked up on several of Haugen's (1950: 211) germinal distinctions. Though she shows conclusively that the use of material from more than two languages is not 'abnormal' in many communities, her notion of the 'matrix' language is an extension of enquiry into Haugen's second and third claims (above) that bilingual speech is essentially the modification of a single language. While Myers-Scotton and Jake (1995, 2000) effectively argue that both languages are on some level active for bilingual speakers involved in CS, and Myers-Scotton (2002: 157) is explicit about the centrality of dual-activation to accounts of code-switching, they maintain that CS is characterised by a fundamental asymmetry in the roles of the different languages involved (Myers-Scotton and Jake 2009: 339).

Under the Matrix Language Frame (MLF) model (Myers-Scotton 1993a) one language has dominance over the utterance, and sets the morphosyntactic frame into which 'content morphemes' from different languages may be inserted. Following the 4M morpheme-classification schema (Jake and Myers-Scotton 1997; Myers-Scotton and Jake 2000) the categories of morphemes which may be included from the different languages involved in CS are carefully defined as including, in addition, 'early system morphemes', such as plural

affixes, and 'late bridge system morphemes', such as some uses of conjunctions. Only 'late outsider system morphemes', such as the copula or other structural features and some uses of prepositions and pronouns, are excluded from Insertion. Where they occur, under the modified MLF model (Myers-Scotton 2002; Myers-Scotton and Jake 2009), they activate the EL System Morpheme Principle, and must be located in an Embedded Language Island, a whole-constituent switch to the EL. This model, and its relevance to Tok Pisin/English data, will be discussed in more depth in Chapter 7.

This concept of a morphosyntactic 'base' language can be seen also in D.Sankoff et. al. (1990a), and it remains superior to other diagnostics such as word order (Nishamura 1985, 1986), volume of speech, or automatic distinctions based on assumptions about the expected, 'unmarked' (Myers-Scotton 1993b, 1997c, 1998b) language in a particular context. In Chapter 7 the morpheme-class distinctions on which Myers-Scotton's MLF model is based are introduced prior to the discussion of my findings of Insertion.

Other observations of bilingual behaviour have highlighted cases where a 'base' or 'matrix' language is not so easy to detect.

3.2.3.4 'Mixing' Languages : the Range of Bilingual behaviour

Muysken (2000) presents an explicit critique of the definition of 'nonce loans' in Tamil, and argues for a typology of CS behaviour that itself forms a continuum from strict intersentential Alternation, via morphosyntactic integration and borrowing, to a mixed grammar. Auer (1998b, 1999) also provides a description of a continuum of CS behaviours (see s3.2.7). Heath (1989) shows that CS is a route to borrowing in Moroccan Arabic. As Poplack admits (1985), it is in practise very difficult to distinguish small amounts of L2 material found in bilingual speech as being either CS or B forms, and the smaller the switched constituent, the harder this is.

Gardener-Chloros cuts to the heart of this debate about the distinction between borrowing and code-switching:

> The reason why this question has so often been raised is partly that in many communities, single-word switching is the commonest kind, and partly that any bilingual data-set contains examples of loans alongside instances of code-switching. The researcher transcribing and analysing code-switched data therefore inevitably has to face the problem of drawing the line between the two categories.
>
> (Gardener-Chloros 1995:73)

She further identifies three criteria which have been shown not to enable the categories to be distinguished: morphophonemic integration, replacement of 'core' lexicon, and restriction to particular grammatical categories. I have considered the first, above, with reference to the different approaches to similar data made by Poplack and by Myers-Scotton. The second is also considered by Myers-Scotton (1993a) who distinguishes 'cultural' loans (s7.2.1), material expressing new objects or concepts, from 'core' loans (s7.2.2), borrowing which replaces existing native synonyms. In terms of the development of Tok Pisin both forms of borrowing have been recorded, though attitudes to the different types of borrowing are often quite different. In regard to the third distinction, attempts to show that only particular categories of grammatical material, e.g. nouns, thematic-role assigner/receivers, discourse tags, can be involved in intra-sentential code-switching have been consistently faced with evidence from bilingual communities where unlisted material, e.g. conjunctions, pronouns, prepositions, are switched, and may become established borrowings (s7.3.4.3).

Myers-Scotton and Jake have modified the former's Matrix Language Frame model in terms of a Four-Morpheme (4M) type distinction (Myers-Scotton and Jake 1995, 2000, 2009), and consistently argued that grammatical functions are significant in limiting CS behaviour, developing the concept of a Uniform Structure Principle underlying bilingual language production (Myers-Scotton and Jake 2009). These accounts attempt to provide an explanation for the switching of 'system' (function) morphemes, grammatically active, structural material, between languages: the 'mixture' that Haugen (1950: 211) stated did not occur.

Haugen's study identified bilingual behaviour that can be characterised as the transfer of vocabulary from one language to another, suggesting that borrowing can lead to relexification and the consequent appearance of new varieties featuring L2 vocabulary within structures of an L1 grammar:

> In becoming bilingual within the American cultural environment they (i.e. the speakers of Norwegian) were forced to modify their Norwegian if they wished to continue using it. At practically every point they maintained the basic phonetic and grammatical structures of their native dialects, but they filled in the lexical content of these structures from the vocabulary of English.
>
> (Haugen 1969: 71)

This suggests that Haugen's speakers' 'English' speech was a form of Norwegian with heavy, fully integrated, borrowing from English.

In terms of pidgin and creole language development, the concept of a 'superstrate' contributing lexicon to fit within a substrate frame presents a parallel case (see s3.1.1). In research on Tok Pisin (see s3.1.4), lexical differences have been key to distinguishing different varieties, and to Mühlhäusler's variously expressed concerns about 'anglicisation' of the national language (Mühlhäusler 1985a, 1985h, 1985e, 1985b). A broader concept of code-switching and borrowing than strictly distinguished Alternation of languages, or of the borrowing of single words, may be necessary to interpret the range of behaviour present in PNG today.

In her introduction to a special issue on 'codeswitching' of the Journal of Multilingual and Multicultural Development, Eastman noted that the collected papers agreed that "efforts to distinguish codeswitching, codemixing and borrowing are doomed" (Eastman 1992: 1). She observes:

> [T]he study of loan words *per se* out of context is a relic of the past. Attention primarily to urban settings where people from diverse linguistic and cultural backgrounds regularly interact makes it abundantly clear that in normal everyday conversation, material from many languages may be embedded in a matrix language regularly and unremarkably.
>
> (Eastman 1992: 1)

The frequent and fluid introduction of material from many languages may occur alongside the strict juxtaposition or 'Alternation' that early definitions of bilingual behaviour sought to characterise as code-switching in contrast to other forms of behaviour. The consequence of this awareness is a recognition that distinguishing CS and B forms is a social, rather than grammatical, issue.

Following on from D.Sankoff and Poplack's early proposed constraints on the 'code-switching grammar', DiScuillo, Muysken, and Singh (1986) proposed additional constraints on CS from within the Government and Binding model of syntax (s8.2.7), which provided an additional theoretical paradigm for the investigation of code-switching. Yet this approach too sought to explain CS as a process of Alternation between different, and discrete, language systems, with only one language being spoken at a time. The status of nonce loans (s7.2.3.3.1) as standing between CS and B forms proved problematic for linguists searching for a complete model of all CS.

The discovery of bilingual communities where spontaneous combinations of "elements of both languages" were consistent characteristics of the way people spoke (see, for examples, the papers in Auer (1998a)) eroded the description of such behaviour as "speech errors" (Poplack 1987: 72) and an awareness of bilingual communities' ability to develop shared phonological systems for the languages they used, to speak both languages with a similar accent, removes, I would argue, the single pillar of phonological integration from the nonce loan/borrowing distinction (see s7.2.3.3).

Myers-Scotton (1993a) highlights the volume of evidence for situations where bilingual communities do not maintain the kind of 'ideal' discreteness between the languages they speak, where distinguishing borrowings from nonce loans is not possible on Poplack's criteria, and where counter-examples to proposed grammatical constraints abound. As Gardener-Chloros observes:

> [C]ode-switching should instead be considered as a much broader, blanket term for a range of interlingual phenomena within which strict alternation between two discrete systems is the exception rather than the rule
>
> (Gardener-Chloros 1995: 68)

Communities where such standard-focussed behaviours take place "are outnumbered by those where the code-switched passages *do* show departures from the two monolingual norms" (Gardener-Chloros 1995: 72).

These 'departures' may move beyond single word 'borrowings' to larger constituents, and the analyses of such 'Insertional' code-switching are the subject of Myers-Scotton's models (see Chapter 7). In some cases, however, they may also indicate a change to the grammatical structure of the language(s) involved: neither the overt juxtaposition of sentences or constituents in different languages nor the integration of words or larger constituents from one language 'into' the system of another, but the production of language based on a mixed system (s3.2.5) (as discussed for Tok Pisin by Mühlhäusler (1985d: 80; 1985h: 248)) as Muysken's (2000)'Congruent Lexicalization'.

3.2.4 Code-switching and Social Meanings

In his discussion of language variation and change Milroy (1992) demonstrates how the choice of particular varieties of language relates to the construction and display of social identity. He explicitly argues that it is speaker interactions in conversational settings that are "the locus of change" (Milroy, J. 1992: 18), and that it is speakers who initiate change. Taking the notion of the linguistic variable as a fundamental, his analysis shows how social identity is negotiated through the adoption of particular forms within the speech community. Changes take place as individuals negotiate a consensus on the norms within the groups that are socially significant for them.

Discussing Labov's treatment of code-switching and variation, Gumperz highlights "listeners' ability to assign speakers to social categories" (Gumperz 1982: 70) as a significant

sociolinguistic factor underlying regular interpretations of mixed language behaviour. Interpretations of social position and identity are possible even when the arrangement of language elements in a particular passage of speech appear irregular and unpredictable. Switching behaviour can act as a "contextualisation cue" (Gumperz 1982: 132-35) by which speakers signal social and situational factors through reference to conventions of language use (Wei 1998: 163). Language choice and the switching behaviour itself provides an additional dimension of variation which can be used to index identity within the speech community.

In Papua New Guinea, this articulation of the relationship between language and identity has been observed at all levels of society (see s2.6). The following section considers research into the relationship between multilingual behaviour and identity, with particular focus also on variation in pidgin/creole language communities.

3.2.4.1 Diglossia and Meaning

The conventions of language use (Wei 1998: 163) may be related to diglossia within the community, or may extend beyond such generalised associations. Britto's (1986) excellent analysis of diglossic theory shows that Fergusson's early formulation of 'diglossia' as referring to standardised relations between varieties of a language such that separate varieties play separate roles, being used in different situations consequently carrying different social associations and meanings (Ferguson 1972 (1959): 232), has been effectively extended by Fishman (1969, 1972a, 1972b, 1972c, 1972d (1962), 1976), to cover also the presence of functional complementarity (Timm 1981) between different languages, dialects, and styles.

Essentially, diglossia is a social concept, covering situations where particular codes, at whatever level, are associated with particular contexts or functions, and it differs from bilingualism, per se, where different 'language' codes may cover the same functional areas (Ferguson 1962: 174). The varieties in diglossic relations may differ in phonological, lexical,

or grammatical features and constructions. They may be as similar as different styles (formal/informal), or as different as distinct languages (Britto 1986: 302).

Closely associated with accounts of diglossia are categorisations of a 'High' (H) and 'Low' (L) code distinction. In Fergusson's work, prestige varieties of a language, typically with a standardised written form, used in contexts of power and authority, such as classical Greek and Classical Arabic, would be the H-code, while vernacular varieties, typically conversational, would be characterised as the L-code. In situations, such as the British Raj, where an H-code (English) was imposed from outside the language community, H/L code distinctions could be associated with identity: as 'they' and 'we' codes (Gumperz 1971b; 1982: 66-67). Specific codes are associated with specific 'domains', and studies of language shift have associated changes in these associations with changes in the linguistic makeup of a community (Gal 1979). In Papua New Guinea the relationships between indigenous vernacular languages, Tok Pisin, and English have been described as diglossic and there is evidence that such characterisations are being transformed (see s2.6). Elsewhere, in typical diglossic situations, 'mixed' H/L varieties have been identified (Householder 1962: 27), indicating that a notion of absolute, fixed functional separation between different languages, dialects, or varieties is not universal.

Myers-Scotton (1986; 1993b: 128) points out that if strict diglossia was the rule, then CS within the sentence would not be possible and 'unmarked' CS should not occur. The presence of such switching within the sentence in many communities shows that Fergusson's strict diglossia is an ideal, and that there is less clear separation between languages in most multilingual communities.

3.2.4.2 Meanings in CS : CA and the Markedness Model

A Conversation Analysis (CA) approach to multilingual behaviour and code-switching may focus on the meanings 'brought about' through the speech acts themselves, rather than the

meanings 'brought along' into the speech act from community expectations (Sebba and Wooton 1998; Wei 1998). While informants in bilingual conversations bring social associations into play, it is only from within the situational context of the speech acts themselves that the meaning of those acts can be determined. Auer used the tools of conversation analysis to show that bilingual conversations generated and developed the frames of relationships between language and meaning dynamically (Auer 1984). Essentially, the situation-bounded rules of diglossic relations could be used metaphorically by bilinguals, and code-switching itself as an expressive tool in conversation, rather than as something that was required by conditions external to the conversation.

The CA approach is explicitly contrasted with Myers-Scotton's (1993b, 1998b) social model of the motivations for code-switching, the 'Markedness Model', by the authors in Auer's (1998a) collection. Myers-Scotton's model more explicitly draws upon diglossic expectations and social patterns of language use, building on a concept of the 'expected' language in the specific context of a particular speech act to show how bilinguals can make 'marked', unexpected, choices in order to convey specific meanings, and to make subtle social negotiations in their conversation. Essentially, languages, under Myers-Scotton's model, are associated with particular social identities, and so with particular sets of rights and obligations within the social context. Language choice in bilingual conversation can work to manipulate these rights and obligations. The CA approach abjures the external ascription of specific meanings to specific languages, and offers an alternative analysis of conversational code-switching where the language-identity associations are created within the speech act rather than being drawn in from outside. Nonetheless, markedness and the local meaning of CS remain key concepts.

The Markedness Model allows Myers-Scotton to identify situations where CS is the 'unmarked choice', where CS does not work to negotiate meanings within the conversation.

Compellingly, this identification can be seen to cover Labov's examples of unmotivated CS, as well as frequent switching in other studies (1971a, 1971b).

In contrast, Auer specifically points out marked, meaningful switching is the only way to distinguish 'codes':

> [H]ow can we, the analysts, prove that a given set of co-occurring linguistic features is perceived by participants as a distinct code? The only valid answer is: by showing that switching between this set and another is employed in a meaningful way in bilingual conversation.
>
> (Auer 1998b: 13)

Consequently I distinguish two valuable trends in the analysis of CS behaviour.

On the one hand are concepts of all mixed language behaviour as being CS, from borrowing to juxtaposition. From this perspective there is a distinction between behaviour which is socially marked as meaningful, and behaviour which may be a marker of social identity though individual switches are unmarked.

The other approach offers concepts which distinguish CS from borrowing or interference phenomena. This places meaningfulness at the centre of our understanding of CS, and considers 'unmarked' CS as characterising a 'mixed code' (Auer 1998b) or "language mixing" (Auer 1999).

Multilingual speakers use the codes within their repertoire with an additional layer of expressiveness characterised by the distinctiveness of the varieties they command.

3.2.4.3 Acts of Identity

The concepts of 'brought about' meanings, 'marked' and 'unmarked' language choice, and the use of switching as a 'contextualisation cue' all combine in what Le Page and Tabouret-Keller call 'Acts of Identity' (Le Page and Tabouret-Keller 1985). Le Page & Tabouret-Keller's work stimulated a growing tradition in creole studies which show that the selection of varieties from a 'continuum' and code-switching between them (whether at the level of

juxtaposition or more intimately), nuances of pronunciation, lexical choice, and grammatical structure are all used to negotiate social identity and expressively define a speaker in context. A good example of this is Sebba's work on London Jamaican (Sebba 1993; Sebba and Wooton 1998), which illustrates and explores selections made from the range of West Indian creole and London and Standard English forms in their speaker's repertoire.

Outside of creole communities, Poplack observed relationships between ethnicity and type and frequency of CS (Poplack 1981), and associations between CS and negotiations of identity emerge from many studies of multilingual communities. In Papua New Guinea, Mühlhäusler observed that the selection of English words over Tok Pisin alternates is used as a marker of urban identity (Mühlhäusler 1985d: 140).

In her study of CS within a creole community in Suriname, Migge (2007) uses the Markedness Model to analyse choices made between different creole varieties without reference to the 'superstrate'. Her work focuses on the way different varieties are deployed to negotiate social interactions, and is comparable to the studies of CS between indigenous languages and Tok Pisin in Papua New Guinea (Sankoff, G. 1980d; Kulick 1992) (see also s3.2.5).

As Auer notes, the presence of such marked switching in the creole establishes that the varieties Migge describes are distinct codes for her speakers. The identification of expressive functions to CS behaviour is suggested by Auer as a diagnostic of whether a single system, or two systems are involved:

> This issue [of meaningful switching] is of particular interest in cases which, linguistically speaking, may be looked upon either as instances of code-switching between closely related varieties, or as one code showing internal variability.
>
> (Auer 1998b: 13)

Le Page and Tabouret-Keller observe that naming a variety of linguistic behaviour tends to give it an identity which may separate it from a group within which it was previously subsumed (Le Page and Tabouret-Keller 1985: 283). The inverse may also apply: a label may

subsume a range of behaviours. In the case of Tok Pisin, we may find, as Mühlhäusler has suggested, that the name 'Tok Pisin' covers a widely divergent range of forms (Mühlhäusler 1985c: 28).

Similarly, the debates on the use and meaning of 'code-switching' to describe linguistic behaviour (s3.2.1) negotiates distinctions between language mixing, lexical adoption, interference, and juxtaposition. As the label 'Tok Pisin' establishes a separation from English not entirely suggested by the term 'pidgin English', so too it covers a range of behaviours in different relationships to English and the indigenous languages of Papua New Guinea, relationships which may not be so easily described as diglossic, especially where the varieties involved do not adhere closely to monolingual norms (Sankoff, G. 1980d: 32).

Nonetheless, it is clear that particular varieties of language will always carry and create, suggest and reinforce, impressions of the identity of the speaker. Particularly in models of marked language Alternation (see s3.3.1) these values can be seen at work at both the pragmatic and the metaphorical levels in CS. In terms of unmarked switching, the social meaning of the mixture itself may be an index of a mixed, novel, multilingual identity of real significance in modern Papua New Guinea. The juxtaposition of language varieties may be at the level of an awareness of community-wide norms of behaviour, and a choice to mix language may carry meaning in its contrast with a decision not to mix (see s3.3.2). In a further development of the analysis of unmarked, mixed language, I will consider accounts of behaviour which explicitly rely on a fundamental combination of previously distinct 'codes' (s3.3.3).

3.3 Part III : An analytic Framework: CS typologies

Within the field of CS research I distinguish different areas of theory, relating to different behaviour. As discussed above (s3.2.1-s3.2.4), while the earliest accounts considered the social functions and meanings of inter-sentential switching, the Alternation or 'juxtaposition'

of language codes, intra-sentential switching (within the sentence) required the development of structural models, defining syntactic constraints on the combination of languages. Here I outline research findings on different types of CS which I will consider in the light of my data in Chapters 6, 7, and 8.

The different research traditions discussed in Part II (s3.3) above are summarised here to establish a typology of CS (s3.3.1-s3.3.4). A final section considers theories on a 'continuum' of CS (s3.2.7). Part III concludes with an appreciation of the variety of behaviour now accepted as CS, and considers its significance for the research questions about how CS behaviour impacts on the development of Tok Pisin.

3.3.1 Diglossia and Alternation

In diglossia, CS is principally Alternation of languages between sentences dependant on the situation, though it can also serve 'metaphorical' functions (Blom and Gumperz 1972). As Myers-Scotton (1986) observed (see s3.2.4.1), the presence of switching within the sentence indicates an erosion of diglossia, except perhaps in cases of the strict borrowing of lexis. Any such intra-sentential switching indicates a more complex relationship of language use to social situation: instead of employing different varieties in different settings, or with different groups of speakers, or during different types of activities, or potentially switching from one to the other in negotiating one's identity and role within a social scene, both 'codes' are used together to form part of the same speech act. This may be termed conversational code-switching, as opposed to situational code-switching.

Gumperz (1982: 59-99) gives examples of intra-sentential switches which show languages alternating in topic-comment, noun-noun complement, and predicate-predicate complement roles, the semantics of the sentences developing from both languages in combination. Additional constraints were proposed by Poplack and associates (e.g. (Poplack 1980; Sankoff,

D. and Poplack 1981)) and by Discuillo Muysken & Singh (1986) on where in the sentence such switches between languages could occur.

Fundamentally, Poplack's (1980) emphasis on linear equivalence developed a new perspective on code-mixing as Alternation between codes. Under her Equivalence Constraint (see Chapter 6) 'code-switching' was defined as occurring at sites of "co-grammaticality" where the order of sentence constituents in each language on each side of the switch point was equivalent (Poplack 1980; Sankoff, D. and Poplack 1981). She interpreted code-switching in terms of the juxtaposition of internally consistent sentence constituents from different languages and claimed that switching was impossible where word order equivalence did not exist. The Free Morpheme Constraint (Poplack 1980) (see Chapter 6) restricted language switching after bound morphemes. The concept of the 'nonce' loan, morphologically and syntactically integrated expressive borrowings, was introduced by Poplack and her associates to overcome counterexamples to the Free Morpheme and Equivalence Constraints. Resulting typologies of code mixing included switching under equivalence, nonce borrowing, constituent Insertion, and flagged switching (Sankoff, D. and Poplack 1981).

These proposed constraints address the syntax of code-switching. The conversational functions of CS were enumerated by Gumperz (1982: 75-81) as quotation, addressee specification, interjection, reiteration, message qualification and emphasis, and personalisation versus objectivisation, which can be included as a non-pragmatic factor alongside the 'we'/'they' code distinctions under diglossia. Of these, he described reiteration as the most common (Gumperz 1982: 59), and demonstrated how addressee specification, personalisation, and interjection work with meanings from a diglossic context. Auer added to this list functions such as supporting topic shift and comments, puns and language play, and changes of activity type (Auer 1995: 120). Table 3.1 provides a list of different categories of Alternation alongside these functions.

Table 3.1 : Types of Functionally Meaningful Alternation (ref. Gumperz (1982: 59) and Auer (1995: 120))

Meaningful Alternation	**Conversational Function** (ref. Gumperz/Auer)
Pragmatic	
Reiteration	reiteration
Adressee/Topic	addressee specification
Interjection	interjection
Qualification	qualification (parenthesis)
Demonstration	qualification (reiteration)
Declarative	emphasis (personalisation)
Focus	emphasis (topicalisation)
New Information	emphasis (topic shift)
Metaphorical	
Quotation	quotation
Topic-comment	topicalisation

These types of switching, whether pragmatic or non-linguistic, situational or metaphorical, are all meaningful, and customarily marked. Such switching indicates a separation between the 'codes' involved. Meaningful Alternations of this type occur in Auer's (1999) 'CS' (see s3.2.7), and are discussed in Chapter 6.

Not all intra-sentential CS of an Alternational A-B pattern (Muysken 2000) has local meaning. Where Alternation occurs without local meaning it falls within Auer's (1999) category of 'LM' (Language-Mixing). Myers-Scotton developed her Insertional account of CS to account for the appearance within sentences in one language (L1) of whole constituents from a second language (L2) (Myers-Scotton 1993a) separately from her work on the meaning of CS (Myers-Scotton 1993b). In some datasets this included the presence of bound L2 morphemes, in violation of D.Sankoff and Poplack's Bound Morpheme constraint. Often this embedding of one language within the other did not appear to convey particular situational or metaphorical meaning, and it required a different structural account to straightforward Alternation: Myers-Scotton's Matrix Language Frame (MLF) model of Insertion.

3.3.2 Dominant Languages: the MLF

The Matrix Language Frame (MLF) model is an argument "that codeswitching is not a helter-skelter production" (Myers-Scotton, 1993a: viii). Dealing with the structure of switching within the sentence where mono-directional Alternation is not an adequate explanation, the MLF is based upon the position that a morpheme-class distinction patterns CS production: the MLF defines structural constraints on intrasentential CS on the basis of an asymmetry between the accessing of 'system' (function) and 'content' morphemes. In the psycholinguistic philosophy of the MLF model "switching is a means to ['increase the flexibility of expression'] index the nuances of social relationships by exploiting the socio-psychological associations of the languages employed" (Myers-Scotton, 1993a: 1).

The MLF model suggests that one language, the 'Matrix Language' (ML) is accessed primarily for the functional (system) part of a grammar: the quantifiers; possessives; tense/aspect system; determiners; complementisers, for structurally assigned agreement; the copula; 'do' verbs; possessive 'of'; dummy pronouns; and parts of some idioms (stock phrases). This language sets the frame, the 'matrix', for an utterance. It is the language that utterance is 'in'. Content words: verbs; nouns; and adjectives, are accessed separately, and may come, in theory, from any language available to the speaker. There may be socially or psychologically pragmatic reasons for <u>not</u> selecting content words from separate codes (L1, L2, L3, ... etc.), but a grammatically well-formed example of CS will both be so <u>and</u> be judged to be so by bilingual listeners on the basis of the principal frame-setting (ML) system morphemes.

Myers-Scotton's conception of a Matrix Language (ML) and an Embedded Language (EL) assume both asymmetrical roles in constituent selection in mixed language production and a high level of typological difference between languages involved. By returning to a view of CS as Insertion, Myers-Scotton was able to renegotiate the difficult territory of nonce borrowings

(s3.2.3.2) by considering them to be mixed ML+EL constituents. Under her model, bilinguals' CS could be viewed as a mechanism for borrowing as switched constituents came to be an accepted part of monolingual production. She contributed to the debate on borrowing and language change by stressing that, since intrasentential switching was limited to content morphemes from the donor (EL) language(s) (L2, L3, etc.) only these would be borrowed into the recipient (ML) language (L1) over time.

The MLF model is judged here to suggest that fundamental changes in the structure of codes in contact are restricted to situations where the ML for mixed language production changes from one language to another. Myers-Scotton's various constraints and hypotheses (see also Chapter 7) limit the appearance of EL system morphemes to constituents acting as embedded islands. Essentially, the MLF model maintains a distance between codes based on the content/system morpheme distinction. In the argument for the later 4M Model composed by the same author and her associate (Myers-Scotton and Jake 2000) this simple division is moderated by an appreciation of different stages in the production of system morphemes, described as falling into three classes: early, late, and late outsider. These are distinguished on the basis of the features [+/- conceptually activated] and [+/- looks outside its own maximal projection].

Under this separate view of language production, it is proposed that the accessing of 'early' system morphemes along with content morphemes is possible in mixed language production because they express concepts fundamental to the sense of the utterance. This modification allows for the presence of structural features of an EL in an ML, and opens the door to language change at the structural level through strategies of Insertion.

3.3.3 Mixed Languages: Congruent Lexicalization

As the field of research into mixed language behaviour has expanded, however, it has become clear from some datasets that neither theories of code-switching as Alternation nor theories of

code-switching as Insertion completely account for all patterns of mixing. Muysken (2000) introduces a third type of behaviour, Congruent Lexicalization (Chapter 8), to account for circumstances where language mixing does not appear based on the Alternation between two grammars or on the Insertion of constituents from one grammar in another, but rather indicates the existence of a shared grammar which can deploy constituents from either code equally throughout the sentence:

> All categories can be switched in congruent lexicalization, including **function words**, since there is no matrix language determining the language of the function words.
> (Muysken 2000: 130) (emphasis in original)

In recognising a third type of pattern for mixed language behaviour, Muysken (2000) develops our understanding of the relationships between the datasets previously reported and social circumstances.

Muysken describes how issues such as levels of congruence and different routes for constituent borrowing combine to motivate the development of mixed grammars in some communities. This is of great significance to the study of Tok Pisin and its relationship with English throughout my Papua New Guinean mixed language data because it completes the association of code-switching with processes of borrowing and language change by indicating circumstances where grammars may converge to form new varieties.

Muysken (2000) lays the groundwork for an integration of the perspectives of Bickerton (1975a) and Mühlhäusler (1975) on the development of Tok Pisin. His typology of CS includes mixed language behaviour involving the use of both content and system morphemes from additional languages, and shows how CS can rely on, and produce, the appearance of shared grammars.

As Auer (1998b) observes, neither the socio-linguistic analysis of the meaning and functions of CS, nor the syntactic analysis of inter-sentential CS provides a full picture. His *Bilingual Conversation* (Auer 1984) argued that bilingual speech had structure that could best be

considered from within the conversation itself, independently of social or syntactic observations. Stroud (1998) questions this position with reference to the crucial role social values play in structuring CS in Gapun village in Papua New Guinea. In contrast, Wei (1998: 170) upholds it with reference to his work in Newcastle.

Myers-Scotton produced two key works on her Zimbabwean data: the syntactic model of *Duelling Languages* (Myers-Scotton 1993a) and the social (markedness) model of *Social Motivations for Codeswitching* (Myers-Scotton 1993b), reflecting an awareness that both approaches were necessary to cover the range of CS behaviour. In her later work, she makes a specific distinction between 'classic' and 'composite' CS (Myers-Scotton 2002: 105), in acknowledgement of the observations by Muysken and others of the 'mixed code' type of behaviour: CS which violates each of Haugen's (1950: 211) three rules (see s3.2.3.1 above). This distinction is a development of the earlier concept of 'changeover of matrix language' and an extension of her early concept of the 'mixed matrix' (Myers-Scotton 1993a). It is significant to note that Muysken too associates his 'Congruent Lexicalization' (Muysken 2000) with bi-directional switching: where speakers incorporate material from each language within the other. For Auer (1999), however, frequent Alternations and Insertions together indicate 'language mixing' (LM), but not necessarily a 'mixed code', by which he understands stabilised 'fused lects' (FLs) rather than the variable use of words from either language in a shared structure (see s3.3.4).

3.3.4 Conclusions : Different 'Types' of CS

Muysken (2000) introduces a typology of CS and argues that the different models in the literature stem from different types of behaviour. First, the 'juxtaposition' of codes without grammatical conflict can be considered Alternation for expressive and meaningful purpose, e.g. (Gumperz 1982). Second, the integration of L2 (EL) material within a 'base' or 'frame' of the L1 (ML) sentence is Insertion, e.g. (Myers-Scotton 1993a). Third, the combination of

L1 and L2 material without a hierarchical (ML/EL) relationship between them, consequent on grammatical convergence, is described as occurring through Congruent Lexicalization, e.g. (Muysken 2000).

Each of these types of behaviour may combine with, and come to indicate, not just different levels of language competence (Poplack 1980; Giacalone Ramat 1995), but also different identities.

While this diverse typology may seem to be an inelegant approach to mixed language, it is a natural consequence of analysts' inquiries into the different layers of bilingual behaviour (Sankoff, G. 1980d: 29). Such a typology has great value in a field of research where diverse terminology (s3.2.1) abounds, and where the same terms ('code-switching', 'codeswitching', etc.) are applied to behaviour which works differently on both social and structural levels. Auer (1999) discusses the presence of both Alternation and Insertion in his locally meaningful 'CS' ("Codeswitching") as well as in his less meaningful 'LM' ("Language Mixing") behaviour, with regularised close grammatical combination a characteristic of his 'FLs' ("Fused Lects").

As Gardener-Chloros (1995: 72) observes, a range of CS behaviours occur in any community, and consequently:

> [c]ode-switching should be viewed as an analyst construct rather than as an observable fact. It is a product of our conceptualisations about language contact and language mixing, and it is not separable, either ideologically or in practice, from borrowing, interference or pidginisation.
>
> (Gardener-Chloros 1995: 86)

An investigation into CS in a community should take into account the presence or absence of each of these different types of behaviour without assuming that they will be entirely distinct from each other.

3.3.5 A 'Continuum' of CS?

In addition to establishing typologies of bilingual behaviour, both Auer (1999) and Muysken (2000), and to a certain extent Myers-Scotton in her concept of the 'matrix language turnover' (Myers-Scotton 1993a), suggest an organisation of types of switching in relationship to each other and to the contact situations of the bilingual speakers who utilize them. For Auer this is represented by a monodirectional scale from CS→LM→FL, with individual communities stabilising at different points along the code-switching continuum (Auer 1998b, 1999). While Auer's typology allows for meaningful CS behaviour to involve either Insertion or Alternation, Muysken locates Insertion as the least intimate bilingual strategy:

> As the extent of language contact grows the type of mixing will shift from insertional to either alternational or congruent lexicalization.
>
> (Muysken 2000: 249)

Oesch-Serra (1998) and Maschler (1998) discuss the important and stable role of discourse markers and conjunctions in the 'mixed codes' they describe. As Auer summarises:

> It is tempting to analyse these discourse markers as discourse-related insertions. Yet, we may ask whether the discourse markers in the 'other language' are cases of code-alternation at all. In order to ascribe meaning or function to them not only qua discourse markers, but qua insertions, we would have to prove that this meaning is created by the usage of the 'other' language. Alternatively, and more plausibly, it may be the case that a mixed code has developed to which, diachronically speaking, language B has contributed discourse markers and possibly other metapragmatic structures.
>
> (Auer 1998b: 19-20)

Where stable rules have been established regarding the use of material from either language, Auer argues, a 'mixed code' can be said to have developed (Auer 1998b: 20), and Auer offers criteria for the assessment of such varieties . These are Auer's 'Fused Lects' (Auer 1999: 324), such as Michif (Bakker and Mous 1994), Copper Island Aleut , or Ma'a or, arguably, Tagdal (Benítez-Torres 2009, 2011). In considering the status of observed bilingual behaviour using the structural varieties of CS as per Muysken's typology, it remains important to retain Auer's concept of FLs and his definitions of the 'mixed code' (Auer 1998b: 20).

> On the surface, a FL may look similar to LM. Often, the difference becomes visible at a deeper grammatical level only. While LM by definition allows variation (languages may be juxtaposed, but they need not be), the use of one "language" or the other for certain constituents is obligatory in FLs; it is part of their grammar, and speakers have no choice.
>
> (Auer 1999: 321)

For Auer, LM is intermediate between meaningful CS and a stable mixed variety:

> [W]e may characterise a certain middle position on the continuum between interactionally meaningful code-alternation and a mixed code in the following way: there are numerous (nonce) borrowings which may cause surrounding materials to be switched as well; there are numerous turn-internal switches which, however, do not change the language-of-interaction (i.e. they are non-consequential on the level of language choice); and the distinction between insertions and code-switching thereby gradually loses its force.
>
> (Auer 1998b: 20)

To be clear: Muysken's 'Congruent Lexicalization', permitting relatively unconstrained incorporation of material from two or more language systems in a shared grammatical structure is not equivalent to Auer's FL, though it does resemble Myers-Scotton's 'composite CS' based on a 'composite Matrix language' (Myers-Scotton 1997b; 2002: 52) that emerges during her Matrix Language Turnover phase. Congruent Lexicalization is a potential feature of Auer's LM.

If there is an overlap between CS and the development of stable intermediate varieties between languages, then recognising different structural types of CS is vital: some types of CS keep the languages involved separate, while other types of CS bring them closer together. For Tok Pisin and English, 'pinpointing' CS (Mühlhäusler 1985d: 148) is no longer enough to establish there is no pidgin (creole) continuum: if the type of CS occurring can be identified as Muysken's Congruent Lexicalization, or the mixed language as an instance of Auer's Fused Lects, then this suggests a single underlying structure, neither Tok Pisin nor English, but something intermediate. In this way, identifying different types of Tok Pisin/English CS in PNG is an important step to exploring the existence of a mesolect, and investigating the formation of a continuum.

3.4 Part IV: Analytic Framework - examples[7]

In order to operationalize the analytic framework described above (s3.3), I conclude this chapter with some examples of bilingual behaviour that have variously been described as 'code-switching', 'codeswitching', 'code-mixing', or 'language Alternation' from communities around the world and compare these with examples from Papua New Guinea.

In relation to Tok Pisin, Romaine has observed:

> Since Tok Pisin, like most pidgins, has very little in the way of inflectional morphology, the criterion of morphological and syntactic integration will not be sufficient to distinguish borrowing from code-switching. Nor will word order serve as a sufficient criterion for determining a base language because both languages are SVO. I do not think there is any reliable basis on which to decide that these ambiguous forms belong to one or the other language. ... There is an inverse relationship between the degree of similarity between any two languages in contact, which will have the effect of maximizing the potential sites for code-switches, and the extent to which code-switching and borrowing can be distinguished.
>
> (Romaine 1989a: 139)

While I agree with her conclusion, the specific 'ambiguous forms' she gives as examples may be amenable to alternative analysis, as suggested below (s3.4.3).

The examples presented here illustrate the behaviours that underpin the typologies discussed in Chapter 3 (s3.3). They demonstrate, with reference to prior research, the foundations of the distinctions and discussions of my data made in Chapters 6, 7, and 8. Some examples have been reformatted for consistency with the conventions used elsewhere in this dissertation.

---

[7] Throughout this section the examples have been reformatted for consistency. When specifying the languages involved in each example, I have placed the first language in sequence (in Alternation) or the matrix language (in Insertion) before the other language involved. The formatting of the language names indicates which language is in normal font, and which in bold italics. Additionally, I have always maintained English in these examples as bold italics and Tok Pisin as normal font in order to remain consistent with the examples from my own data presented throughout this dissertation (see s4.3). Key items referred to in the body text are underlined.

3.4.1 Alternation: Situational and Metaphorical

Strict Alternation between languages occurs at sentence or clause boundaries and is frequently associated with meaningful effects. Blom and Gumperz distinguish situational switching (that required by the context and structured associations between languages and particular speakers or situations) and metaphorical switching (that drawing upon those situated values for expressive purposes) (Blom and Gumperz 1972). As discussed above (s3.2.4) situational switching is most commonly associated with diglossia. Gumperz gives the following examples of metaphorical switching between sentences:

(1) Speakers : Chicano professionals in California, exchanging goodbyes (***English*** - Spanish)

A. ***Well, I'm glad I met you.***
B. Andale pues

'(O.K. swell').

(2) Speaker : a college student in India, telling an anecdote (Hindi - ***English***)

Mai gəya jodhpur m~e ***There is one professor of Hindi there, he is a phonetician.*** To us-ne pronauns kiya əpne vais-se

'(I went to Jodhpur). There is one professor of Hindi there, he is a phonetician. (so he pronounced it in his own voice).'

(3) Family conversation in a Slovenian village in Austria talking about a visiting peddler (***Slovenian*** - German):

A. ***Totə kuarbcə yə mewa***

'(she had such baskets).'

B. ***Nɔ na ǰinyan***

'(no I don't believe it).'

C. Ya ya dim it di kɛrbalan

'(the one with the baskets).'

A. ***Vinarca yə ßoa***

'(she was Viennese).'

B. Na Di mit di kɛrbalan

'(no!) (the one with the baskets)?'

(Gumperz 1982: 59)

In Papua New Guinea, G.Sankoff identifies five similar cases of what she describes as a "gross kind of alternation" (Sankoff, G. 1980b: 43) in her examples of Tok Pisin/Buang CS. Kulick has various similar examples from his work in Gapun:

(4) Text 6.2 : ***Taiap*** – Tok Pisin

Takwat ɛnda. Yundɨ. Em wari bilong yu. ***Yuwɔn nu-nuk mɔsɛrkukubɛtəŋan.***

'What a joke. You're too much. That's your problem. You brought yours after I bought mine.' [i.e., you should still have some left]

(Kulick 1992: 195)

(5) Text 6.5 : ***Taiap*** – Tok Pisin

Ŋa ruru sɛnɛ ia kirawmbrɨ wakarɛ. ɛndɛkarɛ, ɛndɛkarɛ. [Turns to Masito] Mm. Masito. Kisim spun i go givim papa.

'These two poor kids I just don't know. Hungry, hungry. [Turns to Masito] Mm. Masito. Take the spoon and go give it to Papa.'

(Kulick 1992: 203)

Salisbury (1962) discusses the significance of immediate translation in public discourse among the Emenyo. In considering the relationship between Tok Pisin and English, Siegel presents a long text from the documentary film *Cowboy and Maria in Town* which shows similar clear juxtapositions between Tok Pisin and English sections, noting: "[a]s in most cases, it is easy to determine which grammatical system is being used." (Siegel 1997: 196).

From his corpus of children's Tok Pisin, G.Smith gives an example of such "gross … Alternation" (Sankoff, G. 1980b: 43) in switching from Tok Pisin to English:

(6) Tok Pisin – ***English***

Mipla sa kolim mipla yet – mipla ol stap we – ***we live where PNG begins and ends***

'we call ourselves, we live where – we live where PNG begins and ends'

(Smith, G. 2002)

In addition to switches between sentences, Gumperz also includes examples of switching within the sentence:

(7) ***English*** - German

Go and get my coat aus dem Schrank da

'Go and get my coat (out of the closet there)'

(8) ***Slovenian*** - German

Uzeyma ti kafe Oder te?

'(will you take coffee?) (or tea)?'

(9) ***Hindi*** - English

Jo wo əccha tičər hota he'll come straight to Delhi.

'(Anyone who is a good teacher) he'll come straight to Delhi'

(Gumperz 1982: 60)

G.Sankoff has a similar example from Buang:

(10) ***Buang*** - Tok Pisin

ŋau ti ŋmodo bai ol i kot stret long yu

'If you're the only one sitting down (i.e. doing nothing). They'll take you straight to court.'

(Sankoff, G. and Brown 1976: 303)

The examples (1)-(10) can all be characterised as Alternations: monodirectional switching from one language to another.

3.4.2 Insertion: Matrix Language

Of a different order is Insertion: behaviour which integrates material from one language (the Embedded Language (EL)) into the structure of another (the Matrix Language (ML)). As discussed above (s3.2.3), where such 'Insertions' are small, single words or phrases, the problematic area of distinguishing borrowing from CS arises.

Benthalia and Davies present the following example, where French NPs fit neatly into the Arabic sentence:

(11) Moroccan Arabic - ***French***

merra ***la plage***, merra z[-v] z[-v]ebel, merra ***la forêt***, kul merra w fin

'one day the beach, one day the mountains, one day the forest, every time somewhere different'

(Bentahila and Davies 1995: 316)

From PNG, (12)-(15) present examples of similar Insertions of Tok Pisin into indigenous *tokples*, while example (16) illustrates the inverse case in Gapun: of a Taiap word inserted into a Tok Pisin sentence:

(12) Text 1 : ***Abelam*** - Tok Pisin

bikplesbʌ ***vətʌkŋʌy anə dawly*** liklik hanwara ***nʌkŋ pwtʌkŋʌy anə warək***
village-in seeing we-2 descending small stream one crossing we-2 ascended

'[We] looked at the main village, came down, crossed a creek and went up again'

(Laycock 1966: 49)

(13) ***Buang*** - Tok Pisin

Orait, ***nga ti ken be, vu, vu*** as ples. Orait, ***su rek be*** winim ***ke, be*** winim olgeta direkta, ***ga be*** winim ol pepul, ***olo ba*** kamap ***vu bumbum re.***

'Now, this has to do with us, in the villages. And it's not too much to handle for me, or for the directors, or for the people, that it has to be taken to the white man.'

(Sankoff, G. 1980a: 42)

(14) Text 2.5 ***Taiap*** - Tok Pisin

Ŋa basket-***nɨ prukaku wakarɛ***

'I'm not working on a basket'

(15) Text 2.6 ***Taiap*** - Tok Pisin

Masitɔ warak-a-kawt-rɛ, yu passim maus ***-aw-tɛt***
Masito talk-be-HAB-when/if, you shut mouth be-2SG:IMPER

'When Masito talks, you shut up!'

(16) Text 2.3 Tok Pisin - ***Taiap***

Painim wanem long ***sapwar*** bilong me?!

'What are you looking for in my basket?!'

(Kulick 1992: 76)

Here, (15) is particularly notable as it shows the Tok Pisin VP *yu passim maus* marked with the Taiap imperative morphology. The presence of larger constituents from one language

within the structural 'matrix' of another is characteristic of 'Insertion' under Myers-Scotton's MLF framework (s3.2.3.3, also s3.3.2 below).

For Tok Pisin and English, Mühlhäusler (17), Nekitel (18), and G.Smith (19) illustrate similar Insertions:

(17) Tok Pisin - ***English***

Wanpela samting tu ya, sam pipel ol i ***politically minded*** na mipela sampela olsem yupela i ***bin manipulated by others*** ...

'And something else, some people who are politically minded and some people like you and me have been manipulated by others ...'

(Mühlhäusler 1985h: 271) (underline in original)

(18) Tok Pisin - ***English***

Olsem no i nap long Minista blong yut senisim ***policy*** blong yu no lukluk long desala na ***daivet***im mani i go ***it's worthwhile instead of giving*** i go ol i putim i go na ol i putim long IBD

'And so can the Minister of youth change your policy and consider this and divert the money to (where) it is worthwhile instead of giving it to them to deposit into an IBD.'

(Nekitel 1998: 158)

(19) Tok Pisin - ***English***

Ol ***militen*** kam stap long hap na - i no ***militen*** bat em mas ol raskol nambaut o ol laik ***treten***im ol wonem ia nambaut

'The militants came and stayed there – they weren't militants but they must have just been criminals or they wanted to threaten the what's its.'

(Smith, G. 2002: 203)

However, as with other language communities around the world, even in the published data on Tok Pisin/English CS not all cases are so clear cut.

Auer's (1999) typology characterises 'CS' as meaningful juxtaposition, whether it be Alternation between language systems, as in (1)-(10), or Insertions, as in (13) and (15) (though G.Sankoff, Laycock, and Kulick are generally in agreement with Salisbury (1962) that Tok Pisin is used with *tokples* deliberately and meaningfully in most contexts in PNG). LM ('language mixing') is the term Auer reserves for combinations, either Alternations or

Insertions, where the juxtapositions do not have local meanings (Auer 1999), arguably the case for examples (11)-(12), (14), and (16)-(19), all above, as well as (20):

(20) ***English*** - Tok Pisin

Ol Morobe i autim tiket bilongen long wanem ***on the grounds that*** em i no ***under-stand***im. Em i ***ignorant bikos*** em i no ritim dispel pepa.

'The people of Morobe district voted his party out on the ground that he didn't understand [the situation]. He didn't know what was going on because he hadn't read this paper.'

(Mühlhäusler 1985h: 271)

3.4.3 Grammatical Mixing : Congruent Lexicalization?

Congruent Lexicalization is a type of CS that relies on a shared, congruent, grammatical structure. This allows material from the languages involved to be combined without specifically complying with the grammatical structure of one particular ML.

Significantly for investigation of the relationship between Tok Pisin and English are constructions such as (21), from Siegel's long extract from *Cowboy and Maria in Town.* In this text Alternational switches from English to Tok Pisin are at first meaningful in the sense that they negotiate the language of interaction between the four speakers (see, for example (Auer 1984; Myers-Scotton 1993b)). As the exchange continues, however, such Alternations feature smaller and smaller constituents until the main speaker (Speaker 2) introduces the English conditional *if* as an inserted grammatical morpheme prominently in his Tok Pisin:

(21) From Cowboy and Maria in Town : ***English*** - Tok Pisin

Speaker 2: ***Okay, then you see, then*** yu lukim ***if*** mi kisim ol lain Rigo na go slip long Kainantu haus lain bilong yu, bai yu tok wanem?

'Okay, then you see, then you see if I take Rigo people and go and stay at your Kainantu clan's place, what will you say?'

(Siegel 1997)

The Insertion of grammatical morphemes as in (21) is indicative of close mixing, between the languages involved (Myers-Scotton 1993a; Myers-Scotton, Jake and Okasha 1996; Muysken

2000: Chapter 6). Similarly intense mixing can be seen in Mühlhäusler's example of the 'post-pidgin' stage of Tok Pisin, as in (22):

(22) 20 year old UPNG student : ***English*** – Tok Pisin

Nau dispela taim mi kam nupela yu, na mi no nap raun, na ***bikos*** mi laik traiim long painim ples bilong slip ya, ***bat*** laki, taim mi laik kam mi tokim brata bilong mi na em go toksave long brata long ***stesen*** ya na ring i kam long ***faders*** long hap hia long ***bukim sit*** bilong mi na ol i save olsem bai mi kam. ***So when I came***, mi kam stret pundaun long ***epot***, mi go lukim, ***I pay half the price***, mi kam stret ***suppose to be fifty*** na mi ***peim twentifaiv. I brought fifty Kina just to pay for that plane but*** ol i hapim prais na ol i givim ***twentifaiv*** na mi lus ***with twentifaiv***, o nau shit, mi amamas pinis.

'When I first arrived, I wasn't able to stroll around, because I wanted to try to find a place to sleep, but luckily, when I first thought of going there, I mentioned it to one of the (religious) brothers and he informed a brother on this station and he rang up the fathers over there to book a seat for me and so they knew that I was coming. So, when I went, I had just got off at the airport, I went to see them, I paid half price, it was supposed to be fifty and I paid twenty-five. I brought fifty kina just to pay for that plane, but they halved the price and they returned twenty-five and there I was with twenty-five, oh shit, I was really happy.'

(Mühlhäusler 1985d: 147-148) (underlining added)

As Mühlhäusler observes, this speaker uses English subordinating conjunctions *bikos* = *because* and *when*, the coordinating conjunction *bat* = *but*, and the English plural suffix *-s*, in addition to English lexis: *epot* = *airport* for *ples balus* 'airport'; *brata* = *brother* for *bruder* 'religious brother'; and *bukim sit* for 'to book a seat', where the Tok Pisin noun *sit* is 'remnants (ashes, leftovers), faeces' (Mihalic 1971). They also use the English counting system with *twennifaiv* for *tupela ten faiv* 'twenty-five'.

Mühlhäusler offers this example demonstrating the impact of English:

> [R]enewed contact with English has caused considerable changes in the character of Tok Pisin over a relatively short time
>
> (Mühlhäusler 1985d: 148).

Could these changes be a result of the third type of bilingual behaviour: the intimate grammatical mixing Muysken characterises as 'Congruent Lexicalization' (Muysken 2000), and Myers-Scotton as 'Composite Codeswitching' resulting from a 'composite Matrix Language' (Myers-Scotton 2002: 5, 20)?

Bickerton (1975a) asserted that the development of intermediate varieties between Tok Pisin and English was an inevitable stage on the road to a continuum. Romaine connects code-switching behaviours and the development of new varieties:

> Rural and urban Tok Pisin are not discrete varieties, but can be thought of as a continuum. In cases where Tok Pisin and English coexist as closely as they do in urban centres, and speakers code-switch between them, new intermediate varieties have arisen which contain compromise forms.
>
> (Romaine 1992: 321)

Romaine also observes:

> [I]n the intermediate varieties code-switching sites are blurred and it is difficult to distinguish borrowing, transfer, and code-switching[.]
>
> (Romaine 1992: 322)

She presents five examples from the colloquial speech of unmonitored Papua New Guinean children at the University crèche at Unitech, Lae, reproduced here as (23)-(27):

(23)

Yu/***you a bad girl***

'You're a bad girl.'

(24)

Who sa/husa(t) kisim ***brown***?

'Who gets the brown [crayon]?'

(25)

Mi ***place the ball in the left hand.***

'I place the ball in the left hand.'

(26)

Mi no kisim dis wan ia. Yu/***you draw***/dro pato?

'I'm not taking this one. Are you drawing a duck?'

(27)

Gi mi wanpela ia.

'Give me one.'

(Romaine 1992: 322)

As Romaine notes, selecting between Tok Pisin and English orthography in these cases is problematic, as (23), (24), and (26) feature homophonous diamorphs: selecting a single

standard orthographic form to represent [ju] (Tok Pisin *yu* or English *you*) in (23), [husa] (Tok Pisin *husat* 'who' or English *who* + Tok Pisin habitual marker *sa* 'who always') in (24), or [gi mi] (Tok Pisin *givim* 'give to' or English *give me* 'give to me') in (27) would obscure the very question raised by these constructions.

To understand (23)-(27) we need to address questions about the nature of English and Tok Pisin in Papua New Guinea, and to consider the possibility that the 'intermediate varieties' Romaine is describing are not stable forms but rather a result of particularly intimate bilingual behaviour. In Chapter 5 I discuss the convergence of PNG English and Tok Pisin that allows these constructions to appear, and the data presented in Chapters 6-8 reveal that intermediate forms like this may be indicative of a particular type of CS.

Additionally, Auer gives as examples of FLs ('Fused Lects') data from German-English CS in America where the discourse markers and connectives *see*, *now*, and *because* from English are the default:

> **(28)** (Sabine Behling, Hamburg, unpublished interview data) : ***Am English*** – German
>
> B: ***see***, wir sind ausnandergegang, ***see***.
>
> *'you see, we went to different places, you see'*

(Auer 1999: 325)

He also discusses Salmon's (1990) data on *Texas German*:

> 3rd-6th generation speakers almost exclusively use American English discourse markers and conjunctions (such as see, now, because ...). At the same time, all German particles, discourse markers, and conjunctions have been lost, that is, one system of discourse and text organization has been replaced by another.

(Auer 1999: 325)

For other communities, Oesch-Serra (1998) and Maschler (1998) have characterised as 'mixed codes' the varieties featuring regularised use of L2 discourse markers and conjunctions in L1 production. Clearly, the presence of English conjunctions in Mühlhäusler's example of 'post-pidgin' Tok Pisin at (22) is very interesting in the light of this

research, and attention must be paid also to the role of discourse markers in language varieties in PNG.

3.5 Summary and Questions

Throughout the literature on Pidgin and Creole languages run the themes of language contact and language change (s3.1.2). In work on language contact between other languages, CS theory has identified a number of types of behaviour, from overt 'code-switching' to 'mixed languages' (s3.3). These can be seen in examples drawn from various datasets involving various languages, including Tok Pisin (s3.4). In considering Tok Pisin today, an in-depth consideration of the relationship between code-switching and the development of different varieties of the language seems essential.

The existing research evidence of code-switching as a characteristic of the way Tok Pisin is spoken is fragmented, broadly divided between interaction with English in 'Urban' or post-pidgin varieties and interaction with indigenous language in rural communities. To date, no attempt to consider this material in the light of contemporary theories of CS has been made. Furthermore, Mühlhäusler's (1985d: 148) claim that observations of CS behaviour indicate the separation of Tok Pisin and English needs to be evaluated in the light of current accounts of the development of mixed language, in particular Auer's LM and FL, and Muysken's Congruent Lexicalization.

Labov has observed:

> [T]o demonstrate that we have a true case of code-switching, it is necessary to show that the speaker moves from one consistent set of co-occurring rules to another.
> (Labov 1970: 34) (also at s3.2.2, above)

Yet it is in the meaningfulness of the switch that Auer identifies the existence of separate codes (Auer 1998: 13) (see s3.2.4.2). A similar position underlies Mühlhäusler's observation that the identification of CS establishes that a continuum does not yet exist (Mühlhäusler 1985e: 148). However, if CS becomes the un-marked choice (Myers-Scotton 1993b), and the

code-switched variety itself, rather than individual switches, meaningful as a marker of identity, this could erode the distinctiveness of the codes. Both Auer's LM and FLs and Muysken's Congruent Lexicalization are CS without local meaning: they are un-marked CS.

Research on code-switching (Myers-Scotton 1993b) and creole continua (Rickford 1987) has noted a range of variable behaviour exhibited by speakers. Every sort of behaviour possible for individual speakers is of significance when considering the state of Tok Pisin today. This is especially important in the light of Mühlhäusler's position on CS: that such behaviour indicates that Bickerton's (1975a) hypothesis remains unproven.

To observe the full range of bilingual behaviour data needs to be collected that is as naturalistic and varied as possible. CS research relies predominantly on spoken language data, and the development of a more comprehensive appreciation of bilingual speech has been largely the result of observations of established language-contact communities in urban settings (Eastman 1992). In exploring the mesolectal variation in creole continua communities the observation of natural language has also been key (Rickford 1987; Patrick 1997). In contrast, the most recent work on Tok Pisin has relied either on textual sources (Verhaar 1995), on a corpus of stories told by schoolchildren in an interview-type situation (Smith, G. 2002), on adult interviews (Wakizaka 2008), internet postings (Harvey 2007), or local observations supported by song lyrics (Devette-Chee 2011).

In investigating the current situation in Papua New Guinea, therefore, the approach suggested by the literature was to make a detailed and multilayered analysis of an original corpus of spoken language containing examples of anglicised 'Urban' Tok Pisin.

Progressing from an initial interest in language change, the project focused on two broad themes.

First, could urban varieties of Tok Pisin (as spoken in towns, by politicians, and on the radio) be explained in terms of CS with Tok Pisin and English? Second, could decreolization proceed by a process of language mixing?

These key questions are expanded upon and explored in the following chapters.

In Chapter 4 we focus on three key research questions: whether Tok Pisin and English in Papua New Guinea could be characterised as separate systems; whether the borrowing from English described as 'anglicising' Tok Pisin could be characterised as CS under existing models, or if it extended beyond content morphemes to indicate structural borrowing and morphosyntactic change; and whether different types of Tok Pisin/English CS were occurring, especially those which could indicate the existence of intermediate grammars? The research design, fieldwork, technical methodology, the dataset of the core corpus, and procedures of analysis are described.

Chapter 5 presents detailed findings on the varieties of English and Tok Pisin evidenced in the data collected and analysed in the core corpus. This provides a background to the structural analyses of 'anglicised' Tok Pisin in the following chapters, where I apply a tri-partite account of various code-switching behaviours: Alternation, Insertion and integration, and the indicators of Congruent Lexicalization (s3.3) to 'anglicised' Tok Pisin.

Chapters 6, 7, and 8 give detailed descriptions of the different types of code-switching. The functions of language Alternation (Chapter 6) are explored as per the schema of Blom and Gumperz (1972), Gumperz (1982: 59-99), and Auer (1995) and the motivations for language choice in particular interactions in terms of diglossia (Fishman 1969; Ferguson 1972 (1959); Timm 1981; Britto 1986; Myers-Scotton 1986; Heath 1989; Hudson, A. 1994; Alfonzetti 1998; Wei 1998) and the Markedness Model (Myers-Scotton 1986, 1993b, 1998b, 2002, 2006 [1989]). Insertion (Chapter 7) is investigated using the Matrix Language Frame model of code-switching between languages proposed by Myers-Scotton (1993a), incorporating the

modifications of the 4M model (Jake and Myers-Scotton 1997; Myers-Scotton and Jake 2000; Myers-Scotton 2002; Myers-Scotton and Jake 2009), and the potential development of a mixed grammar (Chapter 8) is considered primarily in terms of Muysken's (2000: 122-153) account of Congruent Lexicalization.

Specific observations of different characteristic types of CS behaviour are presented alongside an account of their significance for the Papua New Guinean language scene and an understanding of contemporary Tok Pisin in particular.

These chapters address the role these types of bilingual language behaviour play in processes of language change in a mixed language community (Myers-Scotton 1996): a community of multilinguals who speak both Tok Pisin and English, among other languages. I do not address Myers-Scotton's claims regarding the role of the MLF in the formation of pidgins and creoles (Myers-Scotton 1993a, 1997b, 2002), because my focus is on the on-going development of language, rather than its origins, which have been comprehensively discussed and debated elsewhere (Churchill 1911; Hall 1962; Walsh 1978; Mühlhäusler 1979a; Mühlhäusler, Bennett and Tryon 1979; Sankoff, G. 1979; Mühlhäusler 1985a; Keesing 1988; Thomason and Kaufman 1988; Crowley 1990; Holm 1990; Keesing 1991; Mühlhäusler 1996; Nekitel 1998; Siegel 1999; Mühlhäusler, Dutton and Romaine 2003; Siegel 2008). The potential validity of an approach that considers code-switching as a significant feature of a language continuum has been suggested by Auer (1998b).

Chapter 9 provides qualitative and quantitative feedback from speakers on the different types of mixed language.

The final chapter discusses the implications of the findings for work on code-switching, the processes of development of creole continua, and the current understanding of variation in Tok Pisin and English in Papua New Guinea.

4 Project Design

Mühlhäusler (1985d) argued that what he saw as the distinct social roles of Tok Pisin and English in the 1970s and 1980s ensured that they would remain distinct. He opposed the position held by Bickerton (1975a) that the two languages would become increasingly part of a single language continuum, with bush pidgin at one end, and Papua New Guinean English at the other.

This essential opposition is between an account of variation in Tok Pisin as a case of language mixing or of an internal continuum.

I designed my research around this opposition. Initially intending to investigate the presence of a continuum in Tok Pisin, as I progressed I developed the idea that CS was a better explanation of the data than a continuum model.

This primary data was taken from over 115 hours of my field recordings of interviews, radio broadcasts, political speeches, and natural conversations, from which a set twenty-two texts of everyday speech were selected for multi-level coding and analysis. The c14,500 words of these texts, my 'core corpus' below, provide the examples in this dissertation. A smaller set of twenty-six samples were selected from this core corpus for the exercise discussed in Chapter 9.

Previous work on Tok Pisin has established that a range of processes are active in the development of the language (s3.1.4 above). While creolization has taken place in some communities, it is not clear that creole speakers have developed varieties that can be entirely distinguished from the fluent Tok Pisin of their contemporary second-language speakers. Numerically, most people are multilingual with Tok Pisin alongside Tok Ples and/or English (Paliwala 2012), yet only impressions are available to account for the extent to which these people use Tok Pisin as their primary language. There is variation throughout the country, yet

generational differences are most likely to be patterned by experience in English: the widespread programme of English-based education having greatest impact upon the post-Independence generation.

While an initial concept for this project had been to provide a sociolinguistic survey of Tok Pisin throughout the country, preliminary fieldwork made it clear that key questions around multilingual behaviour needed to be addressed before a meaningful survey could take place. With G.Smith's (2002) work providing an excellent account of children's language, it also became clear that a detailed account of adult language as spoken by the post-Independence urban generation was missing from the literature. Most importantly: an analysis of code-switching in Urban Tok Pisin was required before any investigation of linguistic variation could be made, and was required prior to any assessment of evidence put forward for decreolization or for the development of a Tok Pisin/English continuum.

The necessary next step on the way to a more thorough account of the language, both synchronically and diachronically, was to explore of the notion of a continuum in Papua New Guinea by carefully considering the possibility that the range of behaviour characterised as varieties of Tok Pisin was being patterned through the interaction of separate language systems, rather than through variation within a single system. This 'next step' became my primary objective.

Essentially: it was clear that the impact of English on Tok Pisin was both widespread and significant. Following approaches to multilingual communities elsewhere in the world, this impact could potentially be described in terms of code-switching and the development of new, distinctive, mixed varieties. Such a description could throw new light on creolists' questions about decreolization and the development of a continuum in PNG.

If the urban language first characterised as a 'sociolect' by Mühlhäusler (1975) could be described in terms of the interactions of separate systems, as a code-switched variety such as

that observed by Myers-Scotton in Zimbabwe (Myers-Scotton 1993a), then much of the 'anglicised' variation in Tok Pisin had to be re-interpreted in terms of this interaction. With Tok Pisin/English bilingualism increasingly common due to English-medium education, the reality of language mixing needed to be investigated. Until then, the question of whether the language systems would remain distinct, or would merge to form a continuum, remained entirely open.

Early on, I realised that any analysis developed during the course of this research project would need to be repeatable with other datasets and by other researchers over time. A methodological resource would be of real value as the Papua New Guinean language community continued to grow a new, National, identity in the 21st century. In addition to looking at variation in Tok Pisin in terms of code-switching (CS), I established a secondary objective of developing a technical method adequate to the task of incorporating additional recordings of speech and additional levels of analysis over time.

In the course of this research I found evidence of CS in 'anglicised' Tok Pisin and in PNG English. This raised the question of whether some types of CS may depend on mesolectal grammars, which are a feature of creole continuum models.

This chapter outlines the research questions formulated in response to the literature and the way these were focussed by fieldwork (s4.1). The research method is described (s4.2) with the specific field sites and speaker-selection (s4.2.2), recording methods (s4.2.3), and an outline of the dataset (s4.2.4), before data processing techniques are presented (s4.2.5). A short section (s4.3) describes the formatting of examples in this dissertation.

4.1 Research Questions

Starting with an MPhil project in 1999 I set out to investigate the state of Tok Pisin today. From the existing literature, three key questions presented themselves as central.

The first question was whether Tok Pisin and English in Papua New Guinea could be characterised as separate systems or whether they were part of a continuum. An analysis of code-switching behaviour was vital to addressing this question. However, differences of definition and distinctions in the code-switching literature complicate this approach (s3.2.2). Code-switching was a theoretically muddled area of research that had not often been applied to creole studies. On the other hand, the decreolization and continuum models prevalent in creolistics seemed to miss out on some vital facts about the separate identities of Tok Pisin and English in modern Papua New Guinea.

I decided to proceed from the hypothesis established by Mühlhäusler (1985d: 148) and maintained by Siegel (1997) and G.Smith (2000) that Tok Pisin and English constituted separate systems. I was able to refer to detailed descriptive work of both written (Verhaar 1995) and spoken (Smith, G. 2002) Tok Pisin, and studies of PNG English (Smith, A.-M. 1978; Smithies and Holzknecht 1981; Davani 1984; Smith, A.-M. 1986; Yarupawa 1986; Smith, A.-M. 1988a, 1988b; Holzknecht 1989). These previous resources allowed me to focus on the specific issues of language contact, code-switching, and decreolization revealed by my data.

However, the review of the literature (Chapter 3) suggested that a CS variety could itself become a 'code' of sorts, through Congruent Lexicalization or as a Fused-Lect. Could this be Mühlhäusler's (1985d: 79) third system of behaviour between English and Tok Pisin (s3.1.3)? Proceeding from the code-switching perspective, I might be able to provide some fresh insight on the development of creole continua.

A second question related to making the distinction between borrowing and code-switching which is so debated in the literature (s3.2.3 above). With the limited lexical inventory of even expanded traditional Tok Pisin, the acquisition of new vocabulary by its speakers to fill gaps, create greater distinctions, or describe new cultural objects and concepts would seem

inevitable, and borrowing from English is clearly indicated by Romaine and G.Smith's work to be the major source of new vocabulary (Romaine 1992; Smith, G. 2002). With English vocabulary being specifically identified by Mühlhäusler as a defining characteristic of 'Urban' Tok Pisin (Mühlhäusler 1985d: 140), could this process of borrowing be understood as code-switching, or was borrowing between Tok Pisin and English taking place beyond the lexical level, to include morphology and syntactic structures?

A third question tied the first two together: was code-switching behaviour in Tok Pisin limited to any particular type (s3.2.5 above) of behaviour? Would I find expressive, Alternational switching to be the norm, or would there be a matrix language with inserted material? Would there be evidence of multiple types of "code-mixing", as per Muysken's typology (Muysken 2000)? Would the switches have local meaning and be "codeswitching" in Auer's sense (Auer 1999), or would it be better characterised as language mixing (LM)? Bearing in mind the notion of the 'mixed code' underlying Muysken's observations of Congruent Lexicalization, would evidence of this kind of behaviour be found in Tok Pisin, and could this type of code-switching indicate the existence of a mesolect? Would there be any indications of Auer's fused lects (FL)? How could the 'Urban' Tok Pisin be characterised in terms of continua of CS?

Investigating whether variation in Tok Pisin could be characterised in terms of bilingual behaviour, and the question of whether such behaviours could actually create a (pidgin-creole/English) continuum in Papua New Guinea, these three empirical questions lay at the heart of the research project. Are Tok Pisin and English distinct languages? What is the role of code-switching and borrowing in the 'anglicisation' of Tok Pisin? Vitally, are different types of code-switching behaviour creating the preconditions for intermediate varieties, for a 'creole continuum' in PNG?

4.2 Research Method

The research aim was to collect examples of naturally occurring spoken Tok Pisin in urban English-contact contexts and to give an account of the range of code-switching behaviour in evidence. That task was primarily descriptive, but it has significant implications for the interpretation and understanding of varieties of Tok Pisin previously described in other studies.

In particular, by proceeding from an assumption that Tok Pisin and English do not form a continuum in PNG, I considered whether a CS account of the presence and interaction of two separate systems can explain the range of performance, and what the implications of CS behaviour may be. By investigating Tok Pisin material in this way, I developed a method for understanding and interpreting 'unstable' varieties that throws light on the processes of language change.

The technical approach to the data discussed in this dissertation can be extended to other datasets to build up a flexible, multi-layered record of Tok Pisin varieties. Working with my data in this way from 2000 onwards, my work was innovative in documenting the primary data and associated secondary material (time-indexed transcriptions (s4.2.5.1) and tiered database (s4.2.5.2)) with their metadata as a consistent feature of the research dataset.

The present study is based on a database of largely adult, spoken, bilingual material characteristic of 'mixed' ('anglicised' or 'Urban') Tok Pisin. The research project investigated how this speech data can be sorted out or understood by applying models of mixed language behaviour developed in reference to other language contact communities around the world. The aim was to collect data of a very different quality to that previously recorded: to actively seek out mixed language and CS behaviour in natural communication acts. In the findings chapters below, historical characterisations of Tok Pisin and PNG English will be discussed in

terms of my data (Chapter 5), and in particular new interpretations of some features will be provided by a CS analysis (Chapters 6-8).

The dynamic nature of language is an essential feature of this study, which takes the spontaneous acts of multilingual speech as the dataset. In particular, the presence of English elements in otherwise 'monolingual' Tok Pisin contexts is considered in terms of CS and the establishment of a 'mixed code'. Can the borrowing from English in 'anglicised' Tok Pisin as described by Mühlhäusler (1985d) be seen as creating a new, distinct, system, and is this 'mesolectal' in the sense of intermediate between Rural Tok Pisin and English? Would speakers of these two languages identify different intermediate mixtures somewhere between them (Chapter 9)?

Contact between Tok Pisin and English in the lives of multilinguals now extends to several generations in Papua New Guinea, and so Thomason & Kaufman's (1988) observations that structural borrowing at all levels can take place (the 'Borrowing Scale') must be considered a possibility.

4.2.1 Research Design and Field Trips

This research into the mixing of English and Tok Pisin began in 1999 with my MPhil project: a consideration of material posted on Wantoks Forum, an Internet chatroom.[8] The limitations of working with a written corpus from unknown speakers were quite clear. With the grant of a Northcote Graduate Scholarship, the upgrade to PhD research was approved, and the present

[8] Wantoks forum was active at the url : http://www.niugini.com/wwwboard/ from late 1996 until around 2005. Information on it can be found at: http://www.angelfire.com/zine/niuginiforums/files/forums.html

study developed to target naturalistic spoken language, with data to be collected during field work in Papua New Guinea.

An initial field trip in April-June 2000 collected language material in the national capital, Port Moresby, the major north coast town of Madang, and in the West Sepik capital Vanimo. A follow-up trip to Goroka, in the Eastern Highlands, and once again to Madang and Port Moresby in August 2000 extended the range of field sites considered, and allowed a comparison of the behaviour in different regional language communities. The intention was to consider the similarities and differences between these urban settings in terms of language use in the community, and the styles and registers of Tok Pisin in use.

Alongside natural language and conversation, broadcast radio material (news, advertisements, and in particular radio caller talkback broadcasts) were recorded to compare language use with the conversational data. Material was also collected from political speeches in the National Parliament. This was intended to provide a third dataset alongside the regional conversational and broadcast language, and was expected to exhibit high levels of anglicisation in Tok Pisin as well as evidence of English use. This expectation was formed following various reports by Mühlhäusler (1975, 1985b, 1985d) and in particular Nekitel (1998), as well as many others (Brash 1975; Piau and Holzknecht 1985; Wurm 1985; Dorney 1990; Romaine 1992; Siegel 1997; Smith, G. 2000, 2002)).

While initially expecting to survey and compare different regions and focus on a detailed study of one urban community in particular, the objective of providing an in-depth sociolinguistic survey of a particular representative community changed as the research progressed.

Following preliminary analysis of all datasets from the survey trips it became clear that mixing (CS) between Tok Pisin and English was present in all the communities and all the datasets, and that understanding the nature and structure of this mixed production, and its

significance in the PNG community, was essential to any work on the development of Tok Pisin and on language change at the national level.

As a direct consequence of this finding, the decision was made to focus on language use in bilingual middle class families in Port Moresby in and around Waigani, the centre of government and the location of the University of Papua New Guinea (UPNG). This decision was founded on solid ground: against a background of rising Tok Pisin use even in the Motuan communities of the Papuan region, widespread use of both languages in the National Parliament, on the overtly Tok Pisin radio station YumiFM and the English language station NauFM, as well as in the homes and friendship groups of these people, code-switching with Tok Pisin and English was a conspicuous part of day to day life in this community.

There were good theoretical reasons for focussing on an urban community in particular: other research on code-switching and on mesolectal creole varieties has largely focussed on middle-class urban communities. In thinking about the potential development of a mesolect through such practises, a focus in such communities allowed for observations of the casual way in which languages could be mixed.

In addition, there were some practical considerations. In Port Moresby I was able to live as part of a family unit with six sons ranging, at the time of my first field trip, from my own age to a new-born baby. Spanning two decades in age, these brothers allowed a unique insight into the acquisition, use, and development of relationships between English and Tok Pisin in their community. Along with students at UPNG and members of their social community, this family and their friends also provided crucial interpretations, feedback, and additional observations that helped me develop and refine the validity of my corpus and my findings.

In addition to the convenient access to an extended network of speakers, there were additional considerations of access. Travel to different areas was occasionally prevented by local unrest. As a young adult male researcher, my contact was predominantly with children, other young

men, and older parental householders. As a consequence the majority of my speakers were men. Some excellent recordings of women were made, however, and included in the 'core corpus' (s4.2.4) for close analysis.

In addition to recording natural language behaviour, I also made recordings of a number of interviews and structured conversations between myself and some speakers. I did not, however, include this data in the core corpus for analysis.

In 2002, in the second year of my PhD project, I reviewed G.Smith's incredibly detailed account of adolescents' first-language, 'creole', Tok Pisin of the late 1980s and early 1990s (Smith, G. 2002) for the PACE Newsletter (Paliwala unpublished). The publication of G.Smith's work fundamentally refocused my project by providing a detailed corpus-based description of spoken creole Tok Pisin grammar, and establishing that regional variation in Tok Pisin was not, at that time, very high.

G.Smith's work effectively addressed two of my primary interests in understanding the development of Tok Pisin: describing 'creole' varieties and considering regional variation. However, G.Smith's corpus contained almost no evidence of code-switching behaviour (Smith, G. 2002: 27), and my analyses of CS in my bilingual adult data provided interpretations of some forms that I felt were better informed by an appreciation of the multilingual behaviour that underpinned them.

Consequently, I refined my analysis of the data to a direct consideration of CS between English and Tok Pisin and its role in processes of language change, decreolization, and the potential development of a creole continuum.

As I worked on the core dataset of recordings, I was also able to develop ancillary datasets through experiments and activities (s4.2.3). However, time and again it was the complex integration and interaction between Tok Pisin and English that stood out as socially and grammatically significant.

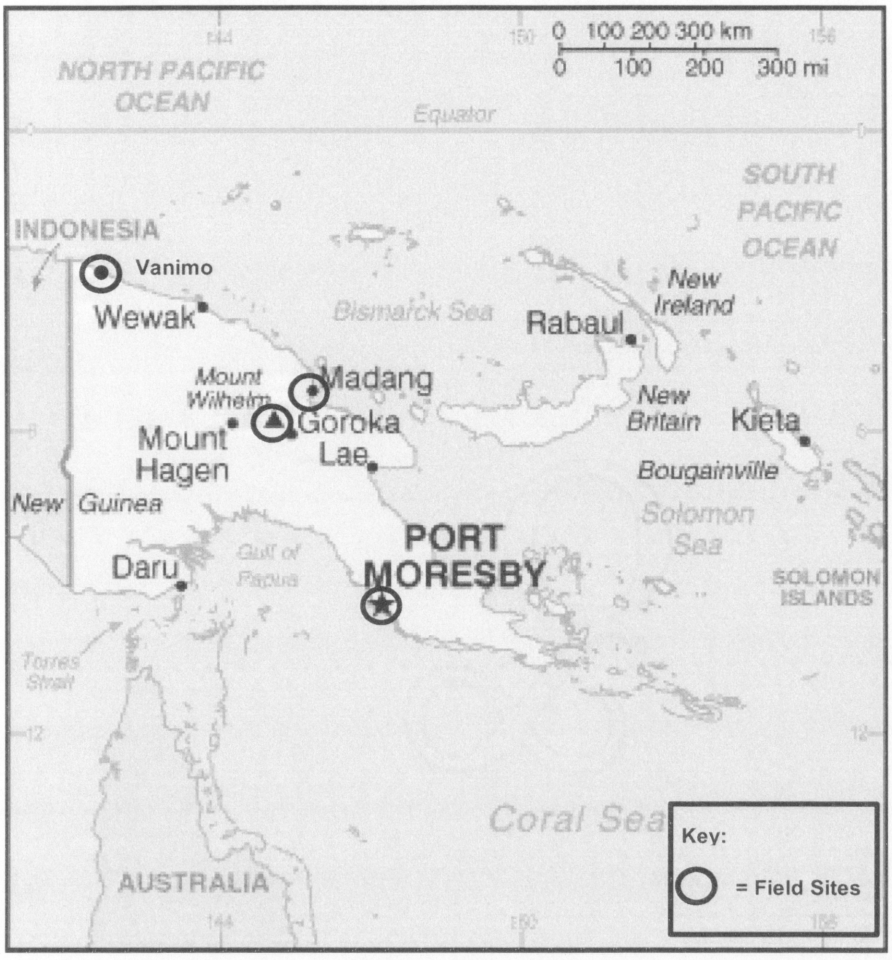
Chapter 5 discusses my findings on the grammatical similarities and differences between Tok Pisin and English in Papua New Guinea, and provides a background to the analyses of different types of code-switching present in my data in Chapters 6, 7, and 8.

In a series of workshops on mixed language in 2003 UPNG student volunteers provided me with quantitative feedback on perceptions of mixed language when they graded twenty-six samples on a scale of 1 to 10 in terms of their perception of how pure the language was, with 'English' given a score of '1' and 'Tok Pisin' a score of '10'. An analysis of the data from this 'Listening Exercise' experiment is discussed in Chapter 9.

4.2.2 Location and Selection of Speakers

Initial recordings were made of natural conversations between Papua New Guinean adults in urban settings. With a focus on mixed language behaviour, I focussed on situations and conversations that included language that appeared unlike the 'Rural' standard, or seemed otherwise influenced by English. In some cases, such as around hotels and guest houses in Madang, English was in general use between staff and international visitors, but the same staff members would use Tok Pisin in exchanges with each other. As a Tok Pisin speaker, I was able to introduce myself into these contexts and gain permission to make recordings of the exchanges. Living in Port Moresby I was able to make many recordings and observations of language use by bilingual families. In some cases, the parents would also use their respective traditional languages (*tok ples*) with each other, in particular to discuss a topic they did not want overheard by their children, who understood both Tok Pisin and English. In Lido village, outside of Vanimo, I spent time with families and young adults in groups and activities in and outside of the village. In Goroka, while at the National Sports Institute I spoke with athletes and their trainers preparing for the 2000 Paralympic Games.

Map 4.1: Papua New Guinea Field Sites (map adapted from Wikipedia) [9]



The aim of data collection was to collect sufficient material to support an analysis of aspects of variation in contemporary Tok Pisin. The field sites overlapped with three of G.Smith's informant locations (Smith, G. 2002: 30) in the MOMASE and Highlands regions: Vanimo, Madang, and Goroka. Following my focus on urban settings, I also worked in the national capital, Port Moresby, which was not one of G.Smith's field sites. While I made a very large number of recordings, once it became clear that similar behaviour was in evidence across the

---

[9] Map provided by Wikipedia under a Creative Commons license:
http://upload.wikimedia.org/wikipedia/commons/4/42/Papua_New_Guinea_map.png

country, and that the use and combination of English with Tok Pisin was something that led to the highest variation in the language I was recording, I focused on the material that illustrated a range of behaviour: from the most commonplace to the most remarkable. This material is referred to below as the 'core corpus' data.

As noted above, social norms meant that I had very limited access to female in-group interactions in the young adult age group. However, in households in Port Moresby I was able to collect some excellent recordings of well-educated mothers talking together in their homes. In the process of selecting and transcribing data (s4.2.5), I prioritised this data to provide some gender balance to the texts for analysis. In many observational situations recordings were not made due to the difficulty of obtaining permissions from all the participants. As a result, the final dataset selected for transcription and coding was necessarily restricted, while carefully chosen as exhibiting a representative range of behaviour.

Many of the younger speakers were first-language speakers of Tok Pisin, and in family groups I was able to observe the transitional phase to English-language acquisition alongside schooling, with parents also shifting their address from Tok Pisin to English as their children aged. In terms of G.Smith's "operational definition of 'first language speaker'" all of the speakers "had used Tok Pisin in a wide range of circumstances for most or all of their lives" (Smith, G. 2002: 29).

The age range of the speakers whose language composed the total dataset of recordings, and those in the selected core corpus, was between 10 and 70 years old, with the majority between 20 and 55 years of age. For Parliamentarians and people participating in Radio broadcasts, the age of the speaker would be estimated, and some rural speakers were unable to provide precise ages. My general focus was on adult language in contexts that could be considered reasonably typical. The question of generational differences in language use was addressed with my speakers, who broadly agreed that younger people could speak more rapidly than

adults, and that older rural people still retained a more conservative variety, but that the majority of Papua New Guineans' Tok Pisin was similar, regardless of age. In accordance with G.Smith's findings (Smith, G. 2000: 202; 2002) of 'anglicisation' in children's creole Tok Pisin, they suggested that it was the use of and mixture with English that made the greatest difference in people's Tok Pisin.

The majority of the 59 speakers who contributed to the conversational recordings were competent bilinguals, with conversational fluency (or better) in both English and Tok Pisin. Six were of doubtful or unknown ability in English, though their Tok Pisin was fluent. All were undergoing or had completed formal schooling in English. Eight had a university level education. Three were still in school. All of the speakers came from what could be considered 'Urban' backgrounds: though six were from villages in the Sandaun province they had electrical power in their homes and access to store commodities in the local urban centre. The remaining speakers were current or former residents of urban centres: either the capital city (Port Moresby), or the second city (Lae), or major urban centres for the Highlands and MOMASE region (Goroka or Madang, respectively). The majority of the speakers (where they were of appropriate age) were wage-earners or lived in a household with a wage-earner, though they may have supported themselves with some cottage industry. Four lived a subsistence lifestyle in village communities.

Overall, therefore, this collection of speakers was heavily skewed towards male, well-educated, urbanised, bilingual members of the cash economy in Papua New Guinea. The set of known speakers is supported by data from radio broadcasts featuring three different

presenters and around ten talkback callers, and rhetorical speaking by five different Members of Parliament and the Speaker of the House.

While broader research questions[10] would have required a more diverse sample, the data collected focuses on those sections of the community most likely to exhibit code-switching mixed language behaviour with Tok Pisin and English. They are not representative of the general population of Papua New Guinea, but of the developing middle class communities of the urban centres and larger, urbanised, villages. On a continuum model of creole language development it is among these speakers' language that examples of the acrolect and aspects of the mesolect would be expected. Urban communities of this type are also commonly among the most intense code-switchers (Eastman 1992). Previous research into Tok Pisin has clearly charted the dimensions of both 'Rural' and 'Urban' varieties, though just what the basilect is in a creolized pidgin like Tok Pisin is a matter that remains worthy of discussion.

4.2.3 Elicitation and Recording of Eata

With the objective of obtaining recordings of naturally occurring conversations between adults, suitable speakers were identified in a range of mostly social situations in the field and the purposes of the project (to investigate contemporary Tok Pisin for my university research) explained. With the speakers' consent, recording devices were activated and simply left running while the conversation developed. Where recordings were made in domestic contexts, the permission of the home-owner was sought, and the nature of the recording project explained by them to their guests. Due to the lengthy nature of the recordings taken, any

---

[10] As we suggest in our conclusions, a sociolinguistic survey of adult Tok Pisin in Papua New Guinea will benefit from the analyses presented in this dissertation, and in particular the coding scheme in Appendix III.

apparent over-compensation by speakers for the presence of either the recorder or the researcher became reduced over time. In many situations, the microphone was simply placed in a central position and the researcher left the area. In some cases I gave the recording device to trusted individuals and allowed them to make their own recordings.

These methods of participant observation (Milroy, L. 1980; Cheshire 1982) follow those used by researchers into both creole language variation (Rickford 1987) and code-switching (Myers-Scotton 1993a). Indeed, my field method is most perfectly described by Wogan's (2004) term: as 'deep hanging out'.

Not all of these recordings of naturally occurring interactions were suitable for further analysis. Often recording quality was too poor for transcription, particularly where the background noise of household appliances, such as ceiling fans, or of the outdoor environment (everything from crickets to road noise) interfered with the recording. On occasion, whole sections of conversation were lost due to technical failure of the microphone or recording device, or when batteries ran out. In general, however, this method was sufficient to obtain the required examples of conversational behaviour.

There were some cases in which I was an active participant in conversations. I consider some of this data to be of a different quality to the naturally observed interactions. Those recordings have more the structure of interviews with speakers, and are of principal value for understanding metalinguistic ideology in the community. I discuss these observations in Chapter 9.

In addition to the conversational interactions observed, this study includes behaviour in specific registers: radio broadcasts, talkback radio, and parliamentary debates. While less was known in these cases about the individual speakers, the behaviour provided the mediated cultural linguistic background to my speakers' lives.

Radio broadcast data was drawn from the news, interviews, advertisements and talkback radio on two national[11] music radio stations: the Tok Pisin language station YumiFM and the English language broadcaster NauFM. Comparing language use by the presenters and newsreaders in scripted and unscripted interaction provided valuable insights into the mediated register. Similarly informative was the recording of talkback radio: presenter/caller interactions on air. These were particularly useful due to the national scope of the coverage: callers from many different regions were identified by name and province before they began to speak, and the mixed language interactions with the presenters were often interesting.

The final set of data from which I drew my texts is recordings made of national parliament debates as broadcast on the television and observations made on a visit to the translators' facility at the National Parliament. This allowed access to high register rhetorical speeches and debates by various professional speakers. Part of the politicians' speech data was recorded from a special EmTV program. With the permission of the Librarian of the National Parliament, access to a translator's booth at the House of Assembly was granted, and 210 minutes of a Grievance debate recorded for analysis. This provided some remarkable records of political speaking in the trilingual National Parliament.[12]

Two sets of recordings were made of sessions specifically designed to elicit different syntactic and phonological forms in Tok Pisin. These comprised Tense, Mood, Aspect elicitation sentences taken from Östren Dahl's study (Dahl 1985) and a combination of word

---

[11] The radio broadcasts of the two national radio stations could not be received in one of the field sites (Vanimo).
[12] All three recognised national languages are officially allowed on the floor of the House: Tok Pisin, Hiri Motu, and English, though regionalist politics can produce vocal criticism of the use of Hiri Motu by MPs from outside of the southern, Papuan, region (see p.252, below).

lists and picture cards that were used to elicit short stories or sentences designed to include specific vowel sounds. Dahl's questionnaire was introduced to collect data on potential variation in the TMA system of contemporary Tok Pisin. This aspect of the grammar was chosen for special attention because work on Caribbean creole continua has previously focused on variation in TMA marking as evidence of multiple syntactic systems structuring a creole continuum. In additional elicitation exercises, picture cards and word lists were used to investigate phonology, lexis, and sentence structure. No data from these sessions were explicitly included in the core corpus, but the recordings were used to address questions about variation in Tok Pisin, which are broadly discussed in Chapter 5.

4.2.4 The Dataset

Over successive field trips over 7,000 minutes of audio was recorded using a combination of tape, minidisc, and hard drive storage media. Contributions ranged from naturally occurring conversations between 32 speakers in Port Moresby, 12 in Madang, and 15 in West Sepik Provinces; 180 minutes of Tense Mood Aspect elicitation; 900 minutes of story-telling and word-list elicitation; 2,970 minutes of Radio broadcasts, primarily from YUMIfm, with 90 minutes from the English-language NauFM station; 300 minutes of political debate from the National Parliament House of Assembly; and around 1,200 minutes of interviews and formal conversations between the researcher and various officials and senior academics. An additional 110 minutes of interviews with 6 men were recorded for me in Iatmul villages in

the Middle Sepik by Adam Philipson in 2004. Table 4.1 gives details of the original dataset, and indicates the size and distribution of texts that form the core corpus.[13]

Table 4.1 : Primary Data and the 'core corpus'

Data type:	**Conversation**	**Radio Broadcasts YUMIfm**	**Political Debate**	*TMA*	*Story-telling and Word Lists*	*Radio NauFM*	*Interviews*	*Iatmul*	**Totals**
Length (minutes)	1330	2890	300	*180*	*900*	*90*	*1200*	*110*	**7000**
Speakers	59	13+	8+	*3*	*5*	*7+*	*8*	*6*	**-**
'Core Corpus' (TiT) Texts	7 texts	4 texts	11 texts	-	-	-	-	-	**22**
TiT Text Length (minutes)	20	7:30	66	-	-	-	-	-	**94**
TiT Text Words (approx.)	3323	1629	9636	-	-	-	-	-	**14588**

From these hours of recordings, 59 separate sections were transcribed from the recordings with the best quality sound. From these 22 texts were selected to form the core corpus on the basis of representative behaviour and subjected to the most detailed coding.

The balanced dataset of the 'Tok Pisin/English core corpus' was based around seven conversational texts. These were in large part narratives or monologues in conversation by two bilingual speakers, one male, one female, aged around forty years old. The audience in each case was also bilingual, though in the case of the male narrative the level of competence of the audience was unknown to the speaker.

---

[13] See also Appendix I for a full listing of audio recordings, and Appendix II for a list of 'core corpus' texts.

Providing comparison and support for these were four texts transcribed from radio broadcasts. The first of these is a radio advertisement, while the other three texts were taken from talkback radio with callers discussing topics of interest with the presenter.

In addition, eleven texts taken from speeches made before the National Parliament were included to provide a substantial sample of widely-recognised 'anglicised' speech. These were taken from a Grievance debate session, and also from Prime Minister's question time during an extended sitting of the House in 2000.

In comparison with G.Smith's corpus of 400,000 words, or substantial corpora of world languages such as English, which often contain hundreds of millions of words, the (mixed language) texts assembled here represent a very small set. However, the c14,500 words spoken in the c94 minutes of these twenty-two texts attest to a range of combinations of Tok Pisin and English that characterise behaviour in Papua New Guinean 'Urban' speech, and, as I discuss in the findings chapters below, illustrate a remarkable range of language mixing phenomena. Vitally, the range of language mixing in these texts is not restricted to any particular category: private conversations and broadcast, political speeches and personal monologues, all show the same variety of types of behaviour.

The diversity of textual types and the typical nature of the urban language recorded in these selected texts provide a unique dataset: adult bilingual Tok Pisin/English conversations, speeches, and radio broadcasts representing an archive of mixed language in Papua New Guinea. Drawn as they are from a more substantial corpus of texts, and being generally unremarkable as examples of Papua New Guinean language in the communities and situations from which they were drawn, they allow us to build a detailed understanding of the way in which English and Tok Pisin relate to each other in the way they are spoken both by people every day, and by those, in politics for example, striving to demonstrate their eloquence and rhetorical flourish in both the pidgin/creole and its lexifier.

4.2.4.1 Significance & Limitations of the Dataset

This dataset is a significant record of speech in PNG at the start of the 21st Century. Digital archives of this primary data, alongside the transcripts and coded database produced from them (s4.2.5), provide a unique resource for on-going work on language in Papua New Guinea. However, the dataset as collected also has some limitations.

First among these may be considered the relatively unstructured makeup of the speakers, and the diversity of the field sites. As observed above: the data collection deliberately targeted educated, bilingual, urban speakers of Tok Pisin and English. In addition, it was notable that the gender of my speakers was strongly biased towards men: they represented all of the political data, most of the broadcast data, and all but 6 of the conversational speakers.

However, the dataset collected here, and particularly the texts analysed as the core corpus, is of a very different quality to that in earlier studies: it is adult, bilingual, conversational, and urban, and it allows us to answer questions about the nature of 'anglicised', 'post-pidgin', 'Urban' Tok Pisin and whether it is, in fact, mesolectal between Rural Tok Pisin and PNG English, or a product of distinct, identifiable, interactions between separate systems. From the work of Hall (1943) to Mühlhäusler (1985h), Romaine (1992), and more recently G.Smith (2002) suggestions about the role of English and the impact of code-switching on the development of Tok Pisin have been made. Specifically collecting and analysing texts that demonstrated this behaviour was not a part of any previous project. This data represents a new resource on the development of the language.

In the process of collecting these texts, the recordings also include a substantial sample of PNG English, and the findings discussed in Chapter 5 develop the work on written English by A.-M.Smith (1986, 1988a, 1988b) to show the existence of a local variety of Papua New Guinean English which has a particular resemblance in some respects, notably phonology, to Tok Pisin.

A second limitation of the data collected may be conceived in the presence of an observer effecting the language of speakers. This is Labov's Observer Paradox (Labov 1972). While this effect of shifting language to a more formal register, may equally take place in the presence of a recording device as well as an outsider, by following a method of participant observation (Milroy, L. 1980; Cheshire 1982) in the communities where my speakers lived I endeavoured to minimise this effect. In addition, by comparing the natural language data with that recorded from broadcast media and from agitated contexts in political debates my findings could be balanced to further reduce any impact of the 'observer'. Furthermore, the analysis resulted in findings that demonstrate the range of behaviour exhibited in my data. This was made possible by focusing on the actual occurrence of code-switching behaviours, rather than on particular features. The consequent dataset presents a record of speech as it occurs throughout the language community as a whole in Papua New Guinea.

The corpus assembled over the course of this project is valuable primarily as a rich and detailed record of communication acts in Urban Tok Pisin in the early years of the 21st century. It has an additional value as an example of the product of a particular kind of methodological approach to the codification and analysis of such data, a method which retains a flexibility and amenability to different research questions while also presenting a rich and reusable corpus of language.

It is my intention that both corpus and method will be made available through the LIPPS project for other researchers on code-switching to use, and as a potential resource for further research into the development of Tok Pisin.[14]

4.2.5 Data Processing Techniques[15]

The primary data collected was in the form of audio recordings of natural language. The earliest recordings were made on D90 cassette tape. This continued to be the medium on which broadcast radio was recorded throughout the project because portable radio/cassette recorders were widely available at the time. Cassette audio was digitised very simply by using Windows SoundRecorder software with a 4.5mm male-to-male cable connecting the headphone jack of a cassette player to the microphone port on a computer sound card. On average, the .wav format files recorded in this way took around 190kb/second of audio. I used a program called CD Wave to edit the long .wav files into separate tracks, which became the texts for transcription. This allowed me to remove the sometimes lengthy stretches of background noise and inaudible audio and identify the useful sections of language data. A selection of audio tapes was also digitised at the University of Sydney by Corin Bone in 2002, and forms one of the earliest deposits for the PARADISEC digital archive.

---

[14] The Language Interaction in Plurilingual and Plurilectal Speakers (LIPPS) Project implements the Language Interaction Data Exchange System (LIDES) (a specialised modification of the CHild Interaction Language Data Exchange System (CHILDES)). It uses the CLAN format to hold transcripts and codes, and to make them available for analysis. See www.ling.lancs.ac.uk/staff/mark/lipps/lipps.htm . See Appendix III for details of the coding scheme used in the core corpus.

[15] A list of the software programs used to process the data, along with links to their download url locations, are given in Appendix V.

As field researchers will know, cassette tape is a cumbersome medium to transport and use. For recording language for linguistic analysis, tape recordings have the additional occasional disadvantage of including the background noise of the cassette reels being wound: a repetitive clicking which cannot later be deleted. One way to solve this problem is to ensure the microphone is kept at a distance from the actual recorder. My field kit for recordings was a SONY radio/Walkman recorder, which I transported in a leather bum-bag with the bidirectional microphone clipped to the outside. This secured the unit from heat and moisture, and kept it close to hand. It could also be given to my speakers to wear when I was absent, or simply placed in a central location without drawing excessive attention. However, the downside of such practicality, convenience, and unobtrusive recording was occasionally poor sound quality and the fact that each cassette had to be carefully reviewed once digitised, and separated into useful tracks.

I was convinced of the value of using the latest technology in the field, and later adopted the SONY MiniDisc system for field recordings. MiniDisc media are much smaller and lighter, and the digital recorder I was using automatically created separate tracks. Creating working copies of MiniDisc recordings was no more convenient that audio cassette however, as SONY and other manufacturers included systems which prevented the rapid transfer of digital media in an effort to prevent music piracy. By 2003 I was able to make some recordings directly to the hard drive on my laptop, due to improvements in hardware storage.

It is worth noting that today, a large range of digital recording devices can hold huge amounts of high-quality data very conveniently, and transfer them to a PC rapidly and accurately.

4.2.5.1 Transcription

Once the original recorded data had been collected I began the task of transcribing the material for analysis. This was initially done when I returned from the field-trips, though on later trips I was able to use my laptop to transcribe material in the field. I used the freeware

specialised tool Transcriber (see Appendix V). This allowed me to create time-indexed transcripts of my audio, which could be visibly linked to the associated sound files for checking. This methodology has become increasingly accepted, and is supported and implemented by digital archivists as facilitating transparency and search of primary language data.

Much research on variation in creoles and other languages has considered phonological form as significant, and in code-switching research too the issue of phonological integration is central to certain distinctions, such as Poplack's (Poplack 1980; Sankoff, D., Poplack and Vanniarajan 1990a) concept of the 'nonce' loan (see s3.2.3.2, above). Furthermore, Tok Pisin and English orthographic standards distinguish shared words, such as *haus* → *house* 'house' by language on the basis of different spellings. Consequently, a broadly phonetic approach was followed for transcription with the aim of retaining phonological distinctions that would have been obscured by using standard English and Tok Pisin orthographies. Because the transcription and data-basing software used only supporting the ASCII character set, a detailed IPA transcript was beyond the technical capabilities of this study.

This decision was very significant: a transcript that follows the phonological form of the actual language spoken retains important details regarding areas of homophony and distinction between languages. Assumptions about 'the language' of an utterance were not coded into the transcripts themselves by adherence to orthographic standards. This reflects the concerns Romaine (1992: 322) expressed about the bias implied by using a particular orthographic standard to record the "intermediate" colloquial language of her Unitech children (s3.4.3). This openness was valuable, though it did mean that additional transcriptions of the data in the orthographic standard had to be added to the corpus data later.

Details of all the conventions followed may be found in the specimen CLAN readme.txt file in Appendix III.

This corpus was translated into electronic format, transcribed, and coded by the method described below to allow statistical comparative analyses to be made, and to make possible automated searches of the texts by lexical item and language feature mixing. The purpose of this was to prove a method for investigating language mixing phenomena in large corpora.

4.2.5.2 CLAN Programs and coding

Working from the hypothesis that Tok Pisin and English remained separate systems (s3.5), and that urban language was entirely characterised by code-switching behaviour, I adopted the LIPPS Group's guidelines on the development, coding, and analysis of mixed language corpora and applied them to my transcripts (LIPPS Group 2000). The files that Transcriber produced were readily converted into the CLAN format, and I began coding the data with various layers of analysis.

The CLAN databasing format is essentially a tiered database, rather than a relational database. In CLAN each utterance is recorded as a single 'main tier', and associated with any number of other, carefully defined, 'tiers' containing alternative transcriptions, translations, and observations.[16] The CLAN software includes analytic programs that enable frequency-counts of individual words, or any detail on any tier. Particularly useful for sequential analysis, it also includes programs that allow you to search for the juxtaposition of different features, or for the appearance of multiple features within a single tier. It presents a powerful analytic toolbox for a researcher to use once their transcripts are coded.

---

[16] See Appendix III for coding scheme. Appendix IV and Appendix VII, Sample 1 contain examples of coded texts.

In general, for every minute of recorded data, an average of 35 minutes were required to digitise the recording, to transcribe it and check the transcript for errors, and to convert the file to CLAN format and add the first tier of coding. On average, it took 15 additional minutes to add each subsequent tier of coding for 1 minute of audio. The 102 minutes of audio in the core corpus (s4.2.4) accounted for around 160 hours of transcription and coding on this basis alone. With additional time being taken defining and refining the coding scheme and for the necessary revisions of the CLAN database data with reference to the transcripts and original audio data, as well as checking the validity of the transcripts with informants, each minute of data could take up to 2 hours of treatment, beyond the original process of reviewing the original audio recordings and select specific sections for analysis.[17]

The result of this investment in the technical methodology, the technical processing and the coding of the data, was a significant resource. Though keying the programs to ask the right questions was a complex process in itself, and analysing the report files was necessarily time-consuming, the process of searching this data for patterns became extremely rapid. Once a question was formulated, answers could be achieved across the dataset with ease. Furthermore, the entire system was scalable: once the coding scheme was established and basic search programs formulated, additional data could be added and reviewed, as individual texts, as part of the whole dataset, or in terms of a particular subset, simply and comprehensively. Questions asked of the core corpus data could be asked of other datasets

[17] This compares favourably with the suggested basic transcription and translation time discussed by Peter Austin in the PARADISEC blog: http://www.paradisec.org.au/blog/2010/04/how-long-is-a-piece-of-string/

coded in this way to enable comparison. This represents a valuable future use of the framework I established.

There are also some practical disadvantages to CLAN. As a text-based databasing program the user interface is essentially a raw text file, with highly specialised strings identifying different tiers and features. A coding scheme must be carefully designed and adhered to by those working with the data to ensure consistency. In addition, the tier-structure means that any feature identified by the programs is only located on its own tier: it always remains the task of the analyst to identify what features on a coding tier relate to which parts of the language transcript on the main tier. Though subsequent versions of the CLAN software have introduced significant improvements (particularly in terms of automating the cross-referencing process) the tiered structure still results in a time-consuming and occasionally opaque dataset.

In terms of providing a structure for the data, and supporting the recording of many levels of observations, CLAN remains a powerful tool. I used it to support up to six layers of analysis of the core corpus data.

During the course of this research, individual words on the main tier were tagged for language and inter-language influences, or as shared words common to both languages, as per the LIPPS guidelines. A standard orthographic transcription was provided on a secondary tier, and an English translation on a third tier. A morphological gloss tier was included. Variable morphological, syntactic, and phonological features were noted on a fifth tier, along with a code for the variant expressed and the language context of expression. This was introduced to test the dataset for a continuum-type analysis of feature-variation. However, a more substantial and varied corpus developed through collection of additional texts by sociolinguistic survey would be better suited to such a future analysis.

Types of language interaction behaviour were recorded on a sixth tier, and it was this level of analysis that proved the most useful in assembling the data for the findings chapters presented below.

4.3 Presentation of Examples

The example sentences in this dissertation drawn directly from other authors or from the core corpus data are presented in Lucidia Sans Unicode font. Identifiable English elements are italicised in bold.

Examples from my primary data are presented following a phonetic orthography, as described in Appendix III. Following each example, where required, are morphological and standard English/Tok Pisin glosses in the same Times New Roman font as the body text. Where a homophone is prominent in the example, both standard English and conventional Tok Pisin orthographic forms are given. Once again, English elements are italicised in bold font. Features of specific interest are underlined in both the phonetic orthography and the standard orthography of each example. An English translation is provided in italics below.

This presentation allows the reader to maintain an awareness of the spoken nature of the primary data: retaining the fundamental ambiguities and variation that characterise these recordings of bilingual speech it includes morpho-syntactic analysis and also retains transparency for those acquainted with the conventional orthographies.

4.4 Summary

This chapter has outlined the research questions (s4.1) and the methods of data collection and recording (s4.2) involved in collecting and analysing the primary data referred to in this dissertation.

In Chapter 5 I discuss findings regarding the grammatical structures of Tok Pisin and English as characterised in previous research and as evidenced in this new data. This provides an up-to-date context for the analyses of CS in the later chapters.

5 Findings 1: Tok Pisin and English in Papua New Guinea

This chapter presents a discussion of various aspects of the grammars of Tok Pisin and English languages. It compares the findings and conclusions of benchmark studies of these languages in Papua New Guinea with the new data, and lays the groundwork for analysis of mixed language behaviour (CS) in the speech of bilingual adults.

These languages are both widely spoken in Papua New Guinea, and substantially spoken together. While no conclusive statistics on Tok Pisin/English multilingualism are available, English-medium education and the nationwide spread of Tok Pisin among the post-Independence generation of Papua New Guineans means that these languages are constantly and regularly used by the same people, in a language-community of 'Nationals'.

In the case of Tok Pisin, several studies and descriptions have been made of the grammar over the years, focusing on particular features, specifically the role of substrate influence, and the development of the language through expansion and creolization. Hall's (1955) description of the language included British and American second-language varieties, and he specifically distinguished these from the language as produced by Melanesian speakers. Later orthographies consequently moved away from an acceptance and incorporation of American, Australian, or British pronunciations of Tok Pisin as normative. Instead we find in Mihalic (1971) a deliberate focus on Melanesian norms of pronunciation. This reflects not only a change in the political situation, as Papua New Guinea moved towards self-government (see Chapter 2), but also an acceptance of the existence of Tok Pisin not only as an interlanguage serving communication between outsiders and locals, but as a significant language for communication between local groups. As a result, in more recent studies it is rare to find Tok Pisin being described as it would be spoken by expatriates. However, this has come at a cost of obscuring some of the long-standing variation in the Tok Pisin of some National speakers, especially where it resembled English pronunciations more closely. The conjunction of this

conservative orthography with English-medium schooling may have increased the view of 'anglicised' Tok Pisin as the result of recent contact, rather than a characteristic of some varieties of the language pre-Independence, or in the pre-WWII period, for which recordings are not available.

As benchmark accounts of the Tok Pisin grammar, Verhaar (1995) represents an account of the written register of Tok Pisin, while Geoff Smith's comprehensive study of schoolchildren's first-language Tok Pisin provides a description of their spoken language (Smith, G. 2002).

I selected G.Smith's work as a point of reference for two reasons. Primarily, G.Smith's work is based on spoken language while Verhaar's study is based on a substantial corpus of written language. As my data is exclusively spoken language, G.Smith's observations provide the closest benchmark. Secondly, G.Smith's survey of schoolchildren's language is based on data collected between 1985 and 1993, with the majority of his samples collected in 1990-1992. With the age range of his speakers being between 10-19 years, this made them effectively the post-independence generation of Papua New Guinean nationals, and the contemporaries of many of the adult generation that I interacted with during my own fieldwork. Schoolchildren of 10 years of age in 1992 would, if they progressed so far in their education, be University candidates during my field trips in 2002 and 2003.

However, the nature of G.Smith's method of data collection, his speakers and the context in which their *stori* were collected, may, as G.Smith admits, have had an impact on his language data. In particular, all of his speakers were in school, by default a place where English language was available and actively being taught. G.Smith observes a variety of potential interference phenomena and conflicts arising from a tension between Tok Pisin and English for his speakers, and notes that his own presence as an English-speaking foreigner may have impacted on their performance. However, while these factors might have been expected to

increase the influence of English on their speech, his explicit request for Tok Pisin speakers and for *stori* in Tok Pisin appears in some cases to have had the effect of limiting the amount of English used by his speakers.

Unlike my study of adult language, G.Smith's work is explicitly not a study of code-switching, or of bilingual effects on Tok Pisin. While his corpus does include a small amount of code-switched material, generally single-words from English or small strings of English words, his methodology restricted any tendency his speakers may have had to alternate between Tok Pisin and English, or to insert large constituents of English into their Tok Pisin *stori*. In my new data, as we will see in Chapters 6 and 7 of this dissertation, such behaviour is both frequent and commonplace. However, while his account is explicitly of the creole variety of Tok Pisin, it cannot be taken as representing a variety uninfluenced by English, due to the context of data collection.

As has been discussed elsewhere (s2.6 and also Paliwala (2012)), the 2000 National Census data reveals that the bilingual Tok Pisin/English community in PNG currently represents at least 30% of the national population, and G.Smith's speakers are likely to be a part of this community today. We are now in a position to test whether the incipient changes he observes as being potentially the result of such sustained bilingualism have emerged.

For English, many grammars are available for standard and dialectal varieties around the world, but the work of A.-M.Smith gives a specific treatment of the Papua New Guinean variety of English (PNG English or PNGE) (Smith, A.-M. 1978, 1986, 1988a, 1988b), noting a number of phonological and syntactic differences between PNG English and an Australian English standard. As discussed above (s3.1.6) several additional accounts of English in PNG have been made (Davani 1984; Barron 1986; Yarupawa 1986; Holzknecht 1989), and Platt, Weber, and Ho (1984) situate PNG English alongside other 'New Englishes', giving a fascinating international context.

My analysis is informed by accounts of standard English and observations on the PNG dialect. As my focus is on the interaction of Tok Pisin and English, I do not provide here an intensive study of the varieties of English spoken today, though it is clear that English-speakers in Papua New Guinea range from those who speak an international standard, to those who command a dialect with a distinct phonology and a number of other non-standard morphosyntactic characteristics. Examples from previous studies provide benchmark observations of this indigenised English from which to evaluate some of the non-standard forms appearing in my data. This is particularly relevant in Chapter 8, where indications of Congruent Lexicalization in Tok Pisin and PNG English are discussed. In some cases, such as a few examples discussed by Romaine (1992: 322) (see s3.4.3 (22)-(27)), making a distinction between Tok Pisin and PNG English is a challenging task. This chapter investigates some similarities between these language systems, which have been previously described in relation to standard varieties of English.

Below I review the phonology (s5.1), morphology (s5.2), morpho-syntax (s5.3), and lexis (s5.4) of Tok Pisin with reference to G.Smith's findings on the creole variety of Tok Pisin, and with reference to my new data. Comparative observations are also made on PNG English, and on the similarities and differences between the two languages.

This chapter sets the stage by establishing that differences that existed at the time of Independence between English and Tok Pisin have lessened. There is now evidence for high levels of congruence between some varieties of Tok Pisin and PNG English. The consequences of such congruence can be seen in subsequent discussion of the on-going interaction between these two languages in bilingual production.

5.1 Phonology

Romaine has observed that "phonology remains the least stable system in otherwise stabilized pidgins" (Romaine 2000: 184), and this is certainly the case in Tok Pisin. Descriptions range

from 24 phonemes (Laycock 1985b) to 26 (Mihalic 1971) with variation explained with reference to speakers' first language as well as what has been generally termed 'anglicisation'.

In common with other New (World) Englishes, PNG English has been characterised as featuring variation from Australian and other older Englishes of the inner circle. The discussion below shows, however, that the distinctive phonological system of PNG English closely resembles that of Tok Pisin. PNG English falls outside of Platt, Weber, and Ho's (1984: 9) division of New Englishes on the basis of their 'background' languages (local/pidgin/creole) precisely because Tok Pisin exists as both a language of wider communication ('pidgin') and as a primary language ('creole') for sections of the community. With the spread of Tok Pisin throughout PNG today and with English remaining the language of education, it is clear that it is the most significant single background language for English speakers, just as English is the most significant single background language for Tok Pisin speakers (Smith, G. 2002; Paliwala 2011, 2012).

G.Smith's (2002) study notes regional variation in the pronunciation of Tok Pisin words due to the influence of local languages on the children's phonology. Laycock (1985b) suggested that because atypical pronunciations are the subject of comment and ridicule in the Kanage cartoons of *Wantok Niuspepa*, and remain a feature of jokes throughout PNG today, there is a standardising pressure that may apply on the Tok Pisin of people when they move around the country. Despite such pressures from the wider community, G.Smith's schoolchildren's language does show some regional variation.

The impact of English on the phonology of Tok Pisin is more widespread. Studies from Hall (1943) to Laycock (1970) and Mihalic (1971), followed by Mühlhäusler (1975), Laycock (1985b), Romaine (1992), Siegel (1997), and G.Smith (2002, 2004b) have all considered the role of English in effecting the way in which Tok Pisin is spoken. In particular, Laycock

discusses Tok Pisin phonology in terms of its relationship to standard English orthography (Laycock 1970: xiv-xvi), later improving on this account with a table showing the relationship between the two systems in terms of the "principal developments of English phonemes in Tok Pisin" (Laycock 1985b: 295). Sections of this table (Laycock 1985b: 296) are reproduced below, and inform the discussion of the two systems. I use Laycock's published account as a point of reference for reasons of practicality: it may be compared with A.-M.Smith's discussion of PNG English phonology (Smith, A.-M. 1978, 1986) because both take a standard English orthography as a starting point for consideration of phonological differences between different systems.

Accounts of the phonology of Tok Pisin commonly include both a 'core' inventory, and an additional set of phonemes in the 'anglicised' varieties. These may include additional diphthongs used to reflect English pronunciation on borrowed words (Smith, G. 2002: 48), as well as introduced consonantal phonemes from English.

Variation in the way that Tok Pisin is pronounced is consequently seen as depending on a number of factors. While the earliest accounts (see e.g. Hall (1943)) saw the key distinction as being between foreigner and local pronunciation, and Mihalic (1971) established a standard around indigenous Morobe, Madang, and Sepik (MOMASE) regional pronunciations, by Independence the rural/urban context was seen as the most significant axis of variation (Mühlhäusler 1975). Post-Independence studies considered the impacts of both substrate influences and urbanisation (Laycock 1985b; Romaine 1992). Lately, increased mobility and population migration has led to G.Smith (2002) claiming that it is exposure to English, rather than location, that is creating the greatest differences. As a result of bilingualism with English, many Tok Pisin speakers have potentially the entire phonemic inventory of English at their disposal (Laycock 1985b: 295), increasing variation.

It is precisely against this context of bi- and multi-lingualism with Tok Pisin and English, as well as indigenous languages, that this study takes place. The initial wide-ranging fieldwork found, like G.Smith, that the impact of English on Tok Pisin was more striking than that of the indigenous language context. Drawing upon, primarily, A.-M.Smith's work (1978, 1986), I also observed that the phonology of PNG English showed some striking resemblances to Tok Pisin, with significant potential impacts on the congruence between some varieties of the two languages.

Observations made by G.Smith and others[18] on the phonology of contemporary Tok Pisin, as opposed to the prestigious 'Rural Pidgin' standard are considered below, and these are connected with the new data and observations on the phonology of PNG English. Particular note is made of the range of variation available in the phonology of both languages, as phonological form has been considered significant in CS research.

5.1.1 Vowels

The standard 'core' set (Laycock 1985b: 297) of Tok Pisin vowels is considered to be a set of five basic phonemes /a/, /e/, /i/, /o/, and /u/, all with a short-mid length (Dutton and Thomas 1985), "fairly close to cardinal IPA values" (Smith, G. 2004b: 719), and three diphthongs /ai/, /au/, and /oi/ (Mihalic 1971), common to all varieties of the language. These are represented as Table 5.1 and Table 5.2:

---

[18] G.Smith (2004b) discusses Tok Pisin phonology on the basis of Mihalic (1971) and Laycock (1970, 1985a), and with reference to his own data. I also consider these early sources in my discussion, below.

Table 5.1: Core Tok Pisin Phonology - five vowel set

Vowel	IPA	*TP*	
/a/[19]	[a]	*hat*	'hot'
/e/	[e]	*bel*	'stomach'
/i/	[i]	*pis*	'fish'
/o/	[o]	*rot*	'road'
/u/	[u]	*pukpuk*	'crocodile'

Table 5.2 : Core Tok Pisin Phonology - three basic diphthongs

Diphthong	IPA	*TP*	
/ai/	[ai]	*taim*	'time'
/au/	[au]	*maus*	'mouth'
/oi/	[oi]	*boi*	'boy'

Hall's (1943) grammar actually allows for a much larger set of vowels than this, but it is significant that he is providing an account of Tok Pisin as a regional lingua franca. As a consequence his phonology for Tok Pisin includes British and American pronunciations of certain words, and the influence of native English-language pronunciations is apparent. Observing changes that had taken place since the 1930s, Laycock (1970) described a ten vowel set in coastal regions, and Mihalic (1971: 4) acknowledges this recent experience, recording nine of these in his dictionary, excluding the long [a:]. Post-Independence, I agree with Mihalic (1971) that it is the Melanesian pronunciations of the vast majority of speakers that are to be considered the most fundamental indicators of the phonology of the language. In his orthography, which has become an unofficial standard for written Tok Pisin (Romaine

---

[19] Examples for contrasts in /a/ are drawn from (Mühlhäusler 1985d)

1989b: 8-9), Mihalic makes no attempt to record the newer, variable, distinctions, though he does acknowledge additional phones. Consequently, the simple five-vowel set of Mihalic's orthography represents an artificial standard, for it is clear from Hall, Laycock, and Mihalic's own revisions to 'the sounds' of his orthography (Mihalic 1971: 4) that a larger set of vowel sounds continued to be available in the spoken Tok Pisin varieties of Papua New Guineans.

Laycock (1985b) allows 12 basic vowels for speakers whose dialect is not, as he says, entirely 'anglicised'. These include open-mid central [ɜ] and back [ɔ] vowels. Table 5.3 combines examples from Mihalic and Laycock of the vowel-sounds in Tok Pisin and equivalent sounds in English to indicate a maximal vowel set for Tok Pisin as indicated by their work. Note that the near-open front vowel [æ] in Standard (British) English *fat* [fæt] is not recorded by Laycock: a fully-open [a] vowel is more usual in Tok Pisin:

Table 5.3 : Tok Pisin Phonology - maximum vowel set illustrated by contrasts (ref. Laycock 1985: 302 & Mihalic 1971: 4)

vowel	IPA	*TP*		*English*	IPA	*TP*		*English*
/a/	[a]	*pat*	'fat'		[a:]	*hat*	'hard'	*arm*
/e/	[e]	*wet*	'wait'	*plate*	[ɛ]	*let*	'belt'	*bed*
/i/	[i]	*nil*	'nail'	*machine*	[ɪ]	*pik*	'pig'	*hit*
/o/	[o]	*kol*	'cold'	*old*	[ɒ]	*dok*	'dog'	*hot*
	[ɔ]	*kot*	'court'	*court*	[ɜ]	*wok*	'work'	*work*
/u/	[u]	*susu*	'breast'	*tulip*	[ʊ]	*puspus*	'copulate'	*put*

In contrast, Verhaar categorically describes the 'core' vowel set as reflecting the absence of distinctions between long/short, high/low vowels (Verhaar 1995: 9). This characteristic

produces the Tok Pisin *ken* [ken] from English *can* [kæn], *kantri* [kantri][20] from English *country* [kəntɹi:][21], *gat* [gat] from *got* [gɒt], and *kot* [kot] from *court* [kɔt]. Verhaar's observation that diphthongs in borrowed words are reduced to simple vowels is reflected in Tok Pisin *nem* [nem] from English *name* [neɪm] and *ston* [ston] from *stone* [stəʊn].

These transformations can be seen in the table Laycock presents of vowels and vocalic nuclei showing the differences and similarities between English source vowels and Tok Pisin pronunciations of equivalent words. These include eight diphthongs /ia/, /ea/, /ai/, /oa/, /ua/, /ei/, /au/, and /oi/. This is reproduced here as Table 5.4[22].

---

[20] Note that I am following standard British, American, and Australian dictionary usage here and indicating the alveolar approximant with the IPA character [r] rather than the formally correct [ɹ]. The use of [r] here should not be taken to indicate a trilled consonant, though trilled pronunciations are possible in PNG English.

[21] For English pronunciations in the body text I use standard British English as a reference, except where otherwise noted.

[22] Laycock's table includes a footnote indicating that he uses some forms to "stay closer to the orthographic representations, except where these ae wildly at variance with the phonetic/phonemic facts" (Laycock 1985b). At no point does he make clear which variety of English he is referring to. There are some indications that Laycock's table follows the older Macquarie Dictionary representations developed from Mitchell (1946), for example in the characterisation of the vowel in English *beads* as [i], as does A-M.Smith's comparison between PNG English and a 'General Australian' standard (Smith, A.-M. 1986), and these are referred to here. Footnotes observe differences between characters presented in Laycock's table and expected IPA forms in British and Australian standard varieties.

Table 5.4 : Tok Pisin Phonology - vowels and vocalic nuclei (ref. Laycock 1985: table 1)

VOWELS AND VOCALIC NUCLEI							
IPA		***English***	***Tok Pisin***	**IPA**		***English***	***Tok Pisin***
i[23]	i	*beads*	*bis*	ɪ	i	*fish*	*pis*
ɪə	ia	*beer*	*bia*	ɛ	e	*bed*	*bet*
ɛə[24]	ea	*prayer*	*prea*	ei[25]	ei	*day*	*dei*
	e	*where*	*we*		e	*wait*	*wet*
æ	a	*fat*	*pat*	æu[26]	au	*house*	*haus*
	e	*can*	*ken*				
a[27]	a	*must*	*mas*	a:[28]	a	*heart*	*hat*
ai[29]	ai	*fight*	*pait*	ɒ	o	*dog*	*dok*
					a	*hot*	*hat*
ə	e	*potato*	*peteita*	ə:[30]	o	*dirty*	*doti*
ɔ:	o	*law*	*lo*	ɔɪ	oi	*boy*	*boi*
	oa	*more*	*moa*				
	ua	*store*	*stua*				
ou[31]	o	*boat*	*bot*	ʊ	u	*cook*	*kuk*
u[32]	u	*room*	*rum*				

[23] British English would have [iː]

[24] British English would have [eə]

[25] British and Australian English would have [eɪ]

[26] British and Australian English would have [aʊ]

[27] British and Australian English would have [ʌ]

[28] British English [ɑː], Australian English would have [a]

[29] British English would have [aɪ] or [ʌɪ], Australian English would have [aɪ]

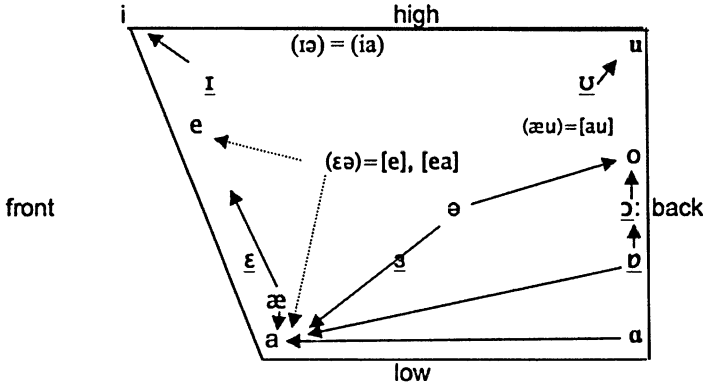
[30] British English would have [ɜː] or [əː], Australian English would have [ɜ]

[31] British English would have [əʊ], Australian English would have [oʊ]

[32] British English would have [uː] or [ʊ], Australian English would have [u]

The differences Laycock records between English and Tok Pisin (Table 5.4) can be reproduced as shifts in a vowel space, as in Figure 5.1[33]:

Figure 5.1 : Tok Pisin Vowel Space (adapted from Laycock 1985: 296 (table 1))



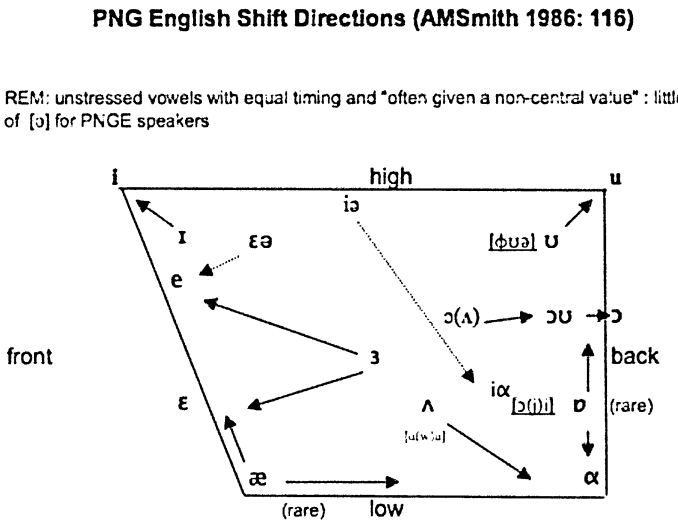
In New/World Englishes in Africa, India, Singapore, and in accounts of PNG English similar vowel transformations and a reduction in distinctions have been observed (Platt, Weber and Ho 1984: 31-37; Smith, A.-M. 1986: pp116; Holzknecht 1989: 184). The fronting of central vowel [ɜ] to [ɛ] and of [ɑ] to 'front a' [a], the movement of [æ] to [a] or [e] and of [ɔ] to [ɒ] or [o], and the avoidance of a central mid vowel [ə] have also been observed. In addition, these indigenized Englishes, like PNG English, tend to reduce diphthongs to simple vowels, and have variable approaches to vowel length. Platt *et al* note "the English of educated Papuans and New Guineans" alongside Indian and Philippine English as making occasional

[33] Note that the back [ɑ], as in British English *heart*, is marked as shifting to Tok Pisin front [a] though an Australian English pronunciation of the same word would not change in Tok Pisin.

distinctions between some words on the basis of vowel length (Platt, Weber and Ho 1984: 32).

A.-M.Smith (1986) provides us with a diagram of the vowel space of PNG English, showing transformations from a standard, Australian English[34], pronunciation, reproduced here as Figure 5.2.

Figure 5.2 : PNG English Vowel Space (from A.-M.Smith 1986: 116)



It is not entirely clear from A.-M.Smith's discussion whether the vowel space indicated is proposed to have a standard shape. For example, the open mid-back unrounded vowel [ʌ] is positioned centrally on her chart. In addition, it is possible that the vowel she indicates in low-back position with the non-standard symbol 'ɑ' represents a low /a/, without a distinction between [a] and [ɑ], in her orthography. This interpretation is supported by the indicated shift in the [iə] dipthong to 'iɑ' on her chart.

---

[34] A-M.Smith describes her point of comparison with PNG English as a 'General Australian' standard (Smith, A.-M. 1986).

There are some problems with the non-standard orthographies used in previous studies. However, comparing the observations of Laycock (Table 5.4 and Figure 5.1) with the literature on PNG English and A.-M.Smith's figure (Figure 5.2) it is clear that Tok Pisin and PNG English show similar tendencies. Combining their observations with contemporary experience of the languages, it is clear that these include: the avoidance of central vowels such as [ɜ] and [ə]; the movement of near-open front [æ] to a fully-open [a] or [ɑ] (ɑ); and the movement of [ʊ] to [u] at top/back. In both Tok Pisin and PNG English the English dipthong [ɪə] becomes [ɪa].

While similarities between various New/World Englishes may lead to general speculation about patterns of articulatory adaptation, A.-M.Smith considers these transformations in the PNG dialect of English to be a possible consequence of the impact of L1 phonology on the pronunciation of the students' English she studied. This is similar to the observations about regional variation in Tok Pisin phonology by G.Smith (2002) and by Laycock (1985b) discussed above. Importantly, however, and as observed above (s5.1), Tok Pisin is the most significant single 'background' language for PNG English. Clearly, because both these languages are spoken by speakers from the same country, similar influences, and similar consequences, may be expected but so too may interactions between them.

What is significant for this study is that the pronunciation of English vowels by PNG English speakers may often resemble equivalent Tok Pisin pronunciations, increasing the similarity in phonological form for many words common in both languages. As a result, where a word is borrowed into Tok Pisin from PNG English, phonological integration may not be an indicator of it being established in Tok Pisin.

This finding has direct bearing on the question about the distinction between borrowing and code-switching, and has an impact on the understanding of anglicised vocabulary observed in Tok Pisin varieties. For example, while G.Smith maintains that syntactic and morphological

integration are his measures for considering whether a word was a Tok Pisin word (Smith, G. 2002: 44), he elsewhere identifies some words in his corpus as being nonce loans from English on the basis of their phonological form (Smith, G. 2002: 72), while in other cases phonological integration is considered alongside morphological and syntactic integration as a marker of Tok Pisin vocabulary borrowed from English (Smith, G. 2002: 98-99).

Integration has been used as a key index to distinguish borrowing from code-switching (s3.2.3.1), and phonological integration is considered more significant than morphological or syntactic integration under the MLF model (Myers-Scotton 1993a) as well as in Poplack and D.Sankoff's definition of a 'nonce' loan (Sankoff, D. and Poplack 1981; Sankoff, D., Poplack and Vanniarajan 1990a). However, if we recognise that the PNG English pronunciations of these 'borrowed' words would already resemble the forms found in Tok Pisin data, then they may, in G.Smith's data from English-educated schoolchildren, represent insertions of words from English, rather than well-established words in Tok Pisin. Convergent pronunciations make distinguishing code-switches from established borrowings more difficult, yet they also make such code-switching a much less taxing and simple route to the adoption of new vocabulary.

G.Smith's transcripts have the five basic vowels and include five diphthongs: /ai/, /au/, /ia/, /iu/, /oi/ (Smith, G. 2002: 44). Though /ia/, in *hia* 'here', is an addition to Mihalic's (1971) set, all of these except /iu/, in *niu* 'new', are included in earlier descriptions of Tok Pisin. This additional dipthong may be considered to have come along with the borrowing of *new* from English for the name of the Independent country of Papua New Guinea: *Niugini* for many Tok Pisin speakers. G.Smith also gives examples (Table 5.5) of non-standard diphthongs /ai/, /ei/, and /ou/ found in his data, and ascribes them to the influence of the English pronunciation of words common to Tok Pisin and English. Of these, only /ou/ is not found in Laycock's expanded inventory (Table 5.4).

Table 5.5: Tok Pisin Phonology - sample non-standard diphthongs (ref. G.Smith 2002: 48)

Meaning	Expected form	Attested form
hole	*hul*	*houl*
road	*rot/rod*	*roud*
tail	*tel*	*teil*
wild	*wel*	*wail(d)*
sew up	*samapim*	*soumapim*

In my data there are many examples of variation in the form of vowels, both in the English recorded there and in Tok Pisin. For the short/long vowel distinction [ɪ] vs. [iː] fewer than 6% of tokens showed long vowels in Tok Pisin, showing on shared or English vocabulary in Tok Pisin utterances, as in examples (29) and (30):

(29) TiT#08-506_L14

piklnlni tu i sliːp ye
[pikɪnɪni tu i sliːp jɛ][35]

child too PM sleep REL

pikinini tu i slip hia

'(the) child too (was) asleep'

(30) TiT#10b-03_L18

traim nupuhla ***gluhooːb kawn biːf fleivuhs***
[tɹaɪm nupəla gləʊːb kɔn biːf fleɪvəs]

try-TR new-ADJ Globe corn beef flavour-PL

traim nupela ***Globe Corn Beef flavours***

'try new Globe Corn Beef flavours'

---

[35] To illustrate the points made here I include an IPA phonetic transcription of the example sentences below the transcription orthography. Elsewhere, throughout this dissertation, the transcription orthography alone is given. See the Transcription conventions in Appendix III for the practical orthography and details of the representation of IPA vowels.

In contrast, in the recorded PNG English, over twice as many long vowels were pronounced in a shortened form than in a long form, as in (31) and (32):

(31) TiT#35b-04_L134

onuhbuhl memba fo xxx # ***pli:s bi sited***
[onəbəl mɛmba fo ... ... pli:s bi sitɛd]

honourable member for xxx, ***please be seated***
'Honourable Member for xxx , please be seated'

(32) TiT#24b-15_L1092

wi biliv In uh splrits a ?
[wi biliv ɪn ə spɪrits a]

we believe in the spirits a ?
'we believe in the spirits eh?'

The English sections of my data therefore show a similar reduction in vowel length in PNG English to that described in Tok Pisin.

Similar trends in the high/low vowel distinction can be observed. There was a tendency towards higher vowels in PNG English speech than would be expected in British or Australian pronunciations, as in (33) and (34):

(33) TiT#35b-02_L196

wi go bek # tu a: pipuhl
[wi go bɛk ... tu a: pipəl]

we go back, to our people
'we go back to our people'

(34) TiT#35b-02_L246

paluhmuhn ken set is on prosi:dyuhs
[paləmən kɛn sɛt is on prosi:djəs]

parliament can set its own proceedures
'parliament can set its own proceedures'

Significantly, in Tok Pisin a low vowel appears on the modal *ken* with some frequency, and in English there was similar variation in the pronunciation of the vowel in the modal *can*. While the canonical pronunciation indicated for this particle is [ken] in Tok Pisin and [kæn] in

standard English varieties, around half of Tok Pisin uses were found with the 'front a' [kan] pronunciation, and, conversely, a similar proportion of English uses were found with the [ken] pronunciation. This strongly indicates that the value of this central vowel is highly variable, with both [e] and [a] realisations being entirely acceptable in either Tok Pisin or PNG English. While Tok Pisin *ken* is characterised as being a marker of permission, G.Smith observed in his corpus instances of it being used to indicate ability, a meaning traditionally carried by the modal *inap* 'be able to'. In the Tok Pisin of my corpus *ken* expresses this additional meaning in around a third of appearances, but the [kan] pronunciation is two and a half times more likely to be associated with physical ability. There is some correlation between form and meaning, indicating that the *ken/kan* distinction in Tok Pisin is not always free variation, though I did not find a similar correlation in the English of the corpus. [36]

The other vocalic variables, including diphthongs, showed similar patterning in this dataset, with shortened sounds without the second vocalic element being present throughout the English in the sample, and with full length English diphthongs appearing on some borrowed or shared words in Tok Pisin. In general my data illustrates that PNG English is characterised by similar vowel sounds to that described as standard in Tok Pisin, and this variation occurs alongside a presence of English-type vowels in Tok Pisin where the words are shared by both languages, or represent a borrowing from English. The significance of this finding in relation to mixed language production will be considered in detail in Chapter 7.

---

[36] One possibility is that the presence of *kan* in Tok Pisin with the meaning 'is able to' is a result of a code-switch from English, though this would be constrained under the MLF and only viable under Congruent Lexicalization (see Chapter 9).

5.1.2 Consonants

Accounts of substrate influences (Goulden 1987; Keesing 1988; Siegel 1999) in Tok Pisin have shown variation in consonants. As discussed above, G.Smith (2002: 43) and Laycock (1985b) specifically address the impact of indigenous vernaculars on the way Tok Pisin is spoken. G.Smith's study suggests that the single most common source for new phonemes in Tok Pisin is English.

As with vowels, previous accounts of PNG English have focussed on similarities between the consonants produced by different New/World Englishes and consistent differences between various indigenized varieties and inner circle 'Standard' English pronunciations. However, we can see that these same differences between inner circle Englishes and outer circle 'New' Englishes are characteristic of Tok Pisin pronunciations also.

The 'core' consonants of Tok Pisin have been variously described as between fifteen (Hall 1943) and eighteen (Mihalic 1971; Laycock 1985b) phonemes. This variation is partly due to the presence/absence of <y> [j] and <j> [ʤ] which only appear word-initially in Tok Pisin, and to the exclusion of <h> [h] from Hall's account (Hall 1943). G.Smith refers to Mihalic as the standard for comparison, but notes that those may not refer to the common spoken forms, especially in the case of rapid speech (Smith, G. 2002: 44). However, where Mihalic includes <j> in his inventory, as a word-initial option, G.Smith leaves it out of the core phonology of the creolized variety, and includes it under his list of new phonemes introduced from English in his data, where [ʤ] appears word-internally also.

Mihalic's inventory of Tok Pisin phonemes is given in Table 5.6:

Table 5.6 : Core Tok Pisin Phonology - consonants (adapted from Mihalic 1971)

orthography	phoneme	*Tok Pisin*	*English*	orthography	phoneme	*Tok Pisin*	*English*
b	[b]	*banana*	*banana*	n	[n]	*nau*	*now*
d	[d]	*dai*	*cease*	ng	[ŋ]	*bringim*	*bring*
f	[f]	*faiv*	*five*	p	[p]	*paip*	*pipe*
g	[g]	*givim*	*road*	r	[r]	*ran*	*run*
h	[h]	*haus*	*house*	s	[s]	*salim*	*send*
j	[ʤ]	*jun*	*june*	t	[t]	*tu*	*also*
k	[k]	*kam*	*come*	v	[v]	*hevi*	*worry*
l	[l]	*longpela*	*long*	w	[w]	*wara*	*water*
m	[m]	*mama*	*mother*	y	[j]	*yu*	*you*

For variable realisations of these phonemes in certain contexts, see Mihalic (1971).

Laycock's account of transformations in consonants between Australian English and Tok Pisin on common words is reproduced below as Table 5.7. However, it should be noted that while Laycock (1985b) does not give examples of /w/ /j/ and /v/ in his column of Tok Pisin options, these phonemes are present in some of the words he gives as examples.

In standard accounts of Tok Pisin fricatives /f/, /v/, /ʃ/, and /ʒ/ are traditionally not recorded, and Laycock describes these as contrastive phonemes only in anglicised varieties. The voiceless stops /p/, /t/, and /g/ are generally unaspirated, and voiced stops, /b/, /d/, and /k/, are de-voiced in word-final positions. Significantly, the place of articulation for the phonemes /t/ and /d/ may be further forward than in 'standard' English, with the tongue placed against the teeth and the alveolar ridge. The voiced alveolar consonant /r/ may be merged with the alveolar lateral approximant /l/, especially where there is no phonemic distinction between them in speakers' traditional languages, however, both are commonly pronounced as flaps in Tok Pisin: /r/ is [ɾ] for many speakers, and may be trilled as [r] in some dialects. Australian,

British, and American English produce /r/ as an approximant [ɹ]. In my data, both in PNG English and Tok Pisin [r], [ɾ], and [ɹ] occur.[37]

Laycock's table (Table 5.7) illustrates expected Tok Pisin options replacing English phonemes /ʃ/, /ð/, /ʧ/, /ʒ/, /ʤ/, /θ/, and /z/. However, accounts from Hall (1943) onwards have noted the presence of these phonemes in the Tok Pisin of some speakers. As is the case with vowels, these phonemes have commonly been ascribed either to foreign accents or to anglicised Papua New Guinean accents. Mihalic leaves this set of 'English' phonemes out of his standardising orthography because it is based on a conservative MOMASE variety.

[37] This is a potential area for further analysis: to explore variation in /r/ in PNGE and in TP.

Table 5.7 : Core Tok Pisin Phonology - consonants (adapted from Laycock 1985 : table 1)

English	Tok Pisin	Examples		English	Tok Pisin	Examples	
p	p	*pig*	*pik*	b	b	*boat*	*bot*
		-			p	*?*	
t	t	*tongue*	*tang*	d	d	*dog*	*dok*
		-			t	*done*	*tan*
		-				*road*	*rot*
tʃ	s	*chain*	*sen*	dʒ	s	*judge*	*jas*
	si	*church*	*sios*		si	*Japan*	*Siapan*
k	k	*keel*	*kil*	g	g	*go*	*go*
		-			k	*bag*	*bek*
f	p	*friend*	*pren*	v	b	*veranda*	*veranda*[38]
	b	*flash*	*bilas*		p	*five*	*faiv*[39]
ʃ	s	*shame*	*sem*	ʒ	t	*measure*[40]	*meta*
	si	*shirt*	*siot*				
θ	t	*think*	*tingim*	ð	t	*brother*	*barata*
		-			r	*other*	*narapela*
s	s	*sun*	*san*	z	s	*razor*	*resa*
m	m	*mouth*	*maus*	n	n	*no*	*na*
ŋ	ŋ	*wrong*	*rong*	w	w	*one*	*wan*
y[41]	y	*young*	*yangpela*	l	l	*louse*	*laus*
r	r	*rope*	*rop*	h	h	*house*	*haus*
ø	ø	*open*	*opim*				
	h	*eye*	*hai*				

[38] Laycock's original table contains this reduplication, though Laycock presumably meant to say *beranda*.
[39] Laycock's original table contains this example, which omits the recorded transformation.
[40] Laycock's original table duplicated English /z/ being transformed to Tok Pisin /s/ here, with empty example slots being filled with a '?' mark. The *measure*→*meta* pronunciations are recorded in Mihalic (1971).
[41] Laycock's original table, favouring standard English orthography, used /y/ for [j] here.

In accounts of PNG English (Smith, A.-M. 1986, 1988a; Holzknecht 1989), and other New/World Englishes (Platt and Weber 1980: 37-43; Platt, Weber and Ho 1984), these same phonemes have been discussed. In particular, the replacement of voiced and voiceless fricatives, /ð/ and /θ/, with the stops /d/ and /t/ has been observed in Indian, Sri Lankan, West Indian, and some African Englishes, as well as in PNG English.[42] In addition, the pronunciation of /ʃ/ as [s] has been noted in Hong Kong, some East African Englishes, and PNG English, along with devoicing of its voiced post alveolar counterpart /ʒ/. Similarly, voicing of the alveolar fricative /z/ may be inconsistently applied, with [s] occurring in PNG English as it does in West African and Indian English, though such variable pronunciations of word-final /s/ allows for the retention of standard English s/z allomorphy by some speakers. In many varieties of indigenized Englishes, as in PNG English, initial voiceless stops may be unaspirated, and in some East African Englishes, as in PNG English, the [h] sound at the beginning of a word is variably omitted. Final stop consonants may be unreleased, and therefore sound as if they have not been articulated, even on words within clauses and phrases. This produces pronunciations of *and* as [æn] in PNG English and educated Ghanaian English, and has an impact on the salience of past tense forms of verbs using the *-ed* ending. Furthermore, PNG English, as with other indigenized Englishes, shows differences in the use and type of /r/ produced: with [r], [ɹ], and [ɾ] all occurring.

Though PNG English shares many of these characteristics of pronunciation with various other indigenized Englishes, it shares all of them with Tok Pisin. Furthermore, while the

---

[42] In the indigenized Englishes of Malaysia, Sri Lanka, Singapore and some African Englishes, [d] or [t] are used and then the air released with some friction, to produce [dð] or [tθ] (Platt, Weber and Ho 1984).

information presented here in Table 5.6 and Table 5.7 represents the common 'core' phonology of Tok Pisin, G.Smith found all of the 'replaced' English phonemes in his creole speech corpus (Smith, G. 2002).

The extent to which such phonemes appear in Tok Pisin as a result of an acquisition of English phonology by schoolchildren and adult bilinguals is something that G.Smith describes as an interference phenomenon, relating it to shared vocabulary items, and echoing the language of Weinreich and Haugen (see s3.2). However, the longevity of such forms in Tok Pisin, albeit limited to certain speakers and the 'Urban' anglicised variety, suggests that these pronunciations are not entirely novel borrowings from English. In fact, the adoption of new phonemes, reversing in some cases the standard transformations detailed by Laycock, again increases the congruence between Tok Pisin and English for some speakers, effectively blurring particular distinctions represented in the standard orthography of Tok Pisin.

In my data there are similarities between the consonants of PNG English and of Tok Pisin, just as there are similarities between their vowels (s5.1.1). However, there are some clear differences in the patterning of different consonants, for example /θ/ and /ð/.

In my data, /θ/ was realised as [θ] (marked in the corpus data as (th) (see Appendix III)) in both Tok Pisin and English, and it was also realised as [t] in both languages. In Tok Pisin /θ/ was sometimes present on words shared with English, as in (35) and (36):

(35) TiT#10b-07_L470

disela kain samthin olsuhn
[disɛla kaɪn samθin olsən]
DET kind something same
dispela kain samting olsem
'this kind of thing'

(36) TiT#10b-07_L233

athin ol i mas hamamas
[aθin ol i mas hamamas]

perhaps 3pl PM MOD happy

ating ol i mas hamamas

'I think they should be happy'

However, there was variation even in the Tok Pisin of a single speaker, compare (36) with (37):

(37) TiT#10b-07_L261

na napla lilik tosawe ah- liklik atin inap lo ol o ?
[na napla lilik tosawɛ ah- liklik atin inap lo ol o]

CONJ another little information ah- small perhaps enough PREP 3pl or

na narapela liklik toksave ah- liklik ating inap long ol o ?

'and another little comment ah- perhaps is (it) okay for them (listeners)'

In the English of my sample, while /θ/ was realised in the majority of cases there was a clear tendency for speakers to produce /t/ or /d/ on words around a third of the time. Examples (38) and (39) show this variation:

(38) TiT#24b-15_L139

jul stei dea faw tu o trri manths
[jul steɪ dɛa fɔ tu o tri manθs]

2sg.MOD.FUT stay PRON.LOC PREP two CONJ three month-PL

you'll stay there for two or three months

'you will stay there for two or three months'

(39) TiT#08-507_L9

fuhs tings fuhd
[fəs tiŋs fəd]

first thing-PL first

first things first

'first things first'

Significantly, this variation appeared most frequently on words that appear in both the English and the Tok Pisin lexicons: that are common to both languages, as in both the examples above.

Where /θ/ showed patterning that indicated some interference from Tok Pisin pronunciations in PNG English and vice versa, a different pattern could be observed for /ð/: variation was far more widely attested in my core corpus data. In particular, the English of this sample showed a strong tendency to realise /ð/ as [d] four times as often as [ð] (marked in the corpus data as (dh)). Though pronunciation of /ð/ as [d] is a feature of other indigenized Englishes it is also characteristic of conservative Tok Pisin. The high frequency of this variable in English corpora is due to its presence on definite determiners. Variation remained, however, even for individual speakers, as in (40):

(40) TiT#35b-02_L106

if yu luk uht dhuh kawst of da palamuhnt miting # hiuh
[if ju luk ət ðə kɔst of da palamənt mitiŋ ... hiə]

COND 2sg look PREP DET cost PREP DET parliament meet-CONT # here

if you look at the cost of the parliament meeting, here

'if you look at the cost of the parliament meeting, here'

In addition, this phoneme was occasionally omitted by speakers altogether when speaking English.

In Tok Pisin utterances, this variable did appear as [ð] on English words, as in (41):

(41) TiT#10b-07_L108

dhas wai suhming i go olsem athing uh- # ***in A we*** em i gudpla na # em nogut a ?
[ðas waɪ səmiŋ i go olsɛm aθiŋ ə ... in æ wɛ ɛm i gudpla na ... ɛm nogut a]

that.is why something PM go same I.think uh- # in a way 3sg PM good-ADJ CONJ # PM no.good Q.Tag

that's why samting i go olsem ating uh- , ***in a way*** em i gudpela na , em nogut a ?

'that is why something like that I think uh-, in a way it (is) good and, not good eh?'

As we will see in Chapter 7, a code-switching analysis would suggest that ['ðæs'waɪ] *that's why* in (41) is not Tok Pisin at all, and the role of phonology as an indicator of code-switching will be considered in detail there.

5.1.2.1 Consonant Clusters

Hall's orthography of Tok Pisin (Hall 1943) gives a large set of consonant clusters evident in Tok Pisin. While his work includes American and British second language Tok Pisin varieties, in distinguishing consonant clusters he acknowledges characteristics of Melanesian pronunciation. The insertion of unstressed vowels between consonants allows for the maintenance of a regular Consonant-Verb (CV) syllable structure to be maintained in the pronunciation of those English-derived words where vowel diphthongs and consonant clusters occur. In a small study of a single Tok Pisin speaker, Pawley (1974) successfully identified the relationship between vowel epenthesis and speed of delivery, uncovering a link between the rate of speech or degree of emphasis being placed on a word and the appearance of the epenthetic vowel. In present day Tok Pisin epenthesis appears more as an expressive strategy than an interference effect. As inserting the extra vowel sound distorts the phonological form of a word derived from English away from a characteristically English shape it is a significant strategy in Tok Pisin, useful for its capacity to clearly identify a word as Tok Pisin, as opposed to English, through its delivery. So, while the sounds [plɛs] might indicate the English noun for a physical location, *that place over there*, or an English verb, *place that over there*, or a Tok Pisin noun, "village", the form [pɛlɛs] can only refer to the Tok Pisin referent. Vowel epenthesis thus becomes a language-indexing feature of Papua New Guinean speech: producing emphatic forms that can be used to distinguish Tok Pisin and English homophones.

The process of adaption of English consonant clusters to CV syllable structure through vowel epenthesis may be seen as an additional 'substrate' feature of Tok Pisin, alongside the removal of fricatives, discussed above (s5.1.2).

Similar rules seem to apply to some varieties of PNG English, where consonant clusters allowed in 'inner circle' Englishes are often modified. As with other New/World Englishes (Platt, Weber and Ho 1984: 43-46), this may be accomplished by the introduction of a short vowel sound, as in Indian English [gəlas] 'glass' [pəlai] 'play'. This resembles the vowel epenthesis of Tok Pisin. Another verbal strategy is not to pronounce the final or middle consonants of the group. This is especially frequent in PNG English. While speakers of inner circle Englishes may reduce the pronunciation of consonant groups in rapid speech, in PNG English, as in conservative Tok Pisin (Laycock 1985b), consonants in groups may not be pronounced at all, or, as discussed above, deliberately separated by vowels in careful speech.

One consequence of such modifications to the production of consonants word-finally (also noted above (s5.1.2) for unreleased stops) is that the English *-ed* morpheme becomes obscured in PNG English where the verb stem ends on a consonant. As A.-M.Smith (1986) observes, the change in phonological form removes the indicator of standard English aspect marking on the verb. Consequently, present and past tense verbs are realised in the same form, and tense is often expressed using adverbs of time, is contextually derived, or uses a dummy auxiliary construction, all of which allow the use of an invariant verb form, as A.-M.Smith illustrates (42):

(42)

The men had used force and so a man had to lose his life.

'The men had used force and so a man lost his life.'

(Smith, A.-M. 1988b: 122)

In this, PNG English resembles basic forms of Tok Pisin where, before the +anteriority marker *bin* was developed, similar strategies of overt marking of past time through adverbial constructions were used (Sankoff, G. and Brown 1976) (see also s5.3.8, below). This phonological effect of reducing word final consonant clusters, sometimes described as T/D

deletion, has been observed in Atlantic creoles (Sebba 1993; Patrick 1999) and African American Vernacular English (Rickford 1999; Bender 2004).

Laycock gives a short list of phonemic consonant clusters in Tok Pisin, for which vowel epenthesis is not a factor. These are given in Table 5.8:

Table 5.8 : Tok Pisin Phonology - phonemic consonant clusters (adapted from Laycock 1985)

CLUSTERS (NOT PERMITTING EPENTHETIC VOWELS)							
English	**Tok Pisin**	***Examples***		**English**	**Tok Pisin**	***Examples***	
mp	mp	*company*	*kampani*	mb	mb	*number*	*namba*
	m	*lamp*	*lam*				
nt	nt	*mountain*	*maunten*	nd	nd	*window*	*windua*
	n	*cent*	*sen*		n	*hand*	*han*
nč	nis	*winch*	*winis*	nǰ	nis	*change*	*senis*
	ns	*?*			ns	*engine*	*ensin*
ŋk	ŋk	*monkey*	*manki*	ŋg	ŋg	*finger*	*pinga*
		bank	*beng*				

Referring to Laycock's table, these homorganic nasal clusters appear in Tok Pisin, when compared with their English counterparts, only when they fall across a syllable boundary. Where the same consonant cluster occurs in the English word-finally, it is, in Laycock's examples, modified: in some cases the final consonant is omitted; in others, such as /nʃ/ or /nʒ/ (Laycock uses the older notation /nč/ and /nǰ/), vowel epenthesis occurs. These are the same techniques observed in PNG English and other New/World Englishes.

In addition to these, G.Smith identifies non-standard consonant clusters in established words in his data, indicating additional impact of English phonology on creole Tok Pisin pronunciation (Smith, G. 2002: 48).

In the corpus data, non-standard consonant clusters appear in borrowings from English, as well as shared English words in Tok Pisin. Examples (43) and (44) show variable realisation

of consonant clusters by a single speaker seconds apart where *ausaid* [ausaɪd] resembles the English *outside* [autsaɪd]. Example (45) presents an identical form produced by a different speaker:

(43) TiT#24b-15_L845

so taim yu kam ausaid em
[so taɪm ju kam aʊsaɪd ɛm]

DISC.CONJ time 2sg come outside 3sg

so taim yu kam ausaid em

'... so when you come outside it ...'

(44) TiT#24b-15_L896

so taim yu kam ***autsait***
[so taɪm ju kam aʊtsaɪt]

DISC.CONJ time 2sg come outside

so taim yu kam ***outside***

'... so when you come outside...'

(45) TiT#10b-05b_L138

ol man no go sali o samti ***autsaid*** i stap
[ol man no go sali o samti aʊtsaɪd i stap]

PL man NEG go sell-TR PL something outside PM CONT

ol man no go salim ol samting ***outside*** i stap

'men (do) not keep selling things outside'

In the English of my sample consonant clusters are often retained, but there is a strong tendency to reduce word-final clusters, as in (46):

(46) TiT#35b-04_L40

bat a dis poin in taim # mista Aktin spika # wi don hav kworuhm
[bat a dis poin in taɪm ... mista æktin spika ... wi don hav kworəm]

CONJ PREP DET point PREP time # mister act-CONT speaker # 1pl do.AUX.NEG have quorum

but at this point in time , mister acting speaker , we don't have quorum

'but at this point in time, mister acting speaker, we don't have quorum'

This finding accords with A.-M.Smith's observations on T/D deletion (Smith, A.-M. 1986), and is a common feature between PNG English and other world English varieties (Kwamwangamalu and Lee 1991; Ho and Platt 1993; Gupta 1994), as well as Caribbean creoles (Sebba 1993; Patrick 1999).

5.1.3 Phonological Reduction

The studies of G.Sankoff and Brown (1976), Romaine (1999), and G.Smith (2002) all identified clear tendencies towards the production of reduced phonological forms of many common words as a standard feature of creole Tok Pisin. Some work has been done to assess whether such reduced forms indicate grammaticalization is underway, or if they are becoming standard clitics (Sankoff, G. and Laberge 1980 (1973); Romaine and Wright 1987: 62). First-language speakers may have adopted these changes most readily, but such innovations have become widespread in Tok Pisin as it is spoken today. The result is a more compact and streamlined verb phrase (Sankoff, G. 1977; Sankoff, G. and Laberge 1980 (1973)) in the characteristically rapid and fluid speech of the majority of fluent speakers.

In my corpus reduced forms for the prepositions *long* and *bilong* and for the determiner *dispela* occur with high levels of consistency, a finding in line with G.Smith's observations. However, there is little support in this adult data for the new forms largely attested in G.Smith's schoolchildren's language. Table 5.9 gives examples of the reduced forms of *dispela* in my data

Table 5.9 : Tok Pisin Phonology – reduced forms of *dispela* in the core corpus

[displa]	[disela]	[ispela]
[disla]	[disel]	[isela]
[disə]	[dise]	[isla]

Across the uses of the suffix in the corpus, however, there are close to twice as many occurrences of *-pela* as of *-pla*, indicating a resistance to generic change.

Where G.Smith (2002) finds many examples of contracted forms in his corpus of schoolchildren's monologues, in my samples of adults' mixing Tok Pisin with English there are similar contracted forms within stretches of rapid Tok Pisin. Example (47) shows reduction effects on the verb phrase *i wok go long tanimtanim* which represent the same phenomena found by G.Smith:

(47) TiT#35b-202_L357

ol ***benks*** blo mi:pela: i: wawk g lo tanitani mi:pela:
[ol bɛnks blo mi:pɛla i: wɔk g lo tanitani mi:pɛla]

PL banks PREP.Poss 1pl.excl PM work go PREP turn(reduplicated) 1pl.excl

ol ***banks*** bilong mipela i wok go long tanimtanim mipela

"our banks are working to corrupt us"

Significantly, this speaker is an elder statesman well known for his fluent Tok Pisin, but who is not a first-language, 'creole' speaker. His reduction of *-im* to *-i* on *tanimtanim* 'to turn and turn' is characteristic of MOMASE regional speech in G.Smith's data. G.Smith's observations in creole speech are found throughout the adult population, reinforcing the view that it is degree and frequency of use, rather than the sequence of language acquisition, which relates to the most developed forms of Tok Pisin.

An additional trend in spoken Tok Pisin is clearly the omission of the final consonant from the *-im* syllable. As with other reduced forms, *-i* appears as much a feature of fluent speech as it is a characteristic of creolization. While G.Smith identifies the reduced form as a notable feature of his corpus of children's' creole (Smith, G. 2002: 60-62), I have found it in the speech of adults also.

The examples used in this dissertation contain many instances of these reduced forms, which are retained in the transcriptions provided.

5.1.4 Summary

In terms of an understanding of the vowels of Tok Pisin and English two observations may be made. Firstly, that Tok Pisin has been spoken with a wider range of vowels than Mihalic's 'core' set of five common vowels and three diphthongs since at least the 1940s, with the expanded set of phonemes available being consistently related by Hall, Laycock, and G.Smith to exposure to English. The Tok Pisin speakers who contributed to the core corpus frequently produce 'anglicised' vowels from this expanded set. Secondly, that the local PNG dialect of English, as described by A.-M.Smith, has a similar vowel space to Tok Pisin, indicating congruence between the vowel systems of these two languages.

In terms of consonants, the corpus data supports a characterisation of PNG English as featuring variation from 'standard' English forms similar to those recorded by Laycock and others between Tok Pisin and Australian English forms, in particular on words shared between Tok Pisin and English, for some of my speakers. This new data also confirms an observed pattern of final consonant cluster simplification in PNG English.

In addition, phonologically reduced forms are as prevalent in this data as they are in earlier spoken corpora, and forms previously observed in children's creole speech can now be observed in the language of non-creole, adult speakers.

These findings have a significant impact on expectations of phonological congruence between some varieties of Tok Pisin and of English as it may be spoken in Papua New Guinea. Just as Tok Pisin pronunciations may be 'anglicised', so too PNG English pronunciations are 'Tok Pisin-ised'. Some of the phonological changes result in the non-realisation of grammatical morphemes, with consequences for the syntactic expression of tense. These changes result in distinctive variation within each language, and also in congruence between them.

The observation that phonological variation is encouraged by exposure to English pronunciations means that, against the backdrop of widespread education in English, the way a word may be pronounced in Tok Pisin can be very variable.

Both these observations have impacts on the similarity between Tok Pisin and English phonology, and on the congruence between them, with consequences for the ease and variety of language mixing behaviour.

5.2 Morphology

Both Tok Pisin and English are characterised by low levels of morphological complexity, especially when compared to other languages in Melanesia and Europe, respectively. In their standard forms they also have little morphology in common, and have quite different verbal, adjectival, and nominal morphology. However, the most recent accounts of written Tok Pisin (Verhaar 1995) and of the spoken language (Smith, G. 2002), as well as those of more varied datasets (Harvey 2007; Wakizaka 2008; Devette-Chee 2011; Walczyński 2012), include examples of distinctively English morphology being used. Discussion of the presence of the English pluralising *-s* suffix and its role in Tok Pisin grammar, and of the past tense *-ed* and progressive *-ing* has centred around accounting for whether such morphology is productive, and the extent to which it represents the introduction of new, active, grammatical particles produced as part of processes of decreolization under influence from English.

However, alternative explanations for the presence of English morphemes in Tok Pisin production are possible, as shown below and in the following chapters.

In accounts of inter-sentential code-switching between languages, such as Myers-Scotton's MLF model (Myers-Scotton 1993a) and subsequent modifications, morphosyntactic integration is a particular focus. The System Morpheme Principle specifically prohibits the presence of certain types of functional morphology in mixed language, unless the items

carrying this morphology are also grammatically well-formed with reference to the L1 grammar. The presence of English morphology in Tok Pisin is therefore particularly important for analysis from the point of view of code-switching theory.

In this section I consider some characteristic features of Tok Pisin morphology and consider whether the presence of English morphology in Tok Pisin corpora can be understood as grammatical borrowing or a characteristic of code-switching. A comparison of Tok Pisin and English morphology will also be made, and examples from this English language data used to show bi-directional mixing in Papua New Guinean speech.

5.2.1 Verbs : marking Transitivity *-im*

The verbal transitivity marker, the *-im* suffix, is one of the most salient parts of the Tok Pisin grammar, attested in early jargons and their successor pidgins throughout the Pacific region. While Keesing (1988) argues that its role is patterned by Melanesian substrate languages, its early appearance in Australia (Baker 1993) suggests that the Melanesian influences are appropriately considered as reinforcing rather than originating its use. The derivation of *-im* is most likely to be from the English third person masculine pronoun *him*, and its role as transitivity marker has developed from this.

In my Tok Pisin data are many examples of the deletion of the object of transitive verbs, a pattern observed and discussed by Lattey (1979). Some of these cases indicate how the *-im* suffix may have developed from the object of transitive predicates to a systematic affix. The transitive verbs (V-TR) *developim* 'develop' and *manesim* 'manage' are Tok Pisin verbs formed from English stems. In (48) *developim* appears in a mixed language sentence without an object noun. The actual object *em* appears in an earlier clause. Example (49) shows the deletion of an expected object noun or the pronoun *em* in the question.

(48) TiT#35b-203_L472

na ie ***develop***im ***tu feis thA neks mileniuhm***

CONJ REL develop-TR Ø PREP face DET next millennium
na iet ***develop***-im ***to face the next millennium***
"and indeed develop (them) to face the next millennium"

(49) TiT#35b-202_L164
sapozi ***manizmen*** blo kantri i bagarap huset i ***maneze***im Ø?
suppose-TR management Prep.POSS country PM fail who PM manage-TR Ø
saposim ***management*** bilong kantri i bagarap husat i ***manage***-im ?
"if (the) management of (the) country is a mess who is managing (it)?"

An object is implied by the presence of the transitive suffix *-im*, facilitating the frequent deletion of objects in transitive clauses.

There is a small class of transitive verbs that conventionally do not receive the suffix, e.g. *gat* 'to have', and a larger set that are followed by the preposition *long* when the *-im* is not included, such as: *tok(-im / +long)* 'to talk'; *spet(-im / +long)* 'to spit'; *wet(-im / +long)* 'to wait'. Significantly, this means that even monolingual Tok Pisin speakers have examples of unmarked transitive verbs in their grammar.[43] Standard English grammar does not mark transitivity on the verb.

In contemporary Tok Pisin, evidence of the reanalysis of *-im* as a universal transitive suffix by at least some children and the appearance of forms such as *gatim* in Highlands Tok Pisin as a consequence of overgeneralisation (Smith, G. 2002: 60) indicate the salience of this suffix in marking Tok Pisin speech.

---

[43] Some Tok Pisin verbs are intransitive in one form and transitive in the *-im* form: e.g. *wara i boil* 'the water boils' (intransitive) and *em i boilim wara* 'he boils the water'.

The core corpus data features a large number of verbs apparently borrowed from English since Mihalic wrote his dictionary (Mihalic 1971). The currency of these verbs outside of the English-speaking community is unknown. However, the adult speakers contributing to the corpus data do show a precision in the application of the transitive suffix that distinguishes them from the younger speakers in G.Smith's corpus. In (50) a speaker introduces the English verb *circumcise* without an object, and also without the transitive suffix.

(50) TiT#24b-15_L289

nau yu: ol i ***suhkuhmsais yu go tu di hospituhl*** a ?

REL 2sg 3pl PM circumcise 2sg go PREP DET hospital Q.tag

nau yu ol i ***circumcise you go to the hospital*** a ?

'so you they (perform the) circumcision(,) you go (to) the hospital eh?'

In (51) the transitive use of the English verb requires the suffix:

(51) TiT#24b-15_L322

so # taim ol i ***sikuhmsais***im yu ...

DISC.CONJ # when 3pl PM circumcise-TR 2sg

so, taim ol i ***circumcise-***im yu ...

'so, when they circumcise you ...'

As I will discuss in depth in Chapter 7, the morphological integration of English verbs borrowed to expand the vocabulary of Tok Pisin is required under the MLF model, so the marking of over 100 different English verbs in Tok Pisin data with *-im* is to be expected. A code-switching analysis of examples such as (50) will also consider the role of the switch to English in encouraging the deletion of the *-im* suffix there.

This morpheme also appears throughout the English data, however, as in (52-53):

(52) TiT#35b-04_L75

hi neva korim ***korom***

2sg.masc NEG call-TR quorum

he never call-im ***quorum***

'he never called quorum'

(53) TiT#08-506_L62

tel him det al wan repotim nau

talk 2sg.masc DET 1sg want report-TR now

tell him that I want report-im now/nau

'tell him that I want to make a report now'

The root forms of both verbs are shared between Tok Pisin and English, and appear in these speaker's English in their Tok Pisin form. As was shown in the discussion of phonology, above, Tok Pisin and PNG English pronunciations, particularly on shared words, are similar. However, the appearance of transitive marking in PNG English is a significant diversion from standard English grammar in the direction of Tok Pisin. The word *kolim* 'to call x' has a clear object, and though *repotim* 'to report x' does not assuming a congruent usage (following Lattey (1979) and the observations above), this does not necessarily make the verb phrase ill-formed in Tok Pisin, or, by extension, PNG English,.

Taking these verbs to be code-switches from Tok Pisin the presence of the *-im* suffix alone does not mean it is an addition to the grammar of PNG English. This will be discussed in Chapter 7 in terms of the MLF model (Myers-Scotton 1993a).

However, as I discuss in Chapter 8, code-switching theory also provides an alternative explanation for such mixtures, which allows for the active presence of the *-im* suffix in PNG English.

At this point we can simply note the existence and role of *-im* in Tok Pisin, the variability of object-deletion, the presence of the suffix on English verbs appearing in Tok Pisin, and its occasional presence in PNG English.

5.2.2 Adjectives : *-pela*

The second productive suffix in Tok Pisin is the modifier *-pela* suffix, which is applied to most monosyllabic adjectives. Hall uses it as a morphological basis for his qualification of the class of adjectives (Hall 1943). However, the suffix may be omitted where a monosyllabic

adjective is used predicatively or is a past participle and is not required on polysyllabic adjectives or on a small class of irregular monosyllabic descriptive adjectives, such as *bruk* 'broken' and *les* 'tired' (Mihalic 1971).

In (54) the *-pela* suffix used distributively in an advertisement to cover the whole AdjP *blu na wait*:

(54) TiT#10b-03_L35

painim long liklik blu: na waitpela ***glowb kawn bi:f ti:n***

find-TR PREP little [blue CONJ white]-ADJ globe corn beef tin

painim long liklik blu na waitpela ***globe corn beef tin***

'find (them) in (the) little blue-and-white globe corn beef tin(s)'

In (55) is an example of a predicative use of *wait* 'white' without the adjectival suffix, though *waitpela* would also be valid:

(55) TiT#24b-04_L92

tesol em no gudpla lo # fainim abus lo nait liklik skin blen wait

but 3sg NEG good-ADJ PREP # find-TR game PREP night little skin PREP.poss.3sg white-ø

tasol em no gudpela long, painim abus long nait liklik skin bilongen wait

'but they (are) no good at finding game at night a bit (because) their skin (is) white'

In addition to its role marking adjectives, *-pela* is a standard part of Tok Pisin nominals, where it is added to English roots to indicate number. Table 5.10 presents the traditional counting paradigm.

Table 5.10 : Tok Pisin Morphology - cardinal numbers

numeral	nominal
1	wanpela
2	tupela
3	tripela
4	popela/fopela
5	paipela/faivpela
6	sikispela
7	sevenpela
8	eitpela
9	nainpela
10	ten
11	wanpela ten wan/ elevenpela
12	tenpela tu / twelvpela
13	tirtinpela
20	tupela ten
21	tupela ten wan
22	tupela ten tu
23	tupela ten tri
97	nainpela ten seven
100	tenpela ten/ wan handret
135	tirtinpela ten faiv/wan handret tenpela ten faiv

Alternative pronunciations for some numbers are commonplace, demonstrating the basic variation in phonology of Tok Pisin speakers. However, while the traditional Tok Pisin paradigm allows for simple counting to ten and the expression of base 10 values through additive logic, it is also increasingly the case that English nominals are used for high values. The word *handret* 'hundred' is now commonplace as are individuated nominals for values in the teens: both systems modelled on English.

It is evident that as the teaching of mathematics in English progresses this is one area of change and clear 'anglicisation' in Tok Pisin. Generally, small numbers are regularly expressed with the traditional *-pela* suffix as indicator, but it is unusual to hear younger speakers using forms like *tupela ten wan*, most would prefer *twenty/tweni wan* (Smith, G. 2002). This may be an increasing tendency for all numbers, with calques from the English being generally preferred for all nominals over four. Mihalic (Mihalic 1971: 21) notes that the *-pela* suffix is required for all cardinal numbers from one to twelve when they are used attributively, as predicate adjectives, or as substantives, the 'bare' forms are used for expressions of worth or price, time, or in

arithmetic, reflecting 'anglicisation'. However, while *wanpela, tupela, tripela,* and *fopela* are still commonly heard, use of the *-pela* suffix or the associated traditional counting system is increasingly marked.[44]

In PNG English, with other New/World Englishes, adjectives formed from the participle V+-*ed* (past) or V+-*ing* (present) forms of verbs in standard English are sometimes used differently: the *-ed* ending may not be pronounced where it follows another consonant (see s5.1.2.1, above), and the *-ing* form tends to be used more overtly (see also s5.3.8 and Smith, A.-M. (1988b) for its role marking aspect). In s5.2.6 these affixes in Tok Pisin corpora are discussed .[45] There is no specific adjectivizing morphology recorded in existing accounts of PNG English equivalent to the Tok Pisin *-pela* suffix.

5.2.3 Pronouns

The *-pela* suffix is also found in the Tok Pisin pronominal system, where it is applied to the first, second, and third person pronouns to indicate plurality.

Tok Pisin pronouns follow the paradigm illustrated by Table 5.11. The first and second singular pronouns *mi* and *yu* are modified by the *-pela* suffix to indicate plurality. In the third person there are separate roots for the singular and plural pronouns: *em* and *ol*. In its dual and trial forms the third person singular pronoun *em* is optional before numeral classifiers with the

---

[44] In my data numbers regularly appear without the *-pela* suffix, with *wanpela* appearing principally as a determiner. This reflects a general adoption of the English counting system over the Traditional Tok Pisin paradigm (see also Wakizaka (2008)). The freeing up of *wanpela* in this way may motivate the future grammaticalization of *wanpela* as a determiner (Sankoff, G. and Mazzie 1991), though I make no comment on this here.

[45] While it may be more properly considered under s5.3 on Morpho-syntax, below, I will mention here that the word order of adjective phrases can be different in PNG English compared to standard English: A-M.Smith gives as an example *This two last years* 'these last two years' (Smith, A.-M. 1978: 35).

-pela suffix. Keesing's detailed discussion of competition between substrate forms in the early life of Melanesian Pidgin English (Keesing 1988) indicates how the present system might have evolved. The stabilisation of variation in the pronominal system is seen as a significant stage in the standardisation of an interlanguage grammar.

Significant features of the pronominal paradigms of some indigenous languages find expression in Tok Pisin. This reflects the role of Tok Pisin as a language in use by many people whose approach to the fundamental categorisation and expression of person are coded by their *tokples* languages. The distinctive expression of a distinction between inclusive and exclusive first person forms is one such characteristic feature. Another is the expression of dual and trial number through an additive formula.

Table 5.11 : Tok Pisin Morphology - standard pronouns

		number		
person	**singular**	**dual**	**trial**	**plural**
1st	mi	(excl.) mitupela	mitripela	mipela
		(incl.) yumitupela	yumitripela	yumi
2nd	yu	yutupela	yutripela	yupela
3rd	em	(em)tupela	emtripela	ol

However, in the Tok Pisin of urban and first language speakers, those more detached from substrate language culture, variation in these areas of the system exists (Smith, G. 2002). For example, the trial pronouns *mitripela, yumitripela, yutripela,* and *emtripela* are not as widely used as the other forms, and are not included in some accounts of the language. As with the dual forms, trial pronouns are more commonly used in communities where traditional languages make use of such a distinction. Even the dual pronouns are more commonly used in some communities than others, though they are far more widespread in use. This is not to say that the dual and trial forms are necessarily regional forms, but they are clearly indicative of particular varieties of Tok Pisin. Similarly, forms such as *yumitupela* and *emtupela* are standard in some descriptions, but are not given in others. Such forms are highly

paradigmatic, and perhaps reflect the logic of second-language, 'pidgin', Tok Pisin. More irregular paradigms may develop within native or first-language communicative networks as a result of modifications to increase fluidity and rate of speech. These modifications could become nativized within a language culture where there are minimal requirements for analytic structures that are systematically transparent to listeners.

Variation in the pronominal paradigm is distinctively governed by regional norms, and there is also some indication of an incipient change in the application of the inclusive/exclusive distinction.

Prior to the present study, Romaine (1992) and G.Smith (2002) have both discussed 'anomalous' uses of the inclusive first person *yumi* where the addressee was clearly not included. In my data, a politician from the MOMASE region provides a similar example:

(56) TiT#35b-05_L57
yu tok yumi no sanap
2sg talk 2pl.incl NEG stand[46]
yu tok yumi no sanap
'you told us(inclusive) not (to) stand'

In (56) the speaker clearly did not include the addressee in the scope of *yumi*, yet the inclusive pronoun was used here and in several other similar utterances. This shows in adult speech the phenomena observed in children by Romaine and by G.Smith. Significantly, just as the lack of this distinction in Austronesian languages may undermine its use by Tok Pisin speakers

[46] This example is taken from a longer section: *yu tok yumi no sanap, mi nau mi rausim* 'you told us not to stand, I then I became angry'. The audio could be parsed as *yu tok yumi no sanap-0 mi, nau mi rausim* 'you told us not to stand me up, then I became angry', with the verb *sanap* taken as transitive with the 1sg pronoun *mi* as its object, rather than (as given here) as the intransitive form. Clearly, the absence of the transitive suffix on *sanap* favours the analysis here, and is reflected in the division of the example as presented.

within Austronesian language speaking communities, so too the lack of such a distinction in English provides an additional resistance to its use. For Tok Pisin speakers with little exposure to traditional vernaculars, the influence of the English pronoun system could be responsible for the loss of this distinction.

However, some examples of this distinction also appear in the English data through the use of Tok Pisin pronouns, as in (57):

(57) TiT#24b-15_L682
wi got trris we mipla ***li:vin***
1pl have trees LOC 1pl.excl live-CONT
we got trees where/we[47] mipela ***living***
'we (have) got trees where we(exclusive) (are) living'

There may be some motivation for the transfer of inclusive/exclusive distinctions into PNG English for at least some speakers as this has a significant cultural role. There is some precedent for culturally-specific modifications to the pronominal paradigm of English in some dialects which retain a plural form of the 2nd person pronoun such as *youse* 'you.pl'. However, given the observed variability in the maintenance of the inclusive/exclusive distinction in Tok Pisin, it remains unclear at this stage what direction their development will take. Significantly, the removal of this grammatical distinction not found in English from some varieties of Tok Pisin increases the congruence in their pronominal systems.

Some speakers also use the Tok Pisin first person singular *mi* instead of standard English *I*:

---

[47] Throughout the standard orthographic glosses for examples from the primary data of this core corpus I will include both Tok Pisin and English forms for homophones. Such homophonous diamorphs are considered particularly significant under e.g. Clyne's (1967) account of German/English CS, and are discussed in Chapter 8 with respect to Muysken's (1997, 2000) model of Congruent Lexicalization.

(58) TiT#08-507_L335

ivan de:n mi ***sti:l kam wiv pis in mai hat***

even then 1sg still come-0Past CONJ peace PREP 1sg.poss heart

even then mi ***still come with peace in my heart***

'even then I still came with peace in my heart'

In the Tok Pisin of my core corpus, there are a number of examples showing the appearance of English pronouns. Example (59) includes the flagged use of forms intermediate between the Tok Pisin 3rd person singular (ungendered) *em* and the English 3rd person singular (masculine) *him*:

(59) TiT#35b-05_L91

mi ***thank*** yu: long ### ***im***

1sg thank 2sg PREP 3sg.masc

mi tenk/***thank*** yu long, ***him***

'I thank you for, him'

Though the form attested here [ɪm], without the continuant /h/, resembles the Tok Pisin pronoun [em] the raising of [e] to [ɪ] has not been observed in Tok Pisin phonology and reanalysis of the 3sg pronoun under influence from English is a more cogent explanation than phonological variation. This analysis is supported by this speaker's use of a standard [em] for *em* elsewhere, by the flagged hesitation before the English pronoun in (59), and by its repetition more fluidly a few lines later (60):

(60) TiT#35b-05_L96

behain ***im*** bai ma:kim yu l em

after 2sg.masc FUT mark-TR 2sg PREP 3sg

behain ***him*** bai makim yu long em

'after he will record you on it'

A further example can be seen at (115), below.

Elsewhere the English pronoun *they* is used in Tok Pisin:

(61) TiT#24b-15_L463

dei guht ah- # wuhnpluh kain ***li:v*** ia

3pl have ah- # one-ADJ type leaf LOC
they got ah- , wanpela kain ***leaf*** ia
'they have ah-, one type of leaf there'

Does the presence of these pronouns in this data suggest an extension to the Tok Pisin paradigm through borrowing from English? A code-switching analysis allows us to view constructions such as those above without assuming a widespread change in the grammar of PNG English, and so the Insertion of pronouns will be discussed in Chapter 7. This data will also be revisited in Chapter 8, where an interpretation in terms of Congruent Lexicalization will be presented.

5.2.4 Plurals

A final significant area of Tok Pisin morphology is the expression of plurality. Where English uses a nominal suffix to indicate plurality, Tok Pisin uses the prenominal particle *ol*. As with many pidgin languages, early speakers of Tok Pisin made no systematic distinction between singular and plural nouns, number being either inferred from context or explicitly defined through the use of a numeral. Plural marking developed in Tok Pisin during the stabilisation phase (Mühlhäusler 1985d: 113). Following a possibly universal tendency also observed in other creoles, the third person plural pronoun became grammaticalized as a marker of plurality, becoming obligatory for first-language speakers and also being applied redundantly in many cases (Smith, G. 2004a). So (62-66) below illustrates the development of pluralisation in Tok Pisin.

(62) Basic pidgin
dog i paitim pik
'dog/dogs fight/s pig/pigs'

(63) Basic pidgin with numerals

wanpela dog i paitim tupela pik

'one dog fights two pigs'

(64) Expanded pidgin plural marking

wanpela dog i paitim ol pik

'one dog fights the pigs'

(65) Grammaticalized obligatory plural

wanpela dog i paitim tupela ol pik

'one dog fights two pigs'

(66) Redundant double marking

wanpela dog i paitim ol piks

'one dog fights the pigs'

Example (66) illustrates a developing tendency in Tok Pisin which has previously been observed in creole speakers (Romaine 1992; Smith, G. 2002) and in anglicised varieties (Mühlhäusler 1985a). This is evidenced in my data: the most common pattern when English plural forms are used in Tok Pisin is for double-marking with *ol* pre-nominally even where plural is indicated on the English form with the *-s* suffix, as in (67) or even with a specific plural form of the noun, as in (68).

(67) TiT#10b-07_L420

ol ***problems*** yumi gat

PL problem-PL 2pl.incl have

ol ***problems*** yumi gat

'all (the) problems we have'

(68) TiT#10b-07_L470

ol ***pipol*** blong ol

PL people PREP.Poss 3pl

ol ***people*** bilong ol

'their people'

However, there is some variation in my data, as in (69), where two out of three English NPs are double-marked for plural and one, *pre-emptive measures*, is not.

(69) TiT#10b-07_L299.

ol i got nau wantem ol ***veri: shot tuhm planz*** a lo teigim samala kain ah- ***priemptiv mesuhs*** l ol ***problems*** yumi gat

3pl PM have EMPH PL very short term plan-PL Q.Tag PREP take-TR some kind ah- pre-emptive measure-PL PREP PL problem-PL 1pl.incl have

ol i got nau ol ***very short term plans*** a ? long ***take***-im sampela kain ah- ***pre-emptive measures*** long ol ***problems*** yumi gat

'they have now very short term plans, eh? to take some kind ah-, pre-emptive measures on the problems we have.'

However, the English plural NP in (69) which occurs without the Tok Pisin pluralising particle *ol* is contextually marked as plural through the Tok Pisin quantifier *sampela*. While *ol* may not be entirely obligatory for all speakers in all contexts, contra G.Smith's observations on some first-language children, in the adult bilingual speech of my data double-marking for plural is the norm where English plural forms appear in Tok Pisin. This will be discussed further in Chapter 7 (s7.2.4.1) and the inverse case, where something like the Tok Pisin plural form appears in PNG English, in Chapter 8 (s8.2.1).

G.Smith (2002: 75) predicts that double marking is an indication of competition between plural forms developed from internal resources, the *ol* marker adapted by universal inclinations to utilize pronouns as plural markers, and those borrowed from other languages, the English *-s* suffix. In pidginization, the borrowing of structural morphology is unlikely, and borrowing as a development process is not compatible with the monolingual comprehension of creole innovators. However, such structural borrowing is considered possible in later stages of development, and Thomason & Kaufman (1988) indicate it is a phenomenon resulting from bilingualism.

The development of redundancy and obligatory marking is therefore a different process from variable double marking, though the use of common forms pluralised with an *-s* suffix is apparently a feature of children's Tok Pisin also.

CS models emphasise the role of bilingualism with English in the appearance of double marking, and encourages the re-evaluation of claims of structural borrowing of the English suffix morpheme. Irregularly applied and double marking with both Tok Pisin and English morphemes strongly indicate strong similarities between the use of *-s* suffixation in Tok Pisin and other cases of interaction between language systems.

As I discuss in Chapter 7, the MLF model accounts for the process of the interaction between Tok Pisin and English which introduces the English plural *-s* morpheme into Tok Pisin data.

5.2.5 Conjunctions

Both Tok Pisin and English have conjunctions with similar syntactic roles. Characteristically for a pidgin language, Hall (1943) describes Tok Pisin as having only one 'true conjunction' *na*. However, Mühlhäusler (1985f) gives three: *na, tasol,* and *o*, and Romaine (1988) includes the conjunctive use of *orait*. As Mühlhäusler notes, co-ordination is a development of the expansion of Tok Pisin, and so older speakers (such as those from the period of Hall's study) have a more limited set. Table 5.12 combines Mühlhäusler's 'true' conjunctions with additional items observed by Mihalic (1971: 40-41), Verhaar (1995: 421-449), and G.Smith (2002: 185-190). It details traditional conjunctions and connectors alongside their English counterparts and indicates their role. The top half of the table lists co-ordinating conjunctions. Subordinating conjunctions are listed in the lower half of the table.

Table 5.12 : Tok Pisin Morphology - conjunctions (compiled from Mihalic (1971), Verhaar (1995) and G.Smith (2002))

Conjunction	English equivalent	Role	Source
bat	*but*	Co-ordinating / adversative	G.Smith
en	*and*	Co-ordinating, neutral	Mihalic
long	*than*	Co-ordinating	Mihalic
na	*nor*	Co-ordinating	Mihalic
na(u)	*and*	Co-ordinating, neutral / additive	Mihalic, Verhaar, G.Smith
o / no	*or*	Co-ordinating, alternative	Mihalic, Verhaar
olsem	*and / than / as*	Co-ordinating, modal	Mihalic, Verhaar
olsem bambai	*'so as to', 'in order to'*	Co-ordinating	Mihalic
olsem na	*so*	Co-ordinating, consecutive / causative	Mihalic, Verhaar, G.Smith
tasol	*but / however*	Co-ordinating, contrastive / adversative	Mihalic, Verhaar, G.Smith
bai	*then*	Subordinating, purposive	Verhaar
bihain	*when*	Subordinating / Temporal	G.Smith
bikos	*because* (anglicised)	Subordinating, causal	Verhaar
bilong	*so, then*	Subordinating, purposive	Verhaar
bilong wanem	*why* (interrogative adverb)	Subordinating, Interrogative adverb	Verhaar
long wanem	*because*	Subordinating, causal / causative	Verhaar, G.Smith
husat	*who*	Subordinating, nominal	Verhaar
inap (long taim)	*until*	Subordinating, temporal	Mihalic, Verhaar
inap long	*so*	Subordinating,	Verhaar

Conjunction	English equivalent	Role	Source
		purposive	
long disla taim	*'at that time'*	Subordinating / temporal	G.Smith
maski (wanem / ...o... / ...sapos... / ...tasol...)	*although*	Subordinating, concessive	Verhaar
na	*and* (serializing)	Subordinating, consecutive	Verhaar
nau	*now*	Subordinating / temporal	G.Smith
nogat	*'is not the case'* (conjoins opposites)	Subordinating, negative	G.Smith[48]
nogut	*lest*	Subordinating	Mihalic
olsem	*that, thus, 'as follows'*	Subordinating, nominal	Verhaar
olsem sapos	*as if*	Subordinating	Mihalic
sapos (...orait... / ...em...)	*if / suppose / 'provided that' / whether*	Subordinating, conditional	Mihalic, Verhaar
sapos nogat	*unless*	Subordinating	Mihalic, G.Smith
taim	*when*	Subordinating, temporal	Mihalic, Verhaar
wanem	*which*	Subordinating, nominal	Verhaar

[48] Smith discusses some unusual data on nogat as indicating an unexpected form of an event (Smith, G. 2002: 189-190). See s5.3.6 on negation for non-conjunctive uses of *nogat*.

Table 5.12 makes it amply clear that G.Smith's observation (Smith, G. 2002) that the basic set of three conjunctions has been expanded is true of other datasets also. As well as, uniquely among the sources considered here, finding *bat* 'but' in his data, G.Smith notes in particular the presence of *bikos* as an alternative for causative *long wanem*, and this is also noted as an 'anglicised' alternative by Verhaar also. In fact, *bikos* is included in the Tok Pisin section of Mihalic's dictionary (Mihalic 1971), indicating its use in at least some varieties of Tok Pisin before Independence. My data includes *bikos* as a clear feature of the Tok Pisin there, as in (70):

(70) TiT#10b-05b_L166

mi # ***putin*** se liklik ***bikoz*** ensidisi mas lulu lo disuhla

1sg # put-CONT this little because NCDC must look PREP this

mi, ***putting*** dispela liklik ***because*** NCDC must lukluk long dispela

'I am saying this little thing because NCDC must look into this'

In addition, in the data English *so* is used frequently in Tok Pisin, primarily as a discourse connector, as in (71):

(71) TiT#35b-203_L597

so duhsela em i: wanpela bigpela suhmthing

DISC.CONJ this 3sg PM one-ADJ bi-ADJ thing

so dispela em i wanpela bigpela samting

'so this (is) a big issue'

Mihalic records *en* (see Table 5.12) as a neutral Tok Pisin co-ordinating conjunction in his dictionary, and English *and* is used in the form [æn] frequently in my data, as in (72):

(72) TiT#35b-205_L48

nau disl ol ***sevis*** em i ***rrun daun*** ### ***an*** mi mi gat ***konsuhn*** tru lo displa .

TIME DET PL service 3sg PM run down ### CONJ 1sg 1sg have concern EMP PREP DET(3sg)

now dispela ol ***service*** em i ***run down***, ***and*** mi mi gat ***concern*** tru long dispela

'now this service is run down, and I I have serious concern about it'

Clearly, some 'English' conjunctions have a long history of use in Tok Pisin, and the presence of English conjunctions in Tok Pisin is considered further in Chapters 6, 7, and 9, and considered alongside discourse markers (s5.4.1) in Chapter 10.

5.2.6 Other Affixes in Corpora : expansion vs. code-switching

As with the English plural *-s* suffix, the presence of other English affixes in Tok Pisin corpora such as G.Smith (2002) and Verhaar (1995) can be accounted for as a consequence of code-switching, rather than a development in the Tok Pisin grammar.

While Verhaar's analysis expands the morphology of Tok Pisin to include a significant number of English function (system) morphemes which appear in his corpus, it is not conclusive. The affixes he identifies appear in a modified form suggestive of Tok Pisin pronunciations rather than English ones. Included in such 'new' morphology is a progressive suffix *-in* English *-ing*, past tense *-d* English *-ed*, and agentive *-a* English *-er*.

Verhaar's carefully constructed argument for the productivity of these morphemes in the Tok Pisin grammar depends on two points. Firstly, the validity of his corpus as representative of what he terms a "virtual standard" for Tok Pisin (Verhaar 1995: 3). The written language of the *Buk Baibel*, the Tok Pisin translation of the Bible, and the other sources he includes in his corpus was decided upon by Papua New Guinean Nationals fluent in Tok Pisin, using the orthography developed from Hall's (1943) work by Mihalic (1971) as based on the conservative 'Rural' Tok Pisin of the Madang region.

The inclusion of forms which appear to indicate the useful application of English language functional morphemes within Tok Pisin is therefore, Verhaar concludes, based on the expert knowledge of native speakers referring to what is widely accepted as 'good' Tok Pisin, and is, consequently, indicative of practical language norms.

Essentially, Verhaar's suggestion is that as a variety of English language words which include inflectional morphemes are introduced into Tok Pisin, even monolingual speakers are able to generalise their functions through systematic comparison between different forms, thereby discovering the paradigm and extending it productively in their own use. In this way his argument identifies bilingual borrowing as the nexus for new forms entering the language, and establishes that structural borrowing is an on-going process in Tok Pisin.

Verhaar's inclusion of new morphology presents a significant claim about the nature of the relationship between Tok Pisin and English, and needs to be considered carefully. Accepting that the written corpus he has assembled indeed represents the "virtual standard" for Tok Pisin widely normative in the national community, and agreeing that "virtual standard" includes English inflectional morphology as a productive part of the grammar of Tok Pisin, would indicate that the processes of decreolization are further advanced than has been otherwise anticipated.

Extensive incorporation of English grammatical morphemes into the core code of Tok Pisin removes one of the significant features of the pidgin: the variable use of a single root as noun or verb without a change in form. Borrowing English forms provides a powerful inhibitory force on the further development of indigenous morphology through internal resources (Mühlhäusler 1985d). Where Mühlhäusler noted that German rule and the later social distance between Australians and local people "tended to promote the growth of Tok Pisin from internal resources and the use of loans from local vernaculars" (Mühlhäusler 1985a: 50), and stressed that "Tok Pisin's status as a second language calls for maximal reliance on internal word-formation processes and minimal reliance on borrowing" (Mühlhäusler 1985e: 614), the 'anglicisation' of Tok Pisin has been an on-going process enhanced by the use of English as the medium of school instruction and by the prestige associated with the international language.

However, there are good reasons to remain sceptical about Verhaar's claims for the expansion of the functional grammar of Tok Pisin. Firstly, his claims that his written corpus is representative of a "virtual standard" can be questioned. While the people who produced the texts in the corpus were undoubtedly excellent Tok Pisin speakers, they were also well-skilled in the English language. Throughout the history of Tok Pisin it has been observed that Tok Pisin/English bilinguals' Tok Pisin can have a different flavour to that of other members of the community, notably pidgin speakers without English language competence and first language, creole, speakers. Mühlhäusler, Wurm, and Laycock's 'anglicised' varieties are exactly this.

A code-switching analysis which views such varieties as a consequence of bilingualism with English suggests that the introduced morphology may not be productive (see Chapter 7, s7.7.3.1 on double-marking).

The incorporation of English inflection morphology is unquestionable 'anglicisation' of the language, but is it 'standard'? Rather than being a "virtual standard", I suggest that the language of Verhaar's corpus represents a 'written standard', or even a 'well-lettered written standard'. Verhaar himself does not discuss to what practical extent it is truly indicative of the actual language used by the majority of Tok Pisin speakers. In this light it is important to recognise that while his corpus may represent some kind of 'ideal' or aspirational standard for the language, many of its features may not be commonplace. Indeed, in spoken Tok Pisin one rarely finds such elaborated language as exists in the *Buk Baibel*.

G.Smith's corpus of the creole Tok Pisin of schoolchildren similarly contains such affixes, yet G.Smith himself (Smith, G. 2002) notes that many such affixes remain unanalysed by most speakers, and specifically suggests that nonce borrowing or code-switching may be responsible for the tokens of *-ing* and-*ed* in his data. He uses examples of self-correction to

show that there remains a tension between the English and Tok Pisin grammars in relation to the grammatical role of affixation (Smith, G. 2002: 89-90).

G.Smith's analysis relates directly to Verhaar's second point, regarding the route by which functional morphology can become nativized. Consequently, we find ourselves on ground familiar to theories of contact-induced language change. Verhaar's argument that Tok Pisin speakers are able to analyse and reproduce English morphological paradigms without prior knowledge of English meanings is made questionable by the broad sweep of literature on language mixing, including pidgin and creole genesis, that illustrates that inflectional morphology is rarely borrowed between languages without the assistance of widespread bilingualism with the donating language (Thomason and Kaufman 1988).

Figures from the National Census discussed in Chapter 3 (s3.1.4) show that such bilingualism does exist in Papua New Guinea. However, the weakness of Verhaar's position is that the "virtual standard" of his corpus is established by Tok Pisin/English bilinguals as arbiters of that standard. Once the hypothesis that bilinguals are the gatekeepers of a language is established, that their understandings reflect the standard or emerging standard of a language, the agents of change are being integrated into the model of the outcomes of change. Furthermore, as Mundhenk (1990) observed, many of the decisions made during the creation of the *Buk Baibel* were closely debated by the people involved. It is consequently plausible that from the monolingual speakers' point of view, English morphology is a part of a borrowed word, and is used, unanalysed, with the whole. This is G.Smith's argument, and the examples of double-marking to be presented in Chapter 7 support this.

This is not to say that English inflectional morphology does not form a part of Tok Pisin. Accepting that the language system is being developed by those bilingual with English suggests that 'anglicised' varieties may well be considered the norm. However, this is resisted both within the national community, and by Verhaar himself.

For fluent speakers of an expanded variety the massive incorporation of English words to support, and for some communities supplant, the fundamental vocabulary of Tok Pisin undeniably exposes them to inflectional forms. However, such borrowings appear to be largely unanalysed by monolinguals, and should not therefore be included in an account of the essential grammar of Tok Pisin. In borrowing whole words without analysing the constituent morphemes Tok Pisin speakers may express inflectional forms without utilizing their function. This view is also expressed by Siegel (1998).

Defining Tok Pisin as a single language system is greatly complicated by the behaviour of bilinguals, especially where their behaviour gives example to non-English speakers. Such behaviour is the focus of this study, and is examined in depth in the following chapters. Indeed, as English-medium education influences more and more Tok Pisin speakers, the issue of whether specific morphology is a part of Tok Pisin or not may become much harder to determine: the monolingual speaker may virtually disappear, or it may be the case that all Tok Pisin speakers can comprehend English inflectional morphology, and no longer find it unusual in Tok Pisin. However, distinctions may continue to be drawn by native speakers between what is Tok Pisin and what is English, between 'good' Tok Pisin and 'mixed' Tok Pisin, just as they are by speakers observed throughout the country (see also Chapter 9).

While the status of Tok Pisin as an inherently mixed language, with continued borrowing from English, must be considered, where alternative analyses of English morphology found within Tok Pisin are viable these should be considered before a change to the grammar is said to have taken place.

5.2.7 Summary

This section has illustrated the appearance of English morphemes on loan words in Tok Pisin, and also some examples of Tok Pisin morphology in PNG English in my data. Close similarities in the sound systems of Tok Pisin and English in Papua New Guinea provide one

inroad into questions about the continued mixing of these two languages. A substantial community of self-reported multiliterates with both of these languages (National Statistical Office 1994, 2002a; Paliwala 2011, 2012) exists. Their code-switching behaviour provides opportunities for the appearance, if not the analysis, of distinctive morphology cross-linguistically.

This section has described the key distinctive features of Tok Pisin morphology, and outlined aspects of the differences and similarities between that and English morphology. In addition it has suggested that some models of code-switching may provide an explanation of morphological 'borrowing' alternative to previous accounts of development and expansion of the core Tok Pisin grammar. This explanation will be presented in more detail in Chapter 7.

5.3 Morpho-syntax

Though both Tok Pisin and English share basic SVO word order, it is in the area of morpho-syntax that some of the clearest differences between them have been established. The languages differ in their use of determiners, obligatory in English though not in Tok Pisin, and in the marking of possession. Strategies for negation and the marking of Tense, Mood, and Aspect (TMA)) in the verb phrase (VP) differ markedly between the languages, and some characteristic syntactic structures are not shared. For example: the presence or absence of the copula; the role of dummy auxiliary *do* in English; and the use of the predicate marker *i* in Tok Pisin. Furthermore, Tok Pisin has an extremely limited set of prepositions, especially when compared with English. These differences have been noted in various grammars (Hall 1943; Mihalic 1971; Sankoff, G. and Brown 1976; Woolford 1979a; Mühlhäusler 1985f; Keesing 1988; Romaine 1992; Smith, G. 2002), where a focus has been on the development of syntactic structure in the process of pidginization, expansion, and creolization, and in the context of substrate influences.

In this section, the principal areas of syntax are considered in order to establish benchmark differences between Tok Pisin and English. We have seen above (s5.1) that the phonology of the Papua New Guinea dialect of English is highly congruent with that of Tok Pisin, and that creole Tok Pisin allows for further similarities with English. In the discussion of morphology (s5.2) I have suggested that English morphemes in Tok Pisin corpora, like Tok Pisin morphemes in English language data, may be explained as the result of bilingual code-switching. Significantly, observations on the local PNG English variety have established that standard English benchmarks may not be appropriate to a consideration of the differences between Tok Pisin and English in Papua New Guinea. With reference to syntactic structures too, a comparison of PNG English, expanded and creole Tok Pisin, and the corpus data reveals similarities between them that provide crucial inroads for code-switching behaviour.

Theories about the grammatical constraints on code-switching are primarily concerned with the limitations on, and consequences of, the interaction between different language systems in bilingual behaviour. Just as in the consideration of morphology, above, syntactic structures are presented here to provide a context for my consideration of these models of mixed language in relation to the new dataset in the following three chapters.

5.3.1 Determiners

English noun phrases are regularly marked for number and specificity of referent by determiners, commonly the definite article *the* or *this* or the indefinite article *a*. Such obligatory marking, though it may include a null determiner in some cases, is not a feature of conventional Tok Pisin, which allows optional overt determiners, but has no specific requirement. This contrast can be seen in (73) and (74):

(73) English noun phrase – RED + HOUSE

The red house

This red house

That red house

A red house

DET ADJ N

(74) Tok Pisin noun phrase – RED + HOUSE

redpela haus

ø ADJ N

Without obligatory application of determiners and articles, the Tok Pisin noun phrase is simpler and more flexible than its English counterpart. There is no definite article in Tok Pisin, and though G.Sankoff and Mazzie (1991) have suggested a role for wanpela as an indefinite article in some varieties, evidence has not yet been presented that confirms this role. While it is possible to express some of the same relationships of specificity and number in Tok Pisin as English, as in (75-77), it is not obligatory to do so. As a consequence, it may appear that Tok Pisin and English NPs are well-distinguished by patterns of use of determiners and articles.

(75) definite reference

dispela redpela haus

'this red house'

(76) alternative reference

narapela redpela haus

'another red house'

(77) indefinite article

wanpela redpela haus

'a red house'

However, in PNG English (PNGE) there is evidence that the 'Standard English' rules for regular marking of nouns have been varied. Platt, Weber, and Ho (1984: 52-59) develop an argument with reference to PNG English and other New Englishes that in several of these

varieties speakers are marking a specific/non-specific referent distinction, rather than the standard 'inner circle' English definite/indefinite distinction. This is similar to what happens in Tok Pisin, where examples such as (77) could be interpreted as marking specificity rather than indefiniteness. For PNG English, specific referents may be marked with determiners, but non-specific nouns are generally unmarked. While standard English uses both *a* and *the* for generic reference, speakers of PNG English may neglect these articles for such non-specific nouns. Examples (78-81) are adapted from A.-M.Smith (1978: 18):

(78) (56a) English – GIVE + SPEECH
give a speech

(79) (56b) PNG English – GIVE + SPEECH
give Ø speech

(80) (56c) English – TAUGHT + (Spec.) LANGUAGE
taught the Yabem language

(81) (56d) PNG English – TAUGHT + (Spec.) LANGUAGE
taught Ø Yabem language

For specific nouns, English contrastively uses *a* for indefinite referents, and *the* for definite referents. In PNG English the morphs *a*, *some*, and *one* are all available to mark asserted, indefinite referents, as in (82-83).

(82) English – (indefinite) HOUSE
a house

(83) PNGE – (indefinite) HOUSE
one house

In the respect of using 'one' to mark indefinite referents, PNG English can be said to reflect the Tok Pisin use of the number *wanpela* 'one' as the indefinite article. Compare (77) and (83).

In PNG English, definite specific referents can be marked with either *the* (often pronounced [da]) for culturally shared nouns or *this* (often pronounced [dis]) for shared or confirmed experience, as in (84-85).

(84) English – (definite) HOUSE
the house

(85) PNGE – (definite) HOUSE (confirmed)
this house

Again, while PNG English shares this characteristic in regard to article use with other indigenized Englishes, it also can be seen to resemble Tok Pisin in this area of the grammar. While Tok Pisin has no obligatory placement of the article, the definite article *dispela* 'this' is readily associated by English speakers with the local pronunciation of the English equivalent: *this* [dis], as in (86):

(86) TiT#35b-203_L627
we heding towuhds a: # uuh- uh disastuh in dis kuhntuhrri:
we heading towards a, uh- ***a disaster in this country***
'we (are) heading towards a, uh- a disaster in this country'

Reanalysis of *dispela* as a bimorphemic structure: *dis* + *-pela*, allows for the patterns of appearance of *this* as [dis] as a definite article in PNG English equivalent to Tok Pisin *dispela.*

The flexibility in use of determiners in Papua New Guinea English may be associated with interference from Tok Pisin. In PNG today, Tok Pisin forms a common 'background' language for the variety of English spoken there. A.-M.Smith clearly associates many of the features of PNG English she identifies with imperfect acquisition of English (Smith, A.-M. 1978; 1986: 236), though Platt, Weber and Ho specifically argue against a similar position for well-established communicative varieties (see, for example, on articles (Platt, Weber and Ho 1984: 52)).

For the present study the significant observation is that a traditionally obligatory component of English noun phrases appears variably, or is commonly neglected, in Papua New Guinean speech. As a result we may necessarily consider this a variable aspect of PNG English.

G.Smith discusses his finding that the Tok Pisin indefinite article *wanpela* and definite article *dispela* are not being grammaticalized at the rate predicted in earlier studies (Smith, G. 2002: 147-150). He does suggest that grammaticalized forms of *dispela* as singular and plural markers, as *sla* (contracted from *dispela*) and *slol* (contracted from *dispela ol*), may be developing, but there is no evidence for this in my corpus data.

5.3.2 Possession

In English possession may be indicated by a possessive form of the noun, either a possessive pronoun or through the addition of the suffix -*'s* on the possessor (posr) noun, which precedes the possessee (pose) noun in the form: NP[posr] + -*'s* NP[pose]. In contrast, Tok Pisin does not have any specific possessive pronouns (prenominal or otherwise) or morphology.

English also allows possession to be expressed through constructions containing a post-nominal pronoun, or full DP, with the use of the preposition *of* placed between the pre-posed possessee and the possessor noun, which remains marked with the -*'s* suffix, in the form: NP[pose] *of* NP[posr] + - 's. Canonical Tok Pisin resembles this second construction, and indicates possession through a prepositional construction using *bilong* where the object precedes the preposition, and the possessor follows it. However, there is no possessive pronoun suffix. The Tok Pisin possessive construction is of the form: NP[pose] *bilong* NP[posr], as in (87):

(87) TiT#08-507_L96

yupla tok papa blong en

2pl talk father PREP.POSS 2sg

yupela tok papa bilong en

'you two tell his father'

This is a marked difference between the two language systems, with the construction where the possessor precedes the possessee and the use of the -*'s* suffix being distinctive characteristics of English.

In PNG English, with other New/World Englishes, this distinctive -*'s* suffix may be omitted. While the word order in the Tok Pisin *bilong* construction remains a prominent difference from the English possessive phrase, in the absence of the required possessive pronoun or suffix, its resemblance to the English *of* construction is another example of congruence between these languages in Papua New Guinea.

5.3.3 Copula

The English copula *be* is a requirement linking the subject to the predicate where the predicate is an NP, an AdjP, a PP, or VPs with Progressive aspect or Passive voice. In PNG English, as in many dialects of English, zero copula is more common than in the standard grammar (Smith, A.-M. 1986), and Tok Pisin does not require a copula at all. While some analyses for the particle *i*, such as Mihalic (1971) assigns it this function, subsequent analyses and its highly variable presence in spoken Tok Pisin make clear that it is by no means a comparable element.

As I will discuss in Chapter 8 (s8.2.6), the presence of English copula *be* forms in Tok Pisin in my data is suggestive of a depth of mixing between the two languages that goes beyond the constraints of Alternation or Insertion, and is suggestive of the presence of a highly congruent, mixed grammar, underlying some speakers' Tok Pisin.

5.3.4 DO

In English *do* can function as a main verb and in possessive constructions, as in (88) and (89), and as an auxiliary in interrogatives (s5.3.5) and in negation (s5.3.6):

(88) TiT#08-502_L8
shi wi du it on da dawr
3sg.fem MOD do 3sg.neu PREP DET door
she will do it on the door
'she will do it on the door'

(89) TiT#24b-15_L871
... wi du owa own gadenin
1pl do 1pl.POSS PRON garden-CONT
... we do our own gardening
'we do our own gardening'

The Tok Pisin verb *mekim* 'to make' is the closest analogy to the verbal use, but forms interrogatives and negatives without the use of the auxiliary (see below).

While Platt, Weber and Ho (1984: 76-78) discuss the avoidance of *do* in negative constructions in some New/World Englishes through the use of *never* (as also in some colloquial styles of established inner circle varieties) and high-frequency, unstressed *do* in Indian English, no particular distinctive patterns of use have been described for PNG English.

5.3.5 Interrogatives

In English questions are formed by inverting the subject and predicate, as in (90-91):

(90) English – declarative
you can call

(91) English interrogative
can you call

The particle *do* is used as a dummy auxiliary if no other auxiliary is present, as in (92-93), though (93) does seem slightly forced, and expresses emphasis:

(92)

you do call

(93)

do you call

In addition to the inverted Yes/No interrogative construction, English WH- interrogatives are formed with a similar inversion of word order and the addition of a question word, of the form: Q + auxiliary + N + Verb. Such standard English WH- questions may use the dummy auxiliary *do* in the absence of any other auxiliary, as in (94):

(94) TiT#24b-15_L410

... ah- wha du yu kol dem ?

ah- what AUX 2sg call 3pl

... ah- what do you call them?

'... ah- what do you call them?'

In some New/World Englishes, however, the word-order inversion is not required in interrogatives, and in PNG English question words such as the particle *a* borrowed from Tok Pisin (Platt, Weber and Ho 1984: 127-129) take invariant forms.

Tok Pisin maintains the standard declarative word order and forms interrogatives through rising intonation or adposition of question particles. It does not have an auxiliary function equivalent to *do*. Questions may be formed by postposing a question particle to the predicate. Compare (91) with (95):

(95)

yu kolim em a ?

2sg call-TR 3sg Q.Tag

'can you call him/her?'

In English an emphatic question can be formed by the use of the interrogative pronoun without movement to the beginning of the sentence. In Tok Pisin a question such as (96) does not convey similar emphasis:

(96)

yu kolim em wanem ?
2sg call-TR 3sg what
'you call him/her what?'

Tok Pisin also allows interrogatives to be formed by pre-posing a question word to the subject, but there is no equivalent to the role taken in English by *do*:

(97)

watpo yu kolim em ?
why 2sg call-TR 3sg
'why (did) you call him/her?'

Tok Pisin forms the interrogative through the application of phrasal interrogative adjuncts such as *bilong wanem* 'why' or *olsem wanem* 'how', through independent question words such as *wanem* 'what', *husat* 'who', or *hamas* 'how much', or through the use of question particles such as *a*, *laka*, or *o nogat* following the predicate, as in (98):

(98) TiT#24b-15_L695

yu gat neim bilo yu iet a ?
2sg have name PREP.POSS 2sg EMP Q.Tag
yu gat nem/***name*** bilong yu yet a ?
'you will have received your name eh?'

More recently, question words borrowed from English have been adopted, such as *wai* 'why' and *hau* 'how' (Verhaar 1995), as in (99-100):

(99) TiT#35b-204_L76

wai no ***austreian kamani*** n inuhp inuhp luksave lo ***kwalifuhkaisens uhr eksperiens*** mipela i ga:t

why NEG Australian company NEG able able look.and.know PREP qualifications CONJ experience 1pl.excl PM have

why no ***Australian company*** no inap inap luksave long ***qualifications*** o ***experience*** mipela i gat

'why (is) not any Australian company able to appreciate (the) qualifications we(exclusive) have?'

(100) TiT#10b-07_L165

so ## suhpos ol i ***helthi*** olsem ***wai*** nau ol no laik lain yupla wetwet i stap iet ?

DISC.CONJ # if 3pl PM healthy same Q.Tag EMP 3pl NEG like clan 2pl.excl wait PM CONT EMP

so, sapos ol i ***healthy*** olsem ***why*** nau ol no laik lain yupela wetwet i stap iet ?

'so if they are as healthy why now do they not like your people to keep waiting?'

In addition the English data shows frequent marking of questions with the Tok Pisin question particle *a?*, as in (101):

(101) TiT#24b-15_L1002

dei hev seks olredi a ?

3pl have sex already Q.Tag

they have sex already a ?

'they have already (had)sex eh?'

This shows the impact of Tok Pisin discourse on PNG English.

A significant difference remains between the two languages, however, due to the invariant word order in Tok Pisin, and the use of *do* as a dummy auxiliary in English.

5.3.6 Negation

In English, negative sentences are formed by the negative adverb *not* following the first auxiliary verb in a sentence. If there are no auxiliaries present, then *do* is introduced in auxiliary position. The English *do+not* construction is often contracted to the form [don], as in (102):

(102) TiT#35b-04_L49

anfawtunali # mistuh depu uh- ***aktin spika wi don hav A koruhm hia***

unfortunately, mister deputy- uh- ***acting speaker we don't have a quorum here***

'unfortunately, mister depu- uh- acting speaker we do not have a quorum here'

In Tok Pisin, negation is achieved by pre-posing *no* to the predicate (103), or by replacing the predicate phrase with *nogat* (104):

(103) TiT#08-507_L359

mi no traim[49]

1sg NEG try-TR

mi: no traim

'I did not try'

(104) TiT#35b-203_L512

wanem samting i kamap lo wanem hap lo kuhntri mi ***muvim***[50] ol ***polis faws*** i go ***rraiet skwod*** i go . i nogat ol i kambek sidaun

wanem samting i kamap long wanem hap long kantri mi ***move***im ol p***olis force*** i go ***riot squad*** i go. i nogat ol i kambek sindaun .

---

[49] The absence of an object for this transitive verb is an example of observations by Lattey (1971) and see s5.2.1. It is supported by the example from Mihalic (1971) of directive *Yu traim!* 'Try it, see if you can do it!'.

[50] Interestingly, in this political context, Mihalic (1971) records *muv* as a verb for 'move (a motion), as in political speech'. Here, however, the transitive muvim clearly relates to physical, rather than metaphorical, motion, and represents a loan from English 'move'.

'whatever things happen in whatever part of the country I move the police force there the riot squad there. If nothing (is happening) they return and rest.'

The different strategies are summarised as the use of the auxiliary *do* in English, and the scope of negation in Tok Pisin.

5.3.7 Predicate Marking

One of the most prominent features of Tok Pisin is the presence in many sentences of the particle *i*. This morpheme has been noted in Melanesian Pidgin since its early history and its functional role has been widely discussed in literature on the language. Keesing (1988) describes how the 3rd person singular pronoun *he* from English was reanalysed to fill a role analogous to substrate syntactic structures for subject-referencing, and ultimately developed into a role as a marker for the beginning of the predicate. Mühlhäusler (1990), however, argues that *i* grammaticalized separately in Tok Pisin to other varieties of Melanesian Pidgin English.

The canonical analysis of the role of *i* is that it is a 'predicate marker'. Mihalic (1971) describes it as such, though he also presents it as a form of auxiliary *be*, the copula, and provides prescriptive rules governing its occurrence. However, the fact that some of these rules are only to be applied 'sometimes' belies the fact that the appearance of *i* is highly variable in Tok Pisin, and its role has been much debated. The standards of the *Buk Baibel* that are fundamental to Verhaar's "virtual standard" (1995) experimental grammar follow regular rules, and it is ever-present in formal written Tok Pisin. In spoken language, however, its role and use are more variable.

Referring to a corpus of spoken Tok Pisin collected by Gillian Sankoff, Smeall (1975) and Woolford (1979b) reach different conclusions about the sources of variation: Smeall favours phonological conditions, and Woolford grammatical environment, providing an implicational pattern for *i* use in variable positions. However, though both sets of conditions may be

expected to be influential in the use of *i*, studies of the creole Tok Pisin of young people have raised questions about both analyses. Outside of G.Sankoff's Buang community, the use of *i* has been observed as a declining feature of the Tok Pisin of children, not only by Woolford (1979b), but also by Lynch (1979), by G.Smith (1990), and by Romaine (1993). Much of this data was drawn from particular regions, or derived evaluations of rates of loss from comparisons between the Tok Pisin spoken by locals and the written Tok Pisin of expatriates, and consequently should not be taken as definitive accounts of the reduced occurrence of *i* throughout Tok Pisin as it has been spoken around the country. However, while Harvey (2007) found very low rates of *i* present in her Internet Tok Pisin when compared with letters to *Wantok* newspaper, the observed differences between written/spoken language data suggest that this could be more an indication of a spoken style on the internet, rather than decreolization, as she asserts. The claim that *i* is in decline goes against both the intuitions of Papua New Guinean speakers of Tok Pisin such as Nekitel (1994), and the evidence of G.Smith's corpus (Smith, G. 2002), where it occurs once every 52 words in the Highlands sample, every 49 words in the Momase sample, and every 13 words in the Islands sample.

My data is characterised by mixture of Tok Pisin and English, and a simple count of language tags in the core corpus indicates this mix is roughly 46% English, 44.5% Tok Pisin, with 8% of words indeterminately common to both languages and 1.5% unrecognised. Despite this mixture of languages, in the core corpus *i* is the most frequently-occurring word, appearing once every 31.45 words overall: significantly more frequently than in either the Highlands or Momase samples of children's monolingual data G.Smith describes.

There are clearly some circumstances and some speakers for whom *i* is not used with anything like the regularity suggested by the prescriptions for use by Mihalic and others, and other circumstances where it is used consistently.

Classification of *i* as morpheme whose primary function is the separation of the subject and predicate, a 'predicate marker', may lead to false analyses where it demonstrates a separate function: for example in modifying the meaning of modal *laik* and *ken* (Dutton 1973). The regular use of *i* with aspect markers appearing after the verb (Sankoff, G. and Brown 1976) has become a part of the verb phrase of Tok Pisin, and indicates a new role for the particle in organising the verb phrase, though it is by no means an obligatory component. It has also been observed introducing relative clauses (Smith, G. 2002), and introducing clauses in sequence through the discourse organising constructions *na i* and *nau i* in the Islands region. Consequently I observe that the use of *i* is clearly not in decline, but that its role and function in the grammar of Tok Pisin shows not only regional variation but also high levels of flexibility.

It is possible that *i* is used in some cases for emphasis, serving to focus hearers' attention on particular elements of an utterance, which may range from the predicate, to particular aspectual distinctions, to significant clauses. This role may account for the relatively high frequency of *i* in the mixed Tok Pisin/English data: it is applied to encourage a focus on the Tok Pisin of speakers as Tok Pisin. This would account for its flexibility, and would require a redefinition of this particle as more than just a predicate marker. In fluent rapid speech, *i* may frequently be deleted through lack of stress in streamlined pronunciation.

These observations combined with those on the Tok Pisin verb phrase (s5.3.8) accord with Romaine's discussion on the 'de-grammaticalization' of the predicate marker *i*. Her argument is that observed omission of this morpheme particularly in TMA constructions is a strategy developed by fluent native speakers of Tok Pisin to achieve a "new streamlined but complex verb phrase, which may consist of a main verb modified both pre- and post-verbally by a number of markers of tense, mood, and aspect tightly concatenated and not interrupted by the predicate marker" (Romaine 1992: 282). That such changes may potentially be conditioned by analogy to the English VP is not something she directly addresses, her focus on

creolization encouraging an explanation based on the facility and fluidity of production over one based around structural borrowing. English has no equivalent particle, and so *i* remains a distinctive feature of Tok Pisin verb phrases when compared with that superstrate. In the following section, I consider the similarities and differences between Tok Pisin and English VPs, and specific areas of congruence between them.

Exclusion of the predicate marker makes the speech of the children Romaine (1992) studied much less perceptually transparent to more traditional users of Tok Pisin, but once again this reasserts the significance of the "community of fluent speakers" (Mühlhäusler 1985d) in providing an environment where such innovations in the direction of ease of production as opposed to ease of comprehension are tolerated. Significantly, G.Smith's evidence for regional variation in predicate marking with *i* (Smith, G. 2002: 119) suggests that claims that its use is declining are probably based on unadjusted expectations of 'normal' use. Written norms for Tok Pisin privilege a standard close to the usage of G.Smith's Island schoolchildren: regular marking of the predicate. However, creole speaking children elsewhere in the country do not appear to require *i* for predicate marking.

For native speakers of Urban Tok Pisin, it seems, the clear expression of structured grammatical relations developed by a grammaticalized predicate marker is no longer necessary. Harvey (2007) considers *i* a clitic, and consequently, a late system morpheme whose omission indicates an ML turnover from Tok Pisin to English is taking place. Where my data shows high, though variable, use of *i* in spoken Tok Pisin, it also includes the use of [hi], a form of 1sg *he*, in PNG English constructions in a similar role to that of Tok Pisin *i* (s8.2.6). As a consequence, I consider Tok Pisin and PNG English as potentially influencing each other, with a Congruent Lexicalization model (see Chapter 8) without the stable norms of a mixed language or mono-directional decreolization the most appropriate characterisation of some multilingual behaviour in my data.

5.3.8 The Verb Phrase : Increasing Congruence

This section compares the VP of classical Tok Pisin with those observed in the expanded variety. These are compared with both Standard English forms, and the forms found in my data. Using data from A.-M.Smith's study of PNG English (Smith, A.-M. 1986) it further demonstrates how changes in the Tok Pisin and PNG English verb phrase can be seen as increasing congruence between them.

Whereas English verbs express tense, mood, and aspect through a combination of morphological suffixes and preverbal auxiliaries, the Tok Pisin verb phrase is built up from an invariant verb form and an assortment of pre and post-verbal particles. The basic preverbal markers of tense and mood are derived from contracted forms of English adverbs, or, in the case of *bai* and *save* from older loans which have been grammaticalized into new functions.

The verb phrase in English can be characterised as following an ordering as indicated in Figure 5.3, with modal auxiliaries preceding morphological markers of tense, aspect, and voice and the main verb, along with any associated prepositions:

Figure 5.3 : English Syntax - Verb Phrase Structure

Modal	Perfect	Progressive	Passive	Main Verb	Prepositions	Noun
(one of) *will* *can* *shall* *must* *do* etc.	*have* + -*ed*	*be* + -*ing*	*be* + -*ed*	**V** + Agreement	**P**	**NP**

In contrast, all marking in the Tok Pisin verb phrase is carried out syntactically, with markers following an ordering characterised by G.Sankoff (1991) as in Figure 5.4 and, based on somewhat later data, by Romaine (1992) as in Figure 5.5.

Figure 5.4 : Tok Pisin syntax - Verb Phrase Structure (ref. Sankoff, G 1991 : 64)

no	*bin, ken*	*save*	*laik*	**aux**	**V**	*igo, istap, ikam*	*pinis*

Figure 5.5 : Tok Pisin Syntax - Verb Phrase Structure (ref. Romaine, 1992 : 277)

bai							
mas					**NP**		
ken	*bin*	*save*	*i*	**V**	**PA**	*i stap*	*pinis*
other adverbials		*laik*		[+*im*]	*sori*	*go*	
				[-*im*]	*bikpela*	*kam*	
				gat			

There are differences between these accounts, primarily in G.Sankoff's placement of the negative *no* within the VP and exclusion of the adverbials *bai* and *mas*, In addition, the optionality in G.Sankoff's ordering of the preverbal particles allows for a placement of *ken* (mood: permission/futurity/ability) before *bin* (tense: past), which contradicts both Romaine's data and claims for the universal ordering of creole preverbal markers of tense, mood, and aspect in that sequence: markers of tense preceding markers of modality, preceding markers of aspect (Bickerton 1984). Also, G.Sankoff does not explicitly include a position for *i* in her schema, with this particle appearing as a bound prefix to aspectual *go*, *kam*, and *stap* post verbally. As observed above, the use of *i* as a regular predicate marker may have been overestimated as an obligatory feature of the grammar. However, as a significant feature of the Tok Pisin verb phrase, its presence or absence in particular structures can have special effects in structuring the relationship between Tok Pisin and English. The role of *i* in Tok Pisin is, in some contexts (Mihalic 1971) similar to the copula *be* (s5.3.3). However, it is never inflected, as in English, and, as observed above (s5.3.7), highly variable in the Tok Pisin verb phrase.

It is possible that these differences reflect greater regularisation by the children Romaine worked with. Certainly, Romaine argues for the inclusion of *bai* in the verb phrase, as she

explicitly discusses the development of this particle from an adverbial *baimbai* to a standard verbal marker in advanced first language Tok Pisin. This process had previously been described by G.Sankoff & Laberge (1980 (1973)) and G.Sankoff & Kay (1974) as a feature of creolization, so it is unusual that G.Sankoff's later (1991) work, from which Figure 5.4 is drawn, neglected this feature.

In standard descriptions of Tok Pisin the particle *i* is incorporated into the verb phrase as a predicate marker, with no functional equivalent in English. Similarly, the English copula has no equivalent in Tok Pisin. As a consequence, there are clear differences between the ways in which these two languages code tense mood and aspect on the verb.

In the development of pidgin languages, the introduction of verbal tense, aspect, and modality distinctions is seen to be a feature of the expansion phase, and in early-creolizing pidgins is universally associated with creolization (Labov 1971b: 18). The original future tense adverbial *baimbai* appears pre-verbally in some older texts (Romaine 1992: 275), which suggests that this position was available for tense markers in Tok Pisin before phonological reduction to *bai* or the growth of a community of native speakers occurred.

With the exception of *sa* (reduced from Tok Pisin *save* 'know', derived from Portuguese *saber*), all of the other preverbal markers have been adopted from English, suggesting that, as elsewhere in the language, processes led by bilingualism with English provide distinct but parallel results to language internal processes. Here too is evidence suggesting that the social processes of creolization may be independent from the linguistic developments in Tok Pisin (see s3.1.1).

Table 5.13 reproduces Romaine's list of preverbal markers.

Table 5.13 : Tok Pisin Syntax - preverbal markers in Tok Pisin and their meanings (following Romaine, 1992 : 277)

bai	irrealis/future
mas	obligation
bin	past
ken	permission/futurity/ability
save	habitual
laik	desirative/futurity

In addition to Romaine's preverbal list, durative marking with preverbal *wok long* + V construction is widely attested (Smith, A.-M. 1988b: 119; Smith, G. 2002: 133). Post-verbal markers are: *(i) go, (i) kam, (i) stap* = progressive; and *pinis* = perfective. [51]

In historical Tok Pisin the resources available for expressing complex relations in the verb phrase were limited to adverbial words or phrases. As the community of speakers has evolved, new requirements for fluency and facility in expressing temporal and aspectual relations have developed. As a consequence, some speakers of second language, 'pidgin', Tok Pisin may still rely on more circuitous means to express themselves. In general, however, and as evidenced in the corpus and other studies of contemporary speakers, the structures discussed here are the commonplace standards of communication.

Verbal morphology in Tok Pisin is limited to the transitive suffix *-im* (s5.2.1), and the presence of this morpheme immediately signals a 'pidgin' use of a verb which often is historically derived from English. Many verbs, however, have no difference in form between

[51] An unpublished paper (c.6,500 words) on Tok Pisin Tense Mood and Aspect markers in my corpus discusses in detail their placement and variable appearance. This material is not reproduced here except where strictly relevant to the dissertation topic.

their representations in English and in Tok Pisin. We can compare the differences between English and Tok Pisin TMA expressions through reference to one such 'shared' verb, *go*, or, in the examples below, a verb with different forms in each language *bilas* 'decorate'.

Where English expresses just two true tenses, past and non-past, the perfective and progressive aspects, and a variety of preverbal modal markers, Tok Pisin marks anterior (past) tense, with *bin*, irrealis (future or conditional) mood, with *bai*, and non-punctual progressive aspect with *(i) stap* post-verbally and *wok long* pre-verbally and non-punctual habitual aspect with *save* pre-verbally. With Tok Pisin modals largely borrowed directly from English, it is particularly in the realisation of past tense forms and aspectual distinctions that the languages differ.

Beginning in the past tense the examples (105-107) compare Tok Pisin and English expressions for 2nd person singular *bilas* 'decorate' (past tense). Example (108) illustrates a form commonly found in the corpus:

(105) classical Tok Pisin – 2sg DECORATE (past)
(asde) yu i bilas

(106) expanded Tok Pisin – 2sg DECORATE (past)
yu bin i bilas

(107) English – 2sg DECORATE (past)
you have decorated

(108) core corpus : Tok Pisin/English – 2sg DECORATE (past)
yu (i) bin bilas

Where English has a special past tense form of the verb *decorate* + *-ed* , *decorated*, and includes a past tense form of the auxiliary verb *have*, Tok Pisin marks the past with a single particle *bin* between the subject and the verb, which here, as in the schema of G.Sankoff and

of Romaine (Figure 5.4 and Figure 5.5) is marked with the particle *i*. The English past as 'perfective' aspect is expressed through the use of *have* as an auxiliary verb and a change in the form of the following verb to past participle.

In my core corpus data, where *i* appears with *bin* it does not separate the particle from the verb. Instead the predicate marker precedes the whole predicate: the verb and its associated TMA particles, though it is regularly found attached to the markers *go, kam,* and *stap*. The marker *i* only very rarely follows the 1sg pronoun *mi* or the 1pl.incl *yumi* due to phonological deletion, and, excluding *bin* constructions without those pronouns, and flagged constructions, *bin* constructions without *i* represent a half of all examples in my data: for many contemporary speakers of Tok Pisin *yu bin bilas* 'you have decorated', as indicated as an option in (108) is acceptable.[52] The past in Tok Pisin can be characterised as being expressed through the anterior-tense marking auxiliary *bin* preceding the unchanged verb.

Examples (109-111) compare present tense continuous forms, using the same pronoun verb combination. Example (112) illustrates forms found in the mixed language data:

(109) pidgin Tok Pisin – 2sg DECORATE (present continuous)
(nau) yu i bilas

(110) expanded Tok Pisin – 2sg DECORATE (present continuous)
yu (i) bilas i stap
yu wok long (i) bilas
yu (i) bilas i go

[52] In discourse a complement to the verb may be expected: *yu bin bilas long haus* 'you have decorated (yourself) at the house'.

(111) English – 2sg DECORATE (present continuous)
you are decorating

(112) TiT Corpus : Tok Pisin/English – 2sg DECORATE (present continuous)
yu (i) bilas i stap
yu (i) wok long bilas
yu (i) bilas i go

Again, in English (111) the form of the verb is changed, with the addition of the continuous suffix *-ing*, and a present tense form of the copula *are* precedes the verb. In English, progressive aspect distinguishes present time. The auxiliary verb *be* is placed before the verb, and the verb takes the morphological *-ing* suffix of the present participle form. In classical Tok Pisin (109) there is no change in the verb phrase, just the simple verb marked with *i*. In expanded Tok Pisin (110), as discussed above, there are various resources available to express continuous aspect through adverbs preceding or following the verb root, whose form remains unchanged. Evidence from the corpus suggests that *i* is optional before the predicate, and that where it does appear with *wok long* as a continuous aspect marker, *i* never separates the marker and the verb. However *i* remains frequent with post-verbal *go* and *stap* as adverbial markers of continuous aspect. The present tense in Tok Pisin is versatile, and may be expressed by either the basic verb, or through the addition of a number of auxiliary markers of aspect that are applied post-verbally, and may be marked with *i*.

Examples (113-115) give constructions relating to future time, again using the same pronoun and verb, and again illustrating the form found in the corpus data (116):

(113) pidgin Tok Pisin – 2sg DECORATE (future)
(baimbai) yu i bilas

(114) expanded Tok Pisin – 2sg DECORATE (future)
bai yu i bilas
yu bai (i) bilas

(115) English – 2sg DECORATE (future)

you will decorate

(116) TiT Corpus : Tok Pisin/English – 2sg DECORATE (future)

bai yu (i) bilas

yu bai (i) bilas

English (115) expresses future time through the use of a modal auxiliary verb, commonly *will,* which occurs directly before the basic verb. The transition from classical pidgin to expanded Tok Pisin marking of future events has occurred through the evolution of the marker *bai* from an adverbial phrase.

In contrast with G.Sankoff's account of Tok Pisin (Sankoff, G. 1977; Sankoff, G. and Brown 1980) which includes the regular placement of *i* pre-verbally and a preverbal placement of *bai* outside of the predicate as delimited by *i*, the evidence of the core corpus (116) is split between *bai* as a phrase marker appearing pre-nominally to mark the entire clause (*bai* + N + V) half the time, and *bai* functioning as an auxiliary to the verb (N + *bai* + V) 38% of the time. In my data for *bai*, *i* occurred with *bai* in around 17% of cases and in those cases always came between *bai* and the verb. The marker *i* was c30% more likely to be deleted where it directly followed *bai.*

With the predicate marker appearing so variably, the significant difference between Tok Pisin and English is in the use of the auxiliaries *have* and *be* along with changes in the morphology of the verb in certain aspects. While the use of *i* distinguished the two languages, as the systematic marking of predicates with *i* becomes less common this difference is being eroded. Furthermore, the placement of *wok long* and *bin* preverbal markers following *i* in my data represents a change from the schema offered by Romaine (Figure 5.5), and results in placement of these markers directly preceding the verb they influence. For *bai*, two positions are possible, and in the absence of *i,* the Tok Pisin irrealis (future) construction with *bai* directly preceding the verb resembles the standard English future construction.

The evidence of the corpus indicates that the Tok Pisin preverbal markers can be given a position in relation to the verb that more closely resembles that of English auxiliary verbs: where *i* does not appear, the past and future tensed verb phrases in English and in Tok Pisin have the same word order and are structurally congruent. As neither English nor Tok Pisin requires changes in verb form in statements about the future, in irrealis expressions Tok Pisin *bai* and English *will* are highly congruent.

In PNG English auxiliaries are used more variably than in Standard English. A.-M.Smith (1986) investigates mood and aspect in PNG English in her 5th chapter. She concludes that there is more overt marking of aspect in PNG English than in the varieties of British and Australian English with which she compares it: Papua New Guineans tend to use adverbial expressions to indicate tense and aspect, possibly as a consequence of phonological deletion of the English past perfective suffixes (see above). There is a general preference to use a single modal auxiliary, such as *can* more generally.

Most critical for this comparison of the two language's methods of coding verb groups, however, is the observed tendency for the ellipsis of *have* in the past perfective (Smith, A.-M. 1986: 169). The ellipsis of *have* in PNG English brings it structurally closer to Tok Pisin, which does not feature a copula (s5.3.3).

A.-M.Smith (1988b) discusses the range of aspect marking in PNG English, and includes the use of adverbial time markers to clarify phonologically "unmarked" VPs and a *have to/had to* + V construction for simple aspect as well as the adoption of a standard English *had* + V + *-ed* past perfective construction for simple aspect. In turn, perfective aspect can be expressed with an adverbial marker *already*, as has been observed in Singapore English (Platt and Weber 1980: 81), and with ellipsis of the auxiliary *have*:

(117)

I been to University
1sg be.Past PREP university
'I have been to University'

(118)

This exercise been done
DET exercise be.Past done
'This exercise has been done'

(Smith, A.-M. 1988b: 123)

In the English statement of one MP (119) several of these features combine in the passive voice: the reduced consonant group on the English modal *must* [mʌs] encourages its resemblance to Tok Pisin *mas* [mas] while the adverbial negator *never* is used in the simple tense as observed by A.-M.Smith (1988b: 123) alongside a perfective form of the English verb *be* that is pronounced with a short vowel and so homophonous with the Tok Pisin past tense *bin* [bin]. Together, these markers accompany an uninflected verb *undermine*, pronounced without the final [nd] consonant cluster which would convey the perfect tense in English:

(119) TiT#35b-02_L171
pa:luhmen muˆs neva bin ***unduhmain***
parliament MOD NEG be undermine-0
parliament mas/***must never*** bin/***been undermine(d)***
'parliament must never be undermined'

In about 50% of cases examined by A.-M.Smith (1986: 195) aspect is more overtly marked in PNG English than in standard English. Simultaneously, Tok Pisin is developing a more highly integrated verb phrase, as G.Sankoff and Romaine (above) both discussed.

In the area of modality there are also tendencies in PNG English which increase the resemblance of the local variety to Tok Pisin. Primary observations are the variable use of modal auxiliaries such as *could*, *would*, and *should* without the specificity understood in

standard English, and a consequent tendency to more indefinite modality resembling the Tok Pisin irrealis (Smith, A.-M. 1986: 200). In PNG English, *can...be able to* constructions like those in (120) create an overlap between modals and habitual aspect.

(120)

when I go to school, I can be able to learn English

(Smith, A.-M. 1986)

Even more blurring of the modal/aspectual distinctions of English are made in PNG English when definite mood conveys more about the manner of the action than about its probability (Smith, A.-M. 1986: 205).

As a result, I suggest that some differences between Tok Pisin and English verb phrases are diminished when the local variety of English is taken into account. For PNG English: as with other New/World Englishes certain standard English constructions (such as *be* + V + *-ing* (Platt, Weber and Ho 1984: 72)) are extended to new contexts, and new innovations (such as "dummy auxiliary" *have* or *have to* with a simple verb (Smith, A.-M. 1988b: 122)) similarly have become distinctive features of Papua New Guinea English. However, A.-M.Smith's work clearly reveals a dynamic between innovative and standard constructions, and the tendency in PNG English to simple verb forms, regularised constructions, the ellipsis of *have* in the past participle, and the overt marking of aspect and time increase the congruence between PNG English and Tok Pisin. Such grammatical congruence between the local variety of English and Tok Pisin has influences on how mixed language is considered in Chapters 6 and 7, and is fundamental to the argument of Chapter 8.

5.3.9 Prepositions

The small set of prepositions in Tok Pisin system of prepositions is one substantial feature of the language that contributes to the ease with which it is learned and used. At its simplest, Tok Pisin uses only two preposition forms: *bilong* expressing close relations such as

possessive (see above), and *long* which is used for all other prepositional roles. Both are demonstrated in (121) in reduced pronunciations:

(121) TiT#10b-05b_L223

siuht blong en ***im*** no wasim long ***twenni yi:uhs***

shirt PREP.POSS 3sg 3sg.masc NEG wash-TR PREP twenty year-PL

shirt bilong em ***him*** no wasim long ***twenty years***

'he hasn't washed his shirt in twenty years'

In fact, and contrary to characterisations such as Romaine (2000: 183), there is a third preposition in common use in Tok Pisin: *wantaim* 'with', indicating coincidence and association, as in (122).

(122) TiT#10b-05b_L184

wi wit bi tokim yu lo ***lastwik*** # mi wantaim suhmpuhluh ***frens*** blo mi

1pl 1pl MOD.past talk-TR 2sg PREP last.week # 1sg PREP.Conj some friends PREP.Poss 1sg

we we bin tokim yu long ***last week***, mi wantaim sampela ***friends*** bilong mi

'we we told you about last week, me with some of my friends'

Hall (1943) characterises *wantaim* as adverbial compound which, along with directional adverbs *antap* "on top, above" and *anenit* "underneath, below", and dismissive *maski* "forget-about, ignore", appears in his data frequently in adverbial constructions including *long* as the active preposition. This pattern is illustrated in (123):

(123) wantaim + long

mi wantaim long sampela poroman bilong mi

'me with some of my friends'

Hall (1943: 28) evaluates the use of these adverbs as prepositions as a contraction of constructions with *long*. However, like comparative *olsem* 'the same, just so' in Hall's data, *wantaim* today has a grammaticalized function as a preposition without ever appearing alongside *long*. There are no constructions like (123) in my core corpus, and, as a consequence, *wantaim* is equivalent to English *with.* The *wantaim* + *long* construction appears

to be antiquated, and possible obsolete. As an independent preposition *wantaim* represents a category shared by both Tok Pisin and English.

It is the primary prepositions *long* and *bilong* that most distinguish Tok Pisin and English. The Tok Pisin possessive construction, using *bilong* contrasts with English (s5.3.2 above). However, prepositional phrases in Tok Pisin using *long* are broadly equivalent to those in English. The notable difference here is that *long* as a single form conveys the same relations as a large range of English prepositions (124-131).

(124) English – GO + TO + HOUSE
go to the house

(125) Tok Pisin – GO + TO + HOUSE
go long haus

(126) English – STAY + AT + HOUSE
stay at the house

(127) Tok Pisin – STAY + AT + HOUSE
stap long haus

(128) English – COME + FROM + HOUSE
come back from the house

(129) Tok Pisin – COME + FROM + HOUSE
kam bek long haus

(130) English – SIT + IN + HOUSE
sit in the house

(131) Tok Pisin – SIT + IN + HOUSE
sindaun long haus

In the creolized Tok Pisin and other fluent expanded varieties significant phonological reduction of the prepositions has been observed (see for examples Smith, G. (2002)). The

phonological reduction of these prepositions seems to be the commonplace behaviour in contemporary Tok Pisin, and this is doubtless related to how frequently used they are. In G.Smith's data *long* appears in the top five most frequent words in all three regional corpora, and *bilong* is the eighth most frequently used word overall. More frequently occurring words are principally the pronouns *em*, and *ol*, the conjunction *na*, and the lexical verb and aspect marker *go*. In the mixed language of my own data the prepositions are even more prominent. Appearing in various forms *long* is the most frequent word in the data, and *bilong* is the seventh most frequent. It seems possible that the storytelling register of G.Smith's corpus may be responsible for the promotion of pronouns, conjunctions, and active verbs over these prepositions, while the various sources for my data leave the grammatical morphemes in a more prominent role.

In PNG English, like other New/World Englishes (Platt, Weber and Ho 1984: 82-85), prepositions may be used differently to the usual patterns. It is possible that such variable use reflects the underlying influence of Tok Pisin as a 'background' language. A.-M.Smith (1978: 17) has observed a tendency for the selection of *to* following adjectives, the preposition apparently serving to associate an observer with their observation, as in (132):

(132) PNG English – Adj + TO
it was difficult to me
'it was difficult for me'

(Smith, A.-M. 1978: 17)

With verbs, there appears to be a high level of variability, with particular English prepositions being used in constructions where Standard English would produce a different word (133), or where English verbs do not normally require a preposition (134).

(133) PNG English – V + Prep (alternative)
help someone in singing
'help someone with singing'

(134) PNG English – V + Prep (unnecessary)

to discuss about

'to discuss'

(Smith, A.-M. 1978: 17)

The preposition *at* is used in expressions regarding time and place, and there is variation in the preposition used in temporal expressions (1978: 17).

This variability in the use of prepositions in PNG English can be understood in terms of influence from the patterns of the Tok Pisin grammar: with only three prepositions operative in Tok Pisin, the formal rules regarding the selection of appropriate prepositions in English are not easily acquired by learners of English. A.-M.Smith's studies of the PNG English dialect (Smith, A.-M. 1978, 1986) reveal that prepositions are used quite variably, and that new patterns, such as the observational quality of *to*, are emerging based on generic associations. In PNG English the choice of prepositions is not as exacting as in Standard English, and potentially reveals a partial assimilation of the semantic differences between different English prepositions in the PNG English grammar. What may be termed 'inappropriate' prepositions by speakers of Standard English maintain, under my analysis, expressive qualities, though along lines which more closely resemble the multi-functionality of the Tok Pisin system.

This shift in the use of prepositions in PNG English represents a substantial change in the English code in Papua New Guinea. As a widespread feature of the local dialect, it indicates a reduction in a major functional difference between Tok Pisin and English. Where similar changes have been observed in other varieties of indigenized English similar connections have been drawn with the alternative systems of prepositional marking in the background languages. Here we note the possible relationship between A.-M.Smith's observations, and the weakening of formal distinctions between Tok Pisin and PNG English.

5.4 Lexis

Mühlhäusler (1979a) provides a comprehensive historical account of the development and sources of the Tok Pisin lexicon, and Mihalic (1971) gives us a historical compendium of vocabulary. In his study of schoolchildren's creole G.Smith (2002) also considers the lexis in his data in some detail.

As observed in the discussion of other aspects of the grammars of Tok Pisin and English, above, and in particular the PNG dialect of English, there are many factors which combine to suggest that the boundaries between these languages are much closer than written standards would indicate. Healey was not able to provide an objective answer to the question about how to define a word as being truly Tok Pisin (Healy 1975).

Chapter 7 considers some of the insights offered by code-switching theory, and their value in helping establish the role of new words in Tok Pisin. In the examples and discussion presented here, all words not recorded in Mihalic's Dictionary (Mihalic 1971) are marked as English. While this does not entirely reflect the practical status of cultural loans, for example political terminology, I have adopted this approach because those English words that have become established in the Tok Pisin lexicon since Independence must have been introduced into the language via code-switching. This position follows Myers-Scotton (1993a: 174) and Muysken (2000: 75), who argue that the development of new vocabulary through borrowing requires bilingual behaviour. This is a pragmatic approach, and in the discussion of Chapter 7 and elsewhere the problematic questions around any such decision are specifically raised.

Fundamentally, the question of new lexis borrowed from English is the status of such loans. Are they analysed morphologically? Does their phonological form indicate that they are established, as opposed to nonce loans? Are new words spoken in Tok Pisin that have a source outside the language part of a monolingual lexicon, or is English education vital to the new vocabulary? These are questions addressed in the following chapters.

Here, I will consider characteristically English and Tok Pisin discourse markers found in the mixed language of my core corpus.

5.4.1 Discourse Markers

The discussion of interrogatives, above, has already noted the presence of one Tok Pisin discourse marker in the English of the core corpus, the question tag *a*. In the Tok Pisin data, English markers are likewise extremely common, such as the connectors *so, bikos*, *oke*, *wai*, and *bat*, observed by G.Smith (2002) in his data. In addition to these, I find English form affirmative *yes* as [yɛæ] 'yeah' as in (135):

(135) TiT#10b-05b_L81

olsem isela man i singaut lon pukpuk ia # i kam paulim mi ***yeA oke***

same DET man PM cried PREP crocodile REL # PM come trouble-TR 2sg yes okay

olsem dispela man i singaut long pukpuk ia, i kam paulim mi ***yes okay***

'like this man (who) cried at that crocodile, (which caused) him trouble yes okay'

Other discourse tags from English appear also, like *wel* 'well' and phrases such as *I mean*, as in (136):

(136) TiT#10b-07_L315

wel ah min ol ### olsem mi tok lo wanpla ol ***lisna***

well 1sg mean 3pl # same 1sg talk PREP one-Adj PL listener

well I mean ol, olsem mi tok long wanpela ol ***listener***

'well I mean they, as I said to one (of our) listeners'

There is also, especially in the Radio Talkback samples, a tendency towards English-form greetings like *hello good morning* and *good bye*, as well as *thank you* as [θænk ju] in place of the more traditional Tok Pisin *moning tru*, *babai nau*, and [tenk ju].

In my data, features like the use of *so* and *I mean* to structure the discourse strongly resemble accounts of the use of *orait* and *tasol* in older corpora. G.Smith (2002) suggests the presence

of *so* as a reduced form of Tok Pisin *tasol*, but in the mixed language of the core corpus congruence between such a reduced form and the English tag reinforces it appearance.

Just as the use of *a* or the emphatic *em nau* in PNG English distinguishes a speaker by indexing a National (Gewertz and Errington 1999) identity, the presence of these features is a marker of an anglicised discourse. Their presence in G.Smith's schoolchildren's creole is an indication of their widespread currency in contexts where English has a presence, and they can most easily be used to recognise 'anglicised' speech.

In Chapter 6 I discuss the presence of discourse tags from a second language as a characteristic of mixed language behaviour.

5.5 Conclusions

The analysis presented in this chapter has shown that the Papua New Guinea dialect of English is a variety that shows significant similarities in phonology to Tok Pisin. This chapter has shown what some might interpret as morphological borrowing. However, Chapter 7 will show that this might not be the case, based on analyses using morphosyntactic models of code-switching. It has also been shown that the streamlined verb phrase of the creole variety of Tok Pisin has permitted convergence with the verb phrase of English. Furthermore, variability in the use of prepositions in PNG English may encourage an increased sense among speakers of its similarity to Tok Pisin.

This chapter has shown both areas of congruence between Tok Pisin and English, and also areas of difference. In the three chapters to follow, the mixed language of the core corpus will be considered in depth from the point of view of code-switching theory to describe and assess the extent to which different types of language mixing are taking place between English and Tok Pisin, and to consider what the possible consequences and impacts of such mixing might be.

6 Findings 2: Alternation between Languages

Alternation is in many ways true 'code switching', as opposed to 'code-switching', in the sense that when speakers alternate languages in speech they are effectively switching each language code on and off while expressing themselves. As discussed in Chapter 3 (s3.2.1), Insertion (the focus of Chapter 7) has been labelled 'code-mixing' (Kachru 1978; Bokamba 1988) because elements from one language are mixed into another. However, following Myers-Scotton's (1993a) concept of a matrix language, in Insertion it is not the actual language codes that are mixed, but the content of each language mixed together within a grammatical frame structured according to the rules of just one code. A true 'mixed-code' underlies the third type of code-switching, Congruent Lexicalization (Muysken 1997, 2000) (see Chapter 8, below).

Alternation between languages is a form of code-switching that can most obviously be understood as demonstrating the use of two languages. Languages can be switched between speaker turns, between sentences within a single turn, or within sentences.

What characterises 'Alternation' as a type of code-switching is an observable A-B structure, where a division between parts of speech in each language can clearly be drawn (Muysken 2000: 7). As discussed in Chapter 3, in studies of multilingual communities (Gumperz 1971a; Reyes 1976; Gal 1979; Poplack 1980, 1981) an awareness of the presence and role of multilingual communication developed, and with it lines of enquiry into the structural constraints on language mixing within the sentence. Inter-sentential, or inter-CP, switching can have significance at the level of discourse in terms of identity and status negotiations, and it is in this light that the examples of Alternation presented below are considered.

Poplack's proposed Free-Morpheme and Equivalence constraints, further developed with D.Sankoff as a structural account of code-switching (Sankoff, D. and Poplack 1981),

represent a germinal attempt to provide an account of the restrictions on where switching can take place.

> **The free morpheme constraint**: a switch may not occur between a bound morpheme and a lexical form unless the latter has been phonologically integrated into the language of the bound morpheme.
>
> (Sankoff, D. and Poplack 1981) (emphasis in original)

Though Tok Pisin and English are both languages with very limited morphology, the validity of the Free Morpheme constraint has been questioned in a variety of works, and not least by Myers-Scotton (1993a) whose Matrix Language Frame (MLF) model removes the requirement of phonological integration through her account of Insertional mixing. The next chapter (Chapter 7) presents numerous examples from the corpus that also call this into question.

> **The equivalence constraint**: the order of sentence constituents immediately adjacent to and on both sides of the switch point must be grammatical with respect to both languages involved simultaneously. This requires some specification: the local co-grammaticality or equivalence of the two languages in the vicinity of the switch holds as long as the order of any two sentence elements, one before and one after the switch point, is not excluded in either language.
>
> (Sankoff, D. and Poplack 1981) (emphasis in original)

Again, observations of different types of mixing have shown that the Equivalence Constraint is not universal, and indeed D.Sankoff and Poplack allow that:

> [T]he validity of any code-switching constraint, including the free morpheme and equivalence constraints, depends strongly on the particular configuration of social factors obtaining in a given community.
>
> (Sankoff, D. and Poplack 1981: 8).

Nonetheless, these two constraints remain fundamental tools in an approach to language Alternation within the sentence.

As I will explore in Chapter 7, additional models have been developed to account for language mixing where elements from one language appear bracketed by or inserted within elements from another, A-B-A 'Insertion' under Muysken's (2000) typology. The consideration of such unbalanced switching along with its relationship to linguistic borrowing was something that Poplack specifically excluded from her model. As is clear from the many

examples in the corpus, both types of switching are commonplace between Tok Pisin and English, indicating a range of relationships between these languages and their speakers in the community.

I consider the Alternation type of switching associated with Poplack (1980) against a social backdrop of diglossic relations that vary from context to context (see s3.2.4.1, s3.2.5.1), and in the light of the varying associations of Tok Pisin and English as 'we' and 'they' codes. In additional to such situational switching, pragmatic and metaphoric linguistic functions of Alternation observed elsewhere (Gumperz 1971b) are illustrated with examples from the corpus.

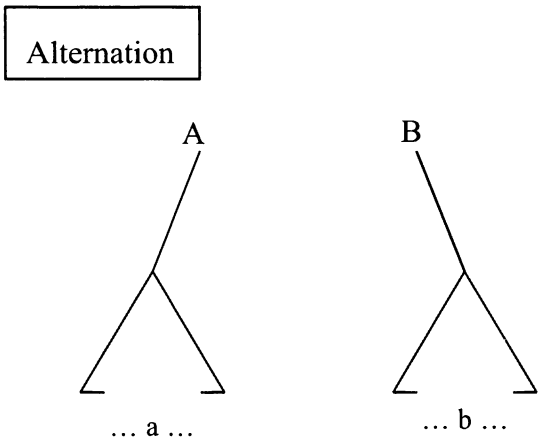
This chapter explores the presence of language switching between Tok Pisin and English as it follows patterns of the Alternational type, and demonstrates the range of social and pragmatic functions such behaviour can fulfil in speech.

First, examples from my data are presented to illustrate what is meant by this kind of code-switching in the community (s6.1). Then decisions about language choice and the significant markedness of such choices are considered, alongside a reassessment of Fergusson's concept of 'High' and 'Low' codes (s6.2). Different types of functions achieved by Alternation, following Gumperz and Auer's Pragmatic and Metaphorical distinctions (see s3.3.1 and Table 3.1), are illustrated that demonstrate that Alternation between Tok Pisin and English in Papua New Guinea is used meaningfully (s6.3). This chapter concludes with some indications of the significance of Alternation in the language community (s6.4).

6.1 Examples of Alternation

Muysken (2000) illustrates the abstract structure of language Alternation with the diagram reproduced here as Figure 6.1:

Figure 6.1 : Alternation - Structure (ref Muysken 2000: 7)



from Musyken (2000: 7) where a constituent from language *A* (with words from the same language) is followed by a constituent from language *B* (with words from the same language). The language of the constituent dominating *A* and *B* is unspecified.

From Figure 6.1 we can see that Alternation within the sentence takes place where constituents can be seen to be separate, indicating that there are clearly two systems in use.

Within the corpus there are 215 examples of switching between languages following the Alternation pattern. Most, 137, of these are switches by a single speaker within their own turn and between sentences, as in (137) where the speaker follows a Tok Pisin sentence with another in English.

(137) TiT#24b-15_L491

bai ol i stap paitim mipela # wol- aw ol ***spesal*** kain brum mipela wokim lo ples. ***so dis spesal diz brumz ah- # dey olredi ah- # madjik voises bin spowken tu it*** a?

FUT 3pl PM CONT hit-TR 1pl.excl # CONJ.PL CONJ PL special type broom 1pl.excl work-TR PREP village. DISC.CONJ DET special DET.PL broom-PL ah- # 3pl already ah- # magic voice-PL Cop.PAST spoke-PAST Prep 3sg Q.Tag

bai ol i stap paitim mipela, wantaim ol o ol special kain brum mipela wokim long ples. ***so this special these special brooms/***brum-***s*** ah- ***, they already*** ah- ***, magic voices been/***bin ***spoken to it*** a ?

'They put (them) up there and then they will be hitting us(excl), wol- our special kind of brooms we(excl) make in the village. So these special these brooms ah-, they already ah- magic voices (have) been spoken to it eh?'

This inter-sentential Alternation is not syntactically interesting, though, as I will illustrate below, it does have a range of expressive and pragmatic functions. However, the omission of the auxiliary verb *have* expected in English *magic voices have been spoken to it* is a feature of

PNG English that is congruent with Tok Pisin grammar (see (117)-(118) and discussion in s5.3.8), and of significance for Chapter 8.

There are also 47 language switches between speaker turns, where one speaker chooses a different language within the discourse, as in (138):

(138) TiT#35b-05_L19

[simultaneous speech]
MPK: ***bikos*** o:l i: longlong i geda longlo
SPK: ***dets h- uhnparluhmentuhri langwidj***
CONJ PL PM mad PM all mad
3sg.NEUT.COP h- NEG-parliament-ADJ language
because ol i longlong i olgeta longlong
that's h- ***unparliamentary language***
'Because they are crazy all of them are crazy'
'That is unparliamentary language'

Again, as discussed below (s6.3), the markedness of language choice can serve a range of functions in discourse.

The remaining 21 examples of Alternation show A-B patterns within the sentence, such as those below:

(139) TiT#35b-203_L620

na pos ol i kalap kalap nabaut # ***mista asisen spika: aim sori we wi heding tu # ai don no***

CONJ if 2pl PM jump jump around # mister assistant speaker 1sg.COP sorry where 1pl travel-CONT PREP # 1sg AUX.NEG know

na sapos ol i kalap kalap nabaut, ***mister assistant speaker I'm sorry where we heading to, I don't know***

'and if they carry on recklessly, mister assistant speaker I am sorry where we (are) heading to I do not know'

(140) TiT#08-507_L19

yupla save olsem wanem mipela ***gat ada things tu du***

2pl know same what 1pl.excl have other things PREP do

yupela save olsem wanem mipela gat/***got other things to do***

'you(pl) know that we(excl) have got other things to do'

(141) TiT#24b-15_L32

an in de vili ol katim skin blo meri

CONJ PREP DET village PL cut-TR skin PREP.Poss woman

and in the village ol katim skin bilong meri

'and in the village they cut the skin of the women'

(142) TiT#35b-203_L28

it doznt mattuh wat wi hav hiuh biznis o wateva buht law an awda em i ***praioriti*** .

3sg AUX.Neg matter 3sg 1pl have here business CONJ whatever CONJ law CONJ order 3sg PM priority

it doesn't matter what we have here business or whatever but law and order em i ***priority***

'it doesn't matter what we have here, business or whatever, but law and order is a priority'

It is precisely the grammatical rules governing such intra-sentential switching that Poplack's Equivalence Constraint and Free-Morpheme constraints are proposed to describe, and it is in order to deal with such Alternation within the sentence that Myers-Scotton and Jake (1997) revised the Insertional MLF model to apply to switching within the CP.

This chapter considers the explanations and interpretations of inter-sentential language change and intra-sentential A-B structures like (139)-(142) as examples of Alternational code-switching.

In the next section (s6.2) intra-sentential Alternation is also considered within the context of its markedness, and the communicative functions it serves.

6.2 Code Choice and Markedness

In PNG, the Language Triangle (

Figure 2.1) illustrates the matrix of languages in the community as a whole. Around Independence, it was maintained that the points of this triangle in PNG (English, Tok Pisin, and Indigenous Vernaculars (*tok ples*)) existed in diglossic relations (Adler 1975; Bickerton 1975a; Mühlhäusler 1985a; Piau and Holzknecht 1985; Wurm 1985). As discussed above (s3.1.6), the presence of a distinctive PNG dialect of English was not considered in linguist's accounts until later, when it came to be considered as an indigenized variety of English.

Today, most people attest to a shifting and blurring of boundaries between the languages in this triangle. In many village communities, Tok Ples vernaculars are giving way to Tok Pisin as children educated outside of the village context acquire the national lingua franca as their 'we' code (Sankoff, G. 1980d; Kulick 1992; Romaine 1992; Foley 1997; Smith, G. 2002). In the urban contexts, and in the peri-urban contexts of villages where Tok Pisin use has become the norm, under the influence of English language education and widespread currency, English has become, for some speakers at least, the 'We' code, though for the majority, Tok Pisin is the language of the community and English the language of wider communication. Many of my speakers attested to a similar pattern: in the village they would use *tok ples* or *tok pisin* (Tok Pisin), and in the town Tok Pisin or *tok inglis* (English language).

Consequently, language choices outside of particular expectations are marked. The associations of language with particular aspects of Indigenous, National, and International identity can be indexed and played with through language choices that convey different roles and responsibilities in particular contexts. In a market in Port Moresby the following exchange between two nationals took place, with the market seller's response indexing her bilingual status while also applying a specificity to her question not so easily expressed in TP.

(143) QL = informant; MKT = market seller

QL: ol kostim hamas?
MKT: ***wich wan***?
3pl cost-PL how.much
DET one
QL: ol kostim hamas ?
MKT: ***which one ?***
'how much (do) they cost?'
'which one?'

The choice of Tok Pisin in the urban context may have been intended to index associations of shared nationality, however the response in English indicated that an identity as English-speakers was more appropriate in this context: Port Moresby, in the Papuan region, has a history of English and Hiri Motu, and Tok Pisin is a more recent introduction. The exchange in (143), therefore, includes a renegotiation of the 'we' and 'they' code dynamic. This is a common use of Alternation both to index an identity, and also a marked choice to add emphasis to her request for greater specificity. In Madang a similar exchange took place between a customer entering a supermarket and a security guard:

(144) SHP = shopper; SGD = security guard

SHP: no ken kisim basket bilong mi!
SGD: ***its mai djob***
NEG MOD.permission take-TR basket PREP.POSS 1sg
3sg.NEU.COP 1sg.POSS job
SHP: no ken kisim ***basket*** bilong mi!
SGD: ***it's my job***
'(you) cannot touch my basket!'
'it is my job'

Again the switch to English is marked, but this exchange also indicates an escalation of claims to authority through language choice: while in the rural context the customer's use of a strong Tok Pisin directive would have indexed values of authority and Government, in the urban context of a supermarket in the state capital the guard's marked choice of English supersedes that authority. In this way, different values can be assigned and assumed for

different languages in different contexts, eroding strict patterns of diglossia. The concept of Markedness is Myers-Scotton's (1993b) alternative to strict assumptions of diglossia.

Her Markedness Model uses the idea that members of a community have observed normal behaviour and internalised a set of principles that allow them to tell when someone is acting in an unusual way. This 'Markedness Evaluator' comes into play in multilingual communities when people switch between languages. Assuming that for most common situations there is a normal 'unmarked' language to use, the claim of the Markedness Model is that we can negotiate our identity in an interaction, in terms of our 'rights and obligations' as a participant by making 'marked' choices of language that will be noticed by others as 'meaning something'.

6.2.1 Code Choice in the Houses of Parliament

In the corpus, such marked choices are particularly clear in the Houses of Parliament. Officially trilingual, the National Assembly is, as Nekitel has observed, "de facto bilingual" (Nekitel 1998: 154) due to the fact that Hiri Motu is rarely used on the Floor of the House without some objection being raised by MPs from outside the Papuan region. The Speaker of the House generally uses English to address the MPs, but MPs may themselves use English, Tok Pisin, or a combination of the two.

Political speeches in general, and in the House in particular, have frequently been referred to as representing some of the most 'anglicised' Tok Pisin (Mühlhäusler 1975, 1985d, 1985b; Piau and Holzknecht 1985; Wurm 1985; Dorney 1990; Romaine 1992; Siegel 1997; Smith, G. 2000, 2002). The extent to which cultural borrowing from English has impacted on the political register is discussed in Chapter 7. However, within the House, as Nekitel discussed in his chapter on the Trilingual Parliament, Alternation between the languages is also a feature that impacts on the ability of Papua New Guineans to follow the debates there.

Significantly here, we can see how language choice becomes a feature of political debate within the House, showing the same features of markedness and indexicality as observed in the examples at (143) and (144), above.

In the following example from a National Parliament session the Speaker switches to Tok Pisin to attract the attention of an MP who is busy in some argument.

(145) TiT#35b-04_L69

SPK: ***honawabl membuh faw xxx pli:s bi sited***
MP: XXX .
SPK: ***ring thuh bel clak ### honawabl membuh faw xxx pli:s***
MP: XXX.
SPK: ***bel*** i ***ring***im nau yu sidaun ***honawabl membuh faw xxx ## pli:s bi sited.***

honour-ADJ member PREP xxx please COP seat-PAST

ring DET bell clerk ### honour-ADJ member PREP xxx please

bell PM ring-TR now 2sg sit honour-ADJ member PREP xxx ## please COP seat-PAST

SPK: ***honourable member for*** xxx ***please be seated***

MP: (unclear)

SPK: ***ring the bell clark , honourable member for*** xxx ***please***

MP: (unclear)

SPK: bell i ringim nau yu sindaun ***honourable member for*** xxx , ***please be seated***

'Honourable Member for xxx please be seated'
'(unclear)'
'Ring the bell clerk, Honourable Member for xxx please'
'(unclear)'
'bell is ringing, now you sit down Honourable Member for xxx please be seated'

The Speaker of the House normally uses English throughout a session, but used Tok Pisin twice to address his speech to this Member, who had ignored earlier comments directed to him in English. By making a choice 'marked' by its difference from his normal usage, the speaker achieved a specific communicative goal through language switching.

A little later in the same debate the Speaker of the House insists upon English, using the Official language in the House to gain control over a Tok Pisin exchange that is taking place between two MPs:

(146) TiT#35b-04_L101

MP1: yupela mas i larim mani blo yu i go i stap we !

MP2: go sidaun!

SPK: ***honruhbuhl membuh fo xxx # an honruhbl membuh fo xxx # ai hav not ruld # yo poin ov oda***

2pl MOD.obligation PM allow money PREP.POSS 2sg PM go PM stay where

go sit

honour-ADJ member PREP xxx, CONJ honour-ADJ member PREP xxx 1sg COP NEG rule-PAST 2pl.POSS point PREP order

MP1: yupela mas i larim mani bilong yu i go i stap we !

MP2: go sindaun !

SPK: ***honourable member for*** xxx , ***and honourable member for*** xxx . ***I have not ruled , your point of order***

'You must reveal where your money went to!'
'Sit down!'
'Honourable member for xxx, and Honourable Member for xxx, I have not ruled, your point of order'

Elsewhere language choice, and the issues of we/they code indexes and Markedness are made explicit. In (147), which includes the exchange in example (139), the speaker is only able to gain control over an MP by speaking in Tok Pisin.

(147) TiT#35b-05_L19

1. SPK: ***onrabuhl memba fo xxx # withdraw dhat wuhd***
2. MP : # noet #
3. SPK: (unclear)
4. MP: ***bikos*** ol i longlong i geda longlo
5. SPK: ***dets*** h- ***uhnpalimentuhri langwidj yu hav tu*** uhh-
6. MP: # biko disela parlimen i longlong (thumps table) na mi gi mipela i stap lo hia !
7. SPK: mi kolim pinis ah- ***koruhm*** na yu xx (unclear)
8. MP: orait na mi se yu tok
9. yu tok yumi no mi rausim
10. em inap lo tokim yumi rausim
11. yu tok yumi no sanap mi nau mi rausim # mi kamautim dispel samting !
12. SPK: ***honruhble memba fo xxx*** em i kam lo maus blo yu ya
13. autim displa # tok nogut ya
14. MP: ***ekskyuz mi suh***
15. ***thank*** yu tru ***suh***
16. SPK: uh- ***gud***

1. SPK ***honourabl member for*** xxx , ***withdraw that word***
2. MP nogat
3.
4. ***because*** ol i longlong i olgeta longlong
5. SPK ***that's*** h- ***unparliamentary language you have to*** uhh-
6. MP ***because*** dispela ***parliament*** i longlong na mi givim mipela i stap long hia !
7. mi kolim pinis ah- ***quorum*** na yu xxx
8. MP orait nau mi save yu tok
9. yu tok yumi na mi rausim
10. em inap long tokim yumi rausim
11. yu tok yumi no sanap mi nau mi rausim , mi kamautim dispela samting !
12. SPK ***honourable member for*** xxx em i kam long maus bilong yu ya
13. autim dispela , tok nogut ia
14. MP ***excuse me sir***
15. ***thank*** you/yu tru ***sir***
16. SPK uh- ***good***

1. *'Honourable Member for xxx, withdraw that word'*
2. *'sorry'*
3. *'...'*
4. *'Because they are crazy all of them are crazy'*
5. *'That is unparliamentary language You have to uhh'*
6. *'Because this Parliament is crazy (thumps table) and I represent my people here!'*
7. *'I have to finish calling quorum and you (unclear)'*
8. *'Alright now I hear you talking*
9. *You talk our talk I will not get angry*
10. *Here it is with words that we get angry*
11. *You said we could not stand up and it got me angry, I came out with this kind thing (speech) !'*
12. *'Honourable Member for xxx it came from your mouth.*
13. *Get rid of this, bad kind of talk.'*
14. *'Excuse me sir'*
15. *'Thank you very much sir'*
16. *'uh- okay'*

Flagging his choice with the non-lexical cue *uhh* the Speaker does not complete his directive in English, but instead alternates languages between lines 5 and 7, using Tok Pisin to make his point clear. This is explicitly acknowledged by the MP at line 8. When, at line 13, the speaker repeats in Tok Pisin his initial, English, direction to withdraw the use of the word *longlong* 'crazy', he finally gets an apology. The apologies at lines 14 and 15 are also marked, with the MP alternating to English to make his acquiescence clear. The Speaker's hesitation

on the last line indicates that this final switch to English was especially marked, and indeed: this MP very rarely uses English in the House.

These examples illustrate a consistent pattern of language use in the National Parliament where Alternation between Tok Pisin and English is both marked for expressive content and index identities associated with a 'Low' and a 'High' code. Negotiations such as in (147) dramatically illustrate the on-going tension between Tok Pisin as 'the language of the people', and English as a 'language of authority' whose use can be opaque to many Papua New Guineans.

In texts taken from Question Time debates language choice in question and answer sessions can be even more marked. Often comments begun in Tok Pisin will be concluded with a final statement in English.

In (148) this Alternation occurs with a repetition, and the Minister concluded his answer from this point in English.

(148) TiT#07a-Hd01_L64

bai mi mikim o e wanem suhmting # uh- em askim ***ail du evrithing pasuhbl tu asist him bikoz xxx iz wan of awa*** ## ah- ***les develop eria*** ### .

FUT 1sg make-TR CONJ 3sg what thing, uh- 3sg ask-TR 1sg.MOD.FUT do everything possible PREP assist 3sg.MASC CONJ xxx COP one PREP 1pl.POSS ## ah- less develop-0 area

bai mi mekim o em wanem samting, uh- em askim ***I'll do everything possible to assist him because*** xxx ***is one of our***, ah- ***less/least develop(-ed) area***

'I will make whatever uh- he asks I will do everything possible to assist him because xxx is one of our ah- least develop(ed) area(s)'

In (149) the same pattern is repeated, with Alternation to English being used in political rhetoric to reinforce the MP's final point.

(149) TiT#07a-Hd01_L76

inap lo askim ***Oseid*** # ***Oseid*** long ***fund***im ***upgreidin*** # i go lo ol # ol disela ***epots*** # osem ah- xxx # ah- chuh- # xxx # xxx # ah- # xxx # xxx # xxx # na xxx # . ***awl dis hav veri veri impoten epots if*** # ***dowz epots bi teiken as*** # ***uh matuh of*** # ***neshnuhl sekuriti*** # ***thenk yu***

able PREP ask-TR Ausaid # Ausaid PREP fund-TR upgrade-CONT # PM go PREP PL # PL DET airport-PL # same ah- xxx # ah- chuh- # xxx # xxx # ah- # xxx # xxx # xxx # na xxx # all DET have very very important airport-PL PREP # DET airport-PL COP take-PAST CONJ DET matter PREP # national security # thank 2sg

inap long askim ***Ausaid***, ***Ausaid*** long ***fund***im ***upgrading***, i go long ol, ol dispela ***airports***, olsem ah- xxx, ah- chuh-, xxx, xxx, ah-, xxx, xxx, xxx, na xxx . ***all this/these have very very important airports if, those airports be taken as, a matter of, national security, thank you***

'(are) able to ask Ausaid Ausaid to fund (the) upgrading of all these airports such as ah- xxx, ah- chuh-, xxx, xxx, xxx, xxx, xxx, and xxx
all these have very important airports if, those airports be taken as, a matter of, national security, thank you'

In an extension of this stylistic technique, elsewhere an MP chooses to address the Speaker in Tok Pisin to excuse his question, and then frames that question in English. Again, this is an example of Alternation indexing the identities associated with the two languages, but more importantly, utilising the code-switch metaphorically to indicate the separation between the MP's aside to the Speaker of the House and his Question to the Minister for Justice:

(150) TiT#07a-Hd03_L36

missa spiga # kwestuhn blo mi no i stap long ah- # ***nowtis pepa*** mi askim ah- # ***juhstis ministuh*** . cough- # ***in duh lait of*** uh- # ***thuh law an awduh situeishin ### dhi inkrising law an awduh situeishin in dhuh kantri #*** ah- ...

mister speaker # question PREP.POSS 1sg NEG PM stay PREP ah- # notice paper 1sg ask-TR ah- # justice minister cough- # PREP DET light PREP uh- # DET law CONJ order situation ### DET increase-CONT law CONJ order situation PREP DET country # ah-

mister speaker, ***question*** bilong mi no i stap long ah-, ***notice paper*** mi askim ah-, ***justice minister***. cough-, ***In the light of*** uh- ***the law and order situation, the increasing law and order situation in the county***, ah- ...

'Mister Speaker, my question is not on ah- (the) notice paper I (am) ask(ing) ah- (the) Justice Minister (coughs), in the light of uh-, the law and order situation, the increasing law and order situation in the country, ah- ...'

Example (150) is drawn from a long sample of speech that also illustrates the significance of language choice in the multilingual setting of the Parliament. When the Minister responds to this question, he does so in English, but a supplementary question from the Deputy Leader of the Opposition is addressed to him in Tok Pisin. Holding his language, the Minister answers

the supplementary question in English, and this leads another MP to ask the same question again in Tok Pisin, forcing the Speaker of the House to state, in his customary English, that "the minister has already made it clear" in response.

These exchanges illustrate the tension between languages and issues of language choice within the National Parliament. The Minister made a marked choice not to switch languages to follow the language of the supplementary question, and another Opposition MP repeating the question highlighted this. Marked Alternation is a clear feature of political discourse.

6.2.2 Discourse-patterning: Tag Switching and Alternation

In other registers, Alternation is not always so strongly marked. There are many examples of tag switching in the data where the discourse is punctuated with switched sentence-peripheral elements. Muysken (2000: 26) identifies peripherality as a key feature facilitating code-switching, and indicates that it is suggestive of Alternation (Muysken 2000: 100). This kind of code-switching behaviour accounts for the presence of discourse tags from English in Tok Pisin, and vice versa.

This is illustrated by the following examples from conversation (151) and radio talkback (152-153):

(151) TiT#24b-15_L700

as a simbol a blong en nau # ***oke*** bai yu:sim disela

PREP DET symbol Q.TAG PREP.POSS 3sg.NEUT now # okay FUT use-TR 3sg.NEUT

as a symbol a bilong en nau, ***okay*** bai yusim dispela

'as a symbol eh? Of them, okay (they) will use that'

(152) TiT#10b-07_L83

wel a min ah- # Jo athin i suhpos i lukluk tu

DISC 1sg mean ah- # Jo I.think PM perhaps PM look PREP

well I mean ah-, Jo ating i sapos i lukluk tu

'well I mean ah- , Jo I think if you look into it , '

(153) TiT#10b-07_L144

wel ai min ahh- yu # suhpos yu yu lu lukluk i go insaid moa # yu ken lukim olsem i gat sampela lain moa lukluk lon weit yet lo kisim ***dei aut*** blon ol

DISC 1sg mean ahh- 2sg # suppose 2sg 2sg lu- look PM go inside more # 2sg MOD.ability look-TR same PM have some-ADJ group more look PREP wait REL PREP take-TR day out PREP.POSS 3pl

well I mean ahh- ***you***/yu, sapos yu yu lu- lukluk i go insaid moa, yu ken lukim olsem i gat sampela lain moa lukluk long weit iet long kisim ***day out*** bilong ol

'well I mean ahh- you, if you you loo- look deeper into this, you can see that there are some other people still waiting to get their day out.'

There are 136 examples of tag switching marked in the core corpus. English speech is marked with high levels of use of the Tok Pisin interrogative tag *a* (47 times). In Tok Pisin contexts, 89 English tags appear, including the examples (151-153) above. Of the remainder, 18 are discourse specific phrases like *mister speaker* in the National Parliament debates, and there are a few occasional tags such as *well* and *okay*. In general, the casual Tok Pisin collected in the core corpus features most frequently the English discourse tags *yeah*, 26 times, and *so*, 23 times. In this sample of Papua New Guinean speech it is clear that discourse tags are used interchangeably in both languages (s5.4.1).

Significantly, G.Smith (2000) considers *so* to be a new Tok Pisin conjunction due to the frequency with which it appears throughout his corpus of children's speech, though Verhaar (1995) does not include it in his grammar based on the written language he surveyed. This difference can best be understood through the lens of code-switching theory, where *so* crosses from English into Tok Pisin discourse as a tag through Alternation at the periphery of the sentence. The presence or absence of *so* in a speaker's Tok Pisin may also be indexical of the degree of anglicisation of their speech. The same argument pertains to the other conjunctions observed cross-linguistically in Chapter 5 (s5.2.5).

Tag switching is an excellent way for a speaker to index a multilingual identity, and is a point of differentiation between different varieties of the language. In PNG English, Tok Pisin tags

such as *em now* to show agreement, or *a* to pose a question clearly index a National identity. In Tok Pisin, using *so* instead of the more conventional *olsem na*, *tasol* or *orait* to structure one's rhetoric indicates modernity, and it is found throughout the urban language of the corpus. As switched tags become commonplace in a register, their markedness within a specific discourse decreases. Chapter 9 presents some data that supports this. For outsiders, however, they remain clear indicators of mixed language.

The prominence of English discourse markers in the Tok Pisin data (see also s5.4.1), taken alongside the presence of English conjunctions (s5.2.5 and 7.3.4.3.1) is particularly interesting in the light of Auer's (1999) typology, as discussed in Chapter 3 (s3.2.5 and s3.2.7). While the regularities of Auer's FLs ('fused lects') may not be in evidence in data, both the Tok Pisin and the English of many speakers is marked with conjunctions and discourse markers from the other language, which elsewhere characterise 'mixed codes' (Salmons 1990; Maschler 1998; Oesch Serra 1998).

6.2.3 A Place for Un-Marked Switching

With its theory of rights and obligations and social negotiation, the Markedness Model places a lot of emphasis on language switching as a tool, and, for many situations, underestimates the customary nature of switching. This is why Myers-Scotton deliberately discusses the notion of "code-switching as the unmarked choice" in her work (Myers-Scotton 1993b). The idea here is that the mixing of languages can become so much the norm that specific switches no longer attract attention. This kind of unmarked switching is a characteristic of a particular kind of multilingual community, and indicates a special relationship between the mixed languages. Most significantly, it is the actual use of the mix itself, rather than particular instances of mixing, which carries a social message. For most people who mix freely and frequently like this, the mixed variety has strong in-group associations.

G.Sankoff specifically identifies "two kinds of tendencies" in Buang: clear Alternational switching and also the Insertion of Tok Pisin into Buang, "marking" it as part of multilingual discourse. Vitally, for G.Sankoff, this multilingualism is the important meaning of mixed language, not the individual acts of CS:

> That the speaker uses both languages in this context is important – he keeps in touch with all segments of his audience by constant Alternation. But exactly what he says in one language or the other is not in itself important[.]
>
> (Sankoff, G. 1980a: 44)

Working later, and elsewhere in PNG, Kulick and Stroud have argued that specific switches are marked, and locally meaningful, against the background of societal phenomena (Kulick and Stroud 1990; Stroud 1998). Between Tok Pisin and *tokples* CS can be both marked and un-marked. Between Tok Pisin and English, especially where the identity-associations of each language are variable (see s2.3, above, Romaine (1992: 54), and the market examples (139) and (140) in s6.2) acts of CS can likewise be marked with added meaning or un-marked, without 'local meaning' in Auer's (1999) terminology.

The reality of markedness in language choice relies on the experience of switching between languages as an extraordinary process. However, as we experience constantly in multilingual communities, the use of different languages to communicate is, for some people, so commonplace in certain situations that actually sticking to a single language throughout a conversation may itself be the extraordinary act. This is the case in some situations in Buang, as highlighted by G.Sankoff (above), and is equally the case in many 'Urban', modernised, contexts in PNG today, where Tok Pisin and English are used together without comment. In this sense we can talk about 'CS as the unmarked code' in a sense explored by Carol Myers-Scotton (1993a, 1993b). That is to say: sometimes language mixing is so normal that we do not need to investigate the meanings behind particular switches, though the reason why the whole conversation is mixed in this way may be very much worth our interest.

Unmarked switching is common among groups of Tok Pisin/English bilinguals, and free and frequent mixing is theoretically the most direct route to language change. Examples of this will be looked at closely in Chapters 7 and 8. In general, however, Alternation switches are not only marked, but also expressively used for pragmatic or metaphorical impact. This is shown by the discussion of examples in the following section.

6.3 Alternation: Expressive Functions of CS

Fine-grained definitions of types of CS such as those suggested above (see s3.2 and s6) help us to understand how Alternation comes to represent a different set of linguistic techniques to Insertion, and how deliberate mixing also achieves some very different social and expressive effects to borrowing and Insertion. Where the discussion of Chapter 7 focuses on 'anglicised' Tok Pisin, where English elements are found in Tok Pisin speech, Alternation between languages requires a command of both language codes, and is only the province of speakers with some multilingual capacity.

This section discusses my findings on the sociolinguistic and expressive functions of Alternation between Tok Pisin and English in the corpus.

6.3.1 Pragmatic Functions

The corpus contains examples of code-switching serving a range of pragmatic functions in communication (Gumperz 1982: 75-81; Auer 1984). These include switching to reiterate a particular statement, to direct a particular utterance to a particular addressee, switching to make an interjection, or to provide focus for interrogation, and to provide message qualification: to provide focus or emphasis on a declaration, explanation, or on new information.

6.3.1.1 Reiteration

One of the most fundamental pragmatic uses of Alternation is switching languages for repetition. Such reiteration is one of the most common types of code-switching. Remarks can be either clarified or strengthened through emphasis by repeating them in a second language.

One example of this behaviour was (148), above. Example (154) is another instance of Alternation from Tok Pisin to English for this purpose:

(154) TiT#24b-15_L783

sapose yu: yusim blo narrapla man ***wen yu: yu:s aduh aduh klans*** uh- # ***dressing***

suppose 2sg use-TR PREP.POSS other-ADJ man ADV 2sg 2sg use other other clan-PL uh- # dressing

sapos yu yusim bilong narapela man ***when you use other other clan's*** uh-, ***dressing***

'if you use what belongs to other men when you use other other clan's uh-, dress-style'

Commonly, code-switching for repetition is also used to make clear the translation of a particular word in each language, as below where a code-switch to Tok Pisin glosses the English word *tree* as *diwai* or *kokonas*:

(155) TiT#24b-15_L695

ai got mai own trri # osem diwai na osem kokonas a ?

1sg have 1sg.POSS own.ADJ tree # same tree CONJ same coconut Q.TAG

I got my own tree, olsem diwai na olsem kokonas a ?

'I (have) got my own tree, like tree and like coconut eh?'

Such switches perform an explanatory function and can work to expand the vocabulary of listeners, or to introduce a borrowed word or phrase to an audience unfamiliar with the donor language:

(156) TiT#35b-202_L262

buhlong wanem yumi kolin ***budo planing*** ?

butum a planin

statim lo has na i go antap lo hed blong en

PREP.POSS why 1pl.INCL call-TR bottom.up planning
bottom up planning
start-TR PREP bottom CONJ PM go above PREP head PREP.POSS 3sg.NEUT

bilong wanem yumi kolim ***bottom-up planning*** ?
bottom up planning
statim long as na i go antap long hed bilong en

'why do we(incl) call (it) budo planning?
Bottom up planning.
Starts at the bottom and travels to its head'

A further excellent example of repetition to explain the meaning of an introduced English word is (274) in Chapter 8, below.

6.3.1.2 Addressee/Topic

Pragmatically also, switching languages can be used to perform the function of addressee-focus. This capacity to target a particular addressee through language choice is a feature unique to multilingual communication, and is a valuable part of a bilingual repertoire. The change in language marks the utterance as an aside, and does not necessarily reflect the default language of the addressee. Frequently, in conversation, a marked language switch can indicate or direct our reference: changing languages to address someone different to the person we are speaking to, or when interrupting, emphasises the change in addressee.

In Port Moresby urban bilingual Tok Pisin/English households, it is common for younger children to be addressed in Tok Pisin, while older children will be addressed in English once they begin to learn it in school. An example of addressee-switching is (157), where the mother switches to Tok Pisin to direct her comments to a baby:

(157) TiT#08-508_L29

KLT: ***hiz wet***
MNM: mmm- ***his wet***
MNM: yu ia wet tumas ia!
3sg.MASC.COP wet
mmm- 3sf.MASC.COP wet
2sg REL wet much REL
KLT: ***he's wet***

MNM: mmm- ***he's wet***

MNM: yu hia wet tumas hia!

'He is wet'
'mmm- he is wet'
'you there, (you are) too wet!'

Such addressee-switching was also a feature of the National Parliament language choices of the Speaker of the House in examples (145-147) above.

6.3.1.3 Interjections and Interrogation

A marked switch between languages can be used to interrupt a speaker with a comment (interjection), or with a question (interrogation). In (158), English is used to mark an interjection, a switch that is echoed by the first speaker before he continues in Tok Pisin:

(158) TiT#24b-15_L471

MSM: wanpla kain ***liv*** ol i kolim salak
sos nao i pulim lo skin blo yu lon -

TMS: ***veri itchi***

MSM: ***veri itchi***
ol i putim antap lo hap ia

one-ADJ kind leaf 3pl PM call-TR salak
suppose now PM pull-TR PREP skin PREP.POSS 2sg PREP

very itchy

very itchy
3pl PM put-TR above PREP part EMP

MSM: wanpela kain ***leaf*** ol i kolim Salak
sapos nau ol i pulim long skin bilong yu long

TMS: ***very itchy***

MSM: ***very itchy***
ol i putim antap long hap hia

'one kind of leaf they call Salak
if they go and put it on your skin nuh-'
'very itchy'
'very itchy'
'they put (it) onto there (your skin)'

The English phrase is marked both by lexical choice and word order: *very itchy* contrasts with conservative Tok Pisin *skirap tru*. The markedness of the switch gives impact to the interruption.

6.3.1.4 Qualification

Examples of marked switching providing focus and emphasis for particular comments in political rhetoric have already been considered, where alternating to English at the end of a Tok Pisin speech appears as a common technique in cases such as (149). This marked switching can also be used to qualify smaller statements, as in (159):

(159) TiT#35b-203_L512

i nogat ol i kambek sidaun ***bat dez no propa plening***

PM NEG 3pl PM come.back sit CONJ 3pl.NEUT.COP NEG proper planning

i nogat ol i kambek sindaun ***but there's no proper planning***

'nothing they come back and debate but there is no proper planning'

This is similar to the Demonstrative, Declarative, and other switches considered below.

6.3.1.5 Demonstrative/Explicative

Marked switching can be used expressively in a range of similar pragmatic functions, as with Qualification above. Particularly frequent in the conversational texts in the corpus is switching to provide explanations, or to demonstrate a point.

In (160) this pattern appears twice over. A mother's question in Tok Pisin (line 1) is followed up with a statement in English (line 2). The teenager's response is in Tok Pisin (line 4) and the mother switches back (line 5) to rephrase the question. At line 6 she repeats her earlier (line 2) English statement in Tok Pisin, and the switches to English to provide a qualifying explanation:

(160) TiT#08-208_L11

1. MNM: eh yupla go mekim wanem ay ay
2. ***no playin*** a ***its madi:***
3. YNT: ***hey ay***
4. KLT: ol lain lo ...
5. MNM: yupela lai go lo we
6. no ken go pilai plis ***its madi: and drrain an ... wi don wan duhti fit runing insaid***

1. Ah- 2pl go make-TR what DISC DISC
2. NEG playing Q.Tag 3sg.NEUT.COP muddy
3. DISC DISC
4. PL group PREP
5. 2pl PROX go PREP where
6. NEG MOD.permission go play please 3sg.NEUT.MOD muddy CONJ rain CONJ 1pl DAUX.NEG want dirty feet run-CONT inside

1. eh- yupela go mekim wanem hey hey!
2. ***no playing*** a ***it's muddy***
3. hey hey !
4. ol lain long ...
5. yupela laik go long we ?
6. no ken go pilai plis/***please it's muddy and rain and ... we don't want dirty feet running inside***

1. *'eh what are you doing? ay ay!'*
2. *'no playing eh? it is muddy'*
3. *'hey ay'*
4. *'the other people from ...'*
5. *'where do you think you are going?'*
6. *'do not go and play please it is muddy and raining and ... we do not want dirty feet running inside'*

Alternating languages here helps the mother get her point across, and by accommodating to the teenager's use of Tok Pisin at line 4, she is also able to soften the severity that her English direction conveys. Such bilingual interaction is both fluent and versatile.

In (161) the speaker provides a qualified explanation of the Tok Pisin phrase *man bilong wok* within a switch to English:

(161) TiT#24b-15_L846

so taim yu kam autsaid em # taim yu kam ausaid em # yu: # yu man blo wok a yu save? ***bikos yu no*** ah- # ***wi*** luh -u # ***Papua Niu Guinians*** a # ***yuno wi liv on had woking*** a ?

DISC ADV 2sg come outside 3sg.MASC # ADV 2sg come outside 3sg.MASC # 2sg # 2sg man PREP.POSS work Q.tag 2sg know? CONJ 2sg NEG ah- # 1pl PREP u- # Papua New Guinean-PL Q.Tag # DISC 1pl live PREP hard work-CONT Q.Tag

so taim yu kam ausaid/***outside*** em, taim yu kam ausaid em, yu, yu man bilong wok a yu save? ***because you know*** ah-, ***we*** long u-, ***Papua New Guineans*** a, ***you know we live on hard working*** a ?

'so when you come outside it, when you come outside it, you, you are a man of work eh? You understand? Because you know ah-, we in u-, Papua New Guineans eh? you know we live on hard work eh ?'

The significance of the title *man bilong wok* for an initiated male is conveyed in English, adding gravitas through the marked language switch.

The same speaker uses the same technique with the language sequence inverted: providing Tok Pisin to explain the significance of an English statement in (162).

(162) TiT#24b-15_L669

Mai mai mai fadas fadha # hi yuzd dat # hi yuzd dat dresin
So yu mas yusim isla ***drres***

1sg.POSS 1sg.POSS 1sg.POSS father.POSS father # 3sg.MASC use-PAST DET # 3sg.MASC use-PAST DET dressing
DISC 2sg MOD.obligation use-TR DET dress

my my my father's father, he used that, he used that dressing
So you/yu ***must***/mas yusim dispela ***dress***

'my my my father's father, he used that, he used that dressing
so you must use that dress'

This flexibility demonstrates that it is the marked switch itself, rather than the status of a particular language, that adds gravitas to the explanation.

6.3.1.6 Declarative

Similarly using code-switching to add emphasis to their words, there are examples of politicians using Alternation to mark declarative statements. At (163) the switch to English at this MPs second line declares the significance of the limited experience he confesses at first (line 1), and the switch back to Tok Pisin at line 3 allows him to repeat and refocus his statement.

(163) TiT#35b-203_L445

1. Mi traim lo liklik ***fuhnding*** blo mi long ah- ***Niu Ailan Provzuhl Gavman*** # i kam lo li:r taso i no inap halavim
2. ***so ai ken onli speek # asisten spika: # on iu ailan siyuaishin***
3. ***so*** lo ol narapluh ***provins*** mi no sawe

1. 1sg try-TR PREP little funding PREP.POSS 1sg PREP ah- New Ireland Provincial Government # PM come PREP here only PM NEG able help-TR
2. DISC 1sg MOD.ability only speak # assistant speaker # PREP New Ireland Situation
3. DISC PREP PL other-ADJ province 1sg NEG know

1. mi traim long liklik ***funding*** bilong mi long ah- Niu/***New*** Ailan/***Ireland Provincial Government***, i kam long hia tasol i no inap halavim
2. ***so I can***/ken ***only speak, assistant speaker, on*** Niu/***New*** Ailan/***Ireland situation***
3. ***so*** long ol narapela ***province*** mi no save

1. *'I tried for a little of my funding for ah- New Ireland Provincial Government, I came here after so I cannot really help*
2. *So I can only speak, assistant Speaker, on New Ireland Situation*
3. *So on the other provinces I do not know'*

Even more striking are switches such as (164), where Alternation serves to strongly mark the point of order as distinct from his personal comments:

(164) TiT#35b-04_L28

atin mi ***hop*** osem biain sampluh taim mila mas # ***palamen*** mas la- larim na em mas kam ***laif o divi*** # lo sampla taim # ***buht A dis poin in taim mista aktin spika # wi don haf kworuhm # uhnfawtyunali***

I.think 1sg hope same FUT some-ADJ time 1pl.EXCL MOD.obligation # parliament MOD.obligation la- allow CONJ 3sg MOD.obligation come live PREP television # PREP some-ADJ time # CONJ PREP DET point PREP time mister acting speaker # 1pl AUX have quorum # unfortunate-ADJ

ating mi ***hope*** olsem bihain sampela taim mipela mas, ***parliament*** mas la- larim na em mas kam ***live on TV***, long sampela time, ***but at this point in time mister acting speaker, we don't have quorum, unfortunately***

'I think I hope that in the future we(incl) must, Parliament must la- allow and it must come live on television, at some time
But at this point in time, mister Acting Speaker, we do not have quorum, unfortunately'

Language switching adds weight and impact to his statement, the juxtaposition of the two languages indexing their associations of inclusiveness and formality, and providing a rhetorical flourish to his point of order in the House.

6.3.1.7 Focus

Similarly, language switching can be used to mark a particular statement as the main focus of attention. In another example from political rhetoric (which includes the example of typical

A-B Alternation given at (142)), the switch to English at line 2 emphasises the significance of the point.

(165) TiT#35b-203_L28

1. Olsem olgeta toktok yumi sa harim # athing i ***beis*** long***law an awde situeisen*** lo disela kantri #
2. ***ah- it duhzn matta wat wi hev hia biznes o wateva: buht law en oda*** # em i ***prioriti***

1. same all talk 1pl.INCL HAB hear-TR # I.think PM base-0 PREP law CONJ order situation PREP DET country
2. ah- 3sg AUX.NEG matter what 1pl have here business CONJ whatever CONJ law CONJ order # 3sg PM priority

1. olsem olgeta toktok yumi save harim, ating i ***base(d)*** long ***law and order situation*** long dispela kantri
2. ah- ***it doesn't matter what we have here business or whatever but law and order***, em i priority

1. *'all of the speeches we(incl) have been hearing, I think are based on (the) law and order situation of this country*
2. *ah- It does not matter what we have here business or whatever but law and order, it is(the) priority'*

The main Alternation here is between the sentences on line 1 and line 2. The statement at 2 gains focus from the language switch. A second switch within the conjoined clause *but law and order em i priority* in line 2 again draws focus from the markedness of Alternation. This switch in line 2 is flagged by a short pause, adding weight to the kernel proposition of his statements. The presence of the English elements *based, law and order situation*, and *priority* will be considered in Chapter 7, as Insertions characterising an anglicised register, and they are also prominently marked through repetition in this speech. It is the Alternational switching here, however, which carries the force of the MPs statement, focusing the listener's attention on his assertion.

Similarly, in a conversational text, particular focus is given to a verb phrase through Alternation, this time in the direction of Tok Pisin:

(166) TiT#08-506_L255

an wen wi keim daun tu da skul wni wi lukluk i go home

CONJ when 1pl came down PREP DET school 1pl look PM go home

and when we came down to the school we we lukluk i go home

'and when we came down to the school we we looked (for him) (he) had gone home'

The switch itself adds focus, through the markedness of the language change.

6.3.1.8 New Information

Alternating languages so actively involves both the speaker and the listener in bilingual behaviour that it is almost always marked. The focus marked switching draws is extremely flexible. While we have already seen how it can be used to draw attention to a core argument, or to a strong statement, it can also be used to flag the introduction of a new topic or significant new information.

In (167) Alternation is flagged by non-lexical vocalisations as the topic of conversation is changed from humorous domestic topics to a more serious concern, the Alternation to English marking the change in topic:

(167) TiT#08-507_L64

... ol laik passim diki antap

a- aah- bifaw yu klos ta wiping mi gat ***aiv got wan kweshuhn***

3pl PROX pass-TR Diki above
a- aah- before 2sg close PREP weep-CONT 1sg have 1sg.POSS have one question
... ol laik pasim Diki antap
a- aah- bifo/***before you close to weeping*** mi gat ***I've got one question***

'... they were about to pass Diki up (to there).
a- aah- Before you (are) close to weeping you there I have got one question'

Here too we see an unconscious switch back to Tok Pisin resisted by the speaker, who self-corrects *mi gat* to English *I've got* in order to maintain the new language selected to introduce this new information.

Elsewhere, in a discussion of traditional practises, the main speaker elaborates on the Tok Pisin statement *bai ol i complain yu*, using Alternation to English to emphasise the new information: details of the likely consequences of a complaint against you (168):

(168) TiT#24b-15_L717

MSM:	ba ol i ***komplen*** yu:
AAA:	yes
MSM:	***an wan sambodi wil bi seyn oh- dets not youhs*** a ?
	FUT 3pl PM complain 2sg
	Yes
	CONJ one somebody MOD.fut COP say-CONT oh- 3sg.NEUT.COP NEG 2sg.POSS Q.Tag
MSM:	bai ol i ***complain*** yu
AAA:	*yes*
MSM:	***and one somebody will be saying oh- that's not yours*** a?
	'they will complain (about) you'
	'yes'
	'somebody will be saying: oh let us get yours eh?'

In (169) new information is again marked by language Alternation, in this case the means by which the traditional method of sterilisation during the actions of the Tok Pisin verbs *wasim* and *klinim* takes place:

(169) TiT#24b-15_L323

em bai wasim han en bai holim em bai klinim a
so wi yus a: # uh- # ***tradisonal medisin*** # ah- ***its*** uh- ***meid of*** # ah- uh- # ***de spesal tanget lif***

3sg FUT wash-TR hand 3sg.NEUT FUT hold-TR 3sg FUT clean-TR Q.Tag DISC 1pl use DET # uh- # traditional medicine # ah- 3sg.NEUT.COP uh- made PREP # ah- uh- # DET special Tanget leaf

em bai wasim han en bai holim em bai klinim a
so we use a, ah-, ***traditional medicine***, ah- ***it's*** *ah-* ***made of***, ah- uh-, ***the special tanget leaf***

'will wash (the) hand they will touch (with) they will clean (them) eh?
'so we use a , ah- , traditional medicine , ah0 its ah- made of , ah- uh- , the special Tanget leaf'

6.3.2 Metaphorical Switching

The examples of Alternation above all use the expressive markedness of code-switching for pragmatic purposes. Some, such as reiteration, are primarily explicative. Others, like addressee-focus, are situational. In some cases the switch itself, regardless of direction, achieves the expressive aim. Other uses enhance communication by combining the impact of language switching with values associated with language choice. By and large these functions work because the switch is both marked and takes place between full sentences in each language.

Essentially extending the pragmatic uses illustrated above, metaphorical switching (Blom and Gumperz 1972; Auer 1984) is used to describe more abstract linguistic functions, where the markedness of the switch is used for rhetorical and semantic flourish.

6.3.2.1 Quotation

In an extension of Alternation for addressee-focus discussed above, switching languages can be used in bilingual discourse to emphasise a second 'voice' for the speaker. This is considered 'metaphorical' switching by Blom & Gumperz and by Auer. In the corpus some Alternational code-switches mark quotations. In (170) from a conversational text, and (171) from political speeches, explicit quotations are given in English.

(170) TiT#08-506_L225
okei mipla kam daun lo # em tok ***okei yu suhvaiv den***
DISC 1pl.excl come down PREP # 3sg say okay 2sg survive ADV
okay mipela kam daun long, em tok ***okay you survive then***
'Ok we (excl) came down to (them), they said "ok, you survive then"'

(171) TiT#35b-203_L632

minista blo ***komunicaisin*** i tok # ***wi siting on uh taim bom*** a ?

minister PREP.POSS communication PM talk # 1pl sit-CONT PREP DET time bomb Q.Tag

minister/minista bilong ***communication*** i tok, ***we sitting on a time bomb*** a?

'minister of communication said, we (are) sitting on a time bomb eh?'

Significantly, the quotations do not have to be in the same language as the original statement. Alternation as a technique gives additional clarity to the act of quotation.

6.3.2.2 Topic-Comment

The use of two languages allows for an additional rhetorical technique, and across the texts in the corpus are examples of metaphoric elaboration of an argument through Alternation. Switching languages allows for a contrast to be established between the topic of a statement and the speaker's comment upon it. This is not necessarily strictly equivalent to a topic-comment structure as expressed in e.g. Sinitic languages.

In (172) the switch from Tok Pisin to English in the conjoined clause shows rhetorical contrast between the desire of the speaker (as topic) and someone else's action (introduced as comment).

(172) TiT#08-506_L193

mi laik ***rast***im ol na ***da fada bi:t mi tu it***

1sg PROX arrest-TR 3pl CONJ DET father beat 1sg PREP 3sg.NEUT

mi laik ***arrest***-im ol na ***the father beat me to it***

'I was going to arrest them but the father beat me to it'

The use of clear A-B Alternation to highlight the separation of topic and comment in the second part of the earlier example (142) and (165) is presented here as (173):

(173) TiT#35b203_L28

... ***buht law en awda*** # em i ***prioriti***

CONJ law CONJ order # 3sg PM priority

... ***but law and order***, em i ***priority***

'... but law and order is a priority'

Isolating this subordinate clause, the Alternation from English to Tok Pisin observed earlier serves to focus attention on the comment *em i priority*. This subordinate clause becomes pivotal as a topic-comment on the whole sentence.

In (174) the prepositional phrase *lo lukluk long ol i sanap wantaim* provides a reason for the action of the predicate *overhaul*. Appearing towards the end of a long speech on the need to re-structure the police and army, and following criticisms of the current resourcing and training provisions, this Alternation to Tok Pisin provides a final comment on the scope of review proposed.

(174) TiT#35b-203_L315

its abaut taim wi ovuhhol # duh hol ah- ***ami # da howl polis faws*** lo lukluk long ol i sanap wantem

3sg.NEUT.COP about time 1pl overhaul # DET whole ah- army # DET whole police force PREP look PREP 3pl PM stand CONJ

it's about time we overhaul, the whole ah- ***army, the whole police force*** long lukluk long ol i sanap wantaim

'it is about time we overhaul, the whole ah- army, the whole police force to look at everything involved in it'

Similarly in (175) (also discussed at (139) as showing typical A-B Alternation structure), the main clause topic introduced in Tok Pisin, regarding the irresponsibility of the administration, is commented on in the PNG English dependent clause, with an omitted copula:

(175) TiT#35b-203_L620

na pos ol i kalap kalap nabaut ***#mista asisen spika: aim sori we wi heding tu # ai don no***

CONJ suppose 3pl PM jump jump about # mister assistant speaker 1sg.COP sorry where 1pl go-CONT Ø PREP # 1sg AUX.NEG know

na sapos ol i kalap kalap nabaut, ***mister assistant speaker I'm sorry where we heading to, I don't know***

'and if they carry on recklessly, mister assistant speaker I am sorry where we (are) heading to I do not know'

This switch is doubly motivated by the focus on the Speaker, with an established formal default language of English, as addressee. However, as has been shown in other examples here, there does not appear to be any community norm regarding language choice when addressing the Speaker of the House.

These kinds of switching are rhetorically very strong as they occur within larger sentences, and have significant impact. Alternation to add force to topic/comment structures is a valuable feature of bilingual discourse, especially when the speaker is making an argument.

6.4 Alternation, Diglossia, and Community

While Tok Pisin and English can be supposed to exist in a diglossic relationship, code choice and Alternation are expected to be governed primarily by situational and indexical factors. In contexts such as the National Parliament or fluent bilingual communities where language indices are more fluid, Alternation as an expressive and rhetorical technique comes to the fore. In these contexts, Alternation between Tok Pisin and English may be considered to form part of what has been described as an 'anglicised' register that is opaque for those who do not understand English. Certainly, in his work on Parliamentary speeches, Nekitel (1998) found switching between Tok Pisin and English reduced the comprehensibility of the language used for his speakers.

It is not clear from Nekitel's study, however, how much of the language he surveyed featured Insertional code-switching, the subject of the next chapter. He was largely concerned with tracking the language with which a speech begins and ends, a difference indicating an Alternational switch at some point, and does provide examples of language that can be categorised as both Alternation and Insertion. Nekitel finds 50% of his Tok Pisin speeches featuring bi-directional switching, meaning switches to English and back to Tok Pisin again, though he categorises the majority of his samples as monolingual. For speeches begun in English, the tendency to remain in English is over 60%. The proportion of uni-directional

switches, where the speaker concludes in Tok Pisin, is higher than for speeches begun in Tok Pisin and ending in English. I observed this same pattern of uni-directional switching in the discussion of s6.2.1. In the core corpus analysed here, with closer attention to the Insertion of vocabulary, and from a sample taken a decade later than Nekitel's, I found 75% of Tok Pisin utterances featuring bidirectional switching, and, conversely, almost 90% of English speeches remaining monolingual (Paliwala 2004).

Across the community as a whole studies have consistently shown that English is making inroads into all but the most remote rural areas (Kulick 1992; Romaine 1992; Siegel 1997; Nekitel 1998; Smith, G. 2000, 2002), with English-medium education a significant vector. Consequently, expectations about language use in a particular setting may need to change.

In the examples above I have shown that bilingual Alternation between Tok Pisin and English serves Papua New Guineans very well, providing them with a repertoire of expressive and communicative functions that they use as fully and fluently as any multilingual community. In PNG I have found situational, indexical, pragmatic, and metaphorical code-switching of the Alternational type. Such Alternation is clearly a widespread feature of the National bilingual speech community.

Such Alternations, following Muysken's typology, have been presented here primarily as examples of meaningful CS, though Auer (see s3.3.5) has noted that not all Alternations bear local meaning. This has been considered in s6.2.3 in terms of Myers-Scotton's notion of un-marked switching. The specific intricacies of unmarked CS will be considered in more detail in the following two chapters.

7 Findings 3 : Insertion

In addition to patterns of Alternation between languages, discussed in the previous chapter, language mixing also takes place within the sentence structure itself. While Poplack and others concerned themselves with establishing regularities and constraints on the conjunction of two language systems, evidence from many language communities (see s3.2.5 and s3.2.3.3) revealed that in some cases one language system dominates in bilingual production.

The MLF model of code-switching (Myers-Scotton 1993a) looks at mixing within the sentence, or the CP in later developments (Jake and Myers-Scotton 1997), in terms of a Matrix Language (ML), which sets the morpho-syntactic frame for production, and an Embedded Language (EL), which can be used to introduce words and larger segments of speech into that framework. Distinguishable from the balanced bilingual production implied by Alternational switching, Insertional code-switching has asymmetry at its heart: one language is dominant in establishing the framework for communication.

In Papua New Guinea, 'anglicised' Tok Pisin is characterised by the inclusion of many features from English. Studies of first language ('creole') Tok Pisin (Smith, G. 1990; Romaine 1992; Smith, G. 2002), and of the expanded grammar of the written register (Verhaar 1995) have included new phonology, morphology, and syntactic structures in their descriptions of the language. While part of efforts to establish the consistency and strength of Tok Pisin as a language around Independence was the reinforcement of its capacity to express new concepts from internal resources, a generation later borrowing from English appears to be the dominant source of new material.

Insertion as Myers-Scotton et. al discuss it (Myers-Scotton 1993a; Jake 1994; Myers-Scotton and Jake 1995; Myers-Scotton 1997b; Myers-Scotton and Jake 2000) can be seen to characterise some typical examples of an 'anglicised' Tok Pisin.

As a lingua franca, Tok Pisin has historically been the only language shared by most of its speakers. Mühlhäusler's assertion of the importance of development based on internal resources (Mühlhäusler 1985a: 50, 60; 1985b: 282-284; 1985d: 140, 145; 1985e: 614) (see also s5.2.6) is predicated on this understanding. However, more recent research such as Romaine (1992), and G.Smith (2002) has made it evident that, especially for the younger generations of Papua New Guineans, English is an additional shared medium of communication. Such a shared resource is a convenient source for innovations, and so, even among second-language speakers of Tok Pisin, borrowing from English is now preferred for the elaboration and development of the language. G.Smith (2002) suggests that borrowing dominates internal processes of language development for creole speakers in his corpus of children's creole speech. In contemporary Papua New Guinea there is effectively a situation where Tok Pisin is no longer a pidgin for communication between language communities but is itself an established language system with its own growing community of first and second-language speakers (s2.6 and s3.1.4). Beyond the level of international visitors and elite contexts, English too exists as an established language used in education, government, and media, with generations of schoolchildren throughout the country being acquainted with its vocabulary and grammar. G.Smith's characterisation of the present situation as "the co-existence of two developed systems in contact" (2002: 199) precisely identifies the circumstances that make code-switching analyses of 'anglicised' Tok Pisin so essential.

Following Muysken (2000), I consider here 'borrowing', 'nonce borrowing', and constituent Insertions as all involving the same processes, and do not distinguish between them for purposes of my code-switching analysis. Where Poplack and associates define and discard borrowed forms from their analysis of code-switching (Poplack 1980; Sankoff, D. and Poplack 1981), only evidence of community-wide acceptance of loan words can prove that borrowing has taken place. While Verhaar (1995) and a range of dictionary projects aim to record new community norms, taking my lead from G.Smith's (2002) larger dataset, I

recognise that my core corpus of language is nowhere near sizeable enough to establish the status of any particular word. Just as G.Smith (2002) identifies words not in Mihalic's dictionary (Mihalic 1971) as being new in Tok Pisin, I take that dictionary as a starting point and analyse 'nonce borrowings' and loan words from English as Insertions in order to illustrate the applicability of code-switching theory to anglicised speech. The relationship between borrowing and Insertion is discussed below, and also the presence of EL 'islands' (Myers-Scotton 1993a).[53]

Borrowing from English has always been associated with the 'anglicisation' of Tok Pisin. While some borrowing is accepted as valuable expansion of a language, such as when new words are introduced for new cultural concepts (like *gavman* from English *government*), other types of borrowing are generally seen as degradation, as when distinctive vocabulary is replaced by new English loans (such as the replacement of Tok Pisin *diwai* by English *tree*) (Romaine 1992; Smith 2002). Significantly then I consider Myers-Scotton's (1993a) distinction between 'cultural' and 'core' borrowings as reflecting the difference between lexical expansion and anglicisation in Tok Pisin. Cultural loans expand the referential adequacy of the language. Core loans replace existing forms with English equivalents. Both are analysed here as Insertions.

Taking Mihalic's admittedly outdated dictionary (1971) as a pre-Independence starting point, this chapter provides a consideration of the Insertion of new English material into Tok Pisin.

---

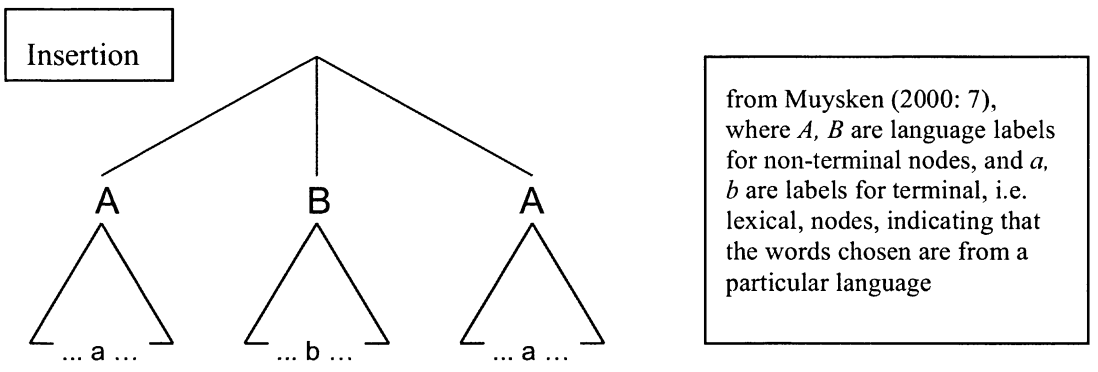
[53] I am undertaking a consideration of EL islands under a lexicalist framework and their relationship to the cultural borrowing of idiomatic phrases such as e.g. *Law and order* without an associated analysis, by Tok Pisin speakers, of their internal structure. However, this theoretical discussion is not included here as it is not central to this thesis.

Examples of Insertion are provided in s7.1 to illustrate the general behaviour described. Section s7.2 discusses how the application of basic criteria of integration to distinguish 'borrowed' and 'nonce' forms is problematic against the backdrop of Tok Pisin's status as an expanded pidgin/creole, and establishes that English elements appearing in the corpus can be analysed as Insertional code-switching. Section s7.3 provides evidence and analyses of Insertional code-switching following Myers-Scotton's MLF model (Myers-Scotton 1993a), and in particular according to her later 4M classification of content and different types of structural morphemes (Myers-Scotton and Jake 2000; Myers-Scotton 2002; Myers-Scotton and Jake 2009).

7.1 Examples of Insertion

The structure of Insertional code-switching is different from that of Alternation, seen in the last chapter. Insertion is characterised by Muysken as the embedding of elements from one language in another, as in Figure 7.1:

Figure 7.1 : Insertion - Structure (ref. Muysken 2000: 7)



Parts of speech from the 'embedded language' (EL) appear bracketed by those from the 'matrix language' (ML). Under the MLF model of Myers-Scotton, the EL elements will be content morphemes structured by ML system, or 'function', morphemes. Prototypical English content morphemes are nouns and verbs, which under the MLF will appear in Tok Pisin well-integrated into its morphosyntactic framework. Under the 4M classification of morpheme

types (Myers-Scotton and Jake 2000; Myers-Scotton 2002; Myers-Scotton and Jake 2009) system morphemes can be classified into 'early system morphemes', such as plural affixes, and 'late bridge system morphemes', such as some uses of conjunctions, which may be inserted into the ML. Only 'late outsider system morphemes', such as the copula in English or the predicate marker in Tok Pisin and some uses of prepositions and pronouns, may not be inserted from an EL into an ML. The development of the 4M model was required in response to datasets that included violations of the simpler content/system morpheme classification of the original MLF (s3.2.3). As we shall see below, an Insertion analysis allows us to understand the presence of a range of English morphemes in Tok Pisin corpora, including my own, and also to understand the presence of some Tok Pisin morphemes in the English data within that corpus.

In my core corpus there are 629 identified examples of A-B-A Insertion sequences. Indeed, Insertion of constituents from one language into another is the most common strategy of language combination in this dataset. The makeup of the inserted constituents ranges from singly-occurring nouns as cultural loans, to EL islands of considerable complexity.

Where G.Smith (2002) identified only 80 of his 536 speakers as making some kind of switch to English, my data collection and text-selection was specifically targeted at recording and uncovering mixed language (see Chapter 4), and all of the core corpus texts contain code-switching at some level. Significantly, however, this code-switching is recorded taking place not just in political debates or in the conversation of urban bilinguals, but also in Radio talkback and broadcast on the national Tok Pisin radio station.

The following examples from a political speech illustrate these kinds of code-switching, where individual nouns and verbs as well as multi-word constituents from English appear fitted into a Tok Pisin sentence structure.

(176) TiT#35b-203_L539

olsem ***let uhs look*** nau ## ***seriuhsli*** long ***sitiuaishin*** # lo ***law an awda sitiuaishuhn*** # na ***reimprovais*** # ***improvais*** o ***ovahol*** olgeta ***disiplin fawses*** blo yumi

same let 1pl look now ### seriously PREP situation # PREP law CONJ order situation # CONJ reimprovise # improvise CONJ overhaul all discipline-0 force-PL PREP.POSS 1pl.INCL

olsem ***let us look now***/nau, ***seriously*** long ***situation***, long ***law and order situation***, na ***reimprovise***, ***improvise*** o/***or overhaul*** olgeta ***discipline forces*** bilong yumi

'so let us look now, seriously at (the) situation, of law and order situation, and reimprovise [sic], improvise [sic] or overhaul all our discipline(d) forces'

Apart from the conjunction present in what is discussed at s7.3.4.2 as an EL island, the English phrase *law and order situation,* and the similar idiomatic EL constituent island *let us look*, all the system morphemes in (176) come from Tok Pisin: emphatic *nau*, prepositional *long*, the conjunctions *na* and *o*, the quantifier *olgeta*, the possessive *bilong* and the pronoun *yumi*. The adverb *seriously,* the noun *situation*, the verbs *reimprovise* and *overhaul* and the adjective and noun *disciplined forces*, the majority of the content, is from English. In (180) Tok Pisin provides the Matrix for the presentation of English-derived content.

Following Myers-Scotton (1993a) and Myers-Scotton and Jake (2000) I consider single English words appearing in the transcripts as falling within three sets, illustrated in (177).

(177) TiT#35b-203_L60

we i ***kos***im planti ah-, ah- ***konsuhn*** long ***gavman*** dede

where 3sg PM cause-TR plenty ah- # ah- concern-0 PREP government today

we/***where*** i ***cause***-im planti ah-, ah- ***concern*** long ***government*** tede

'where (it) causes many ah-, ah- concern(s) for (the) government today'

From right to left the English elements in (177) represent examples of Cultural Insertions (*gavman* 'government'), Core Insertions (*konsuhn* 'concern'), and minimal Mixed Constituents (*kosim* = 'cause'+-*TR*).

'Cultural' nouns are borrowed for non-indigenous objects or concepts, for example *government.*

'Core' nouns replace words that are already established in Tok Pisin. Here *concern* from English has a pre-existing Tok Pisin equivalent: *wari.*

'Mixed Constituents' are the third category of single-word Insertion, and cover in (177) the English verb *cause* that appears with the Tok Pisin transitive suffix *-im.* Transitive marking is a feature of Tok Pisin grammar that distinguishes it from English (see s5.2.1). The use of Tok Pisin morphology on borrowed English verbs creates minimal "mixed constituents" (Myers-Scotton 1993a, 1997), and essentially integrates the English Insertions with the Tok Pisin matrix grammar. The coined verb *kosim* replaces the established Tok Pisin verb *mekim* 'cause'.

These different categories of inserted elements, and their impacts on the Tok Pisin language, are discussed in detail in this chapter.

In terms of language mixing, Insertion is syntactically challenging in a way that straightforward Alternation is not. Furthermore, in contrast to the significance of markedness in the performative functions of alternation, Insertional code-switching can take place as an "unmarked code" (Myers-Scotton 1993a, 1993b, 1998b, 2002). While speech 'anglicised' with code-switched elements may be itself marked, individual Insertions are not necessarily marked in the same way as the particular language Alternations presented in Chapter 6.

This chapter addresses the analysis of such sections of speech in terms of the integration of English elements under an Insertional model.

7.2 Borrowing and Expansion

As discussed in Chapter 5, Tok Pisin, through its origins as a pidgin language, shows evidence of borrowing throughout the grammar. In the lexicon, Mühlhäusler identified the sources for much of the basic vocabulary as being the English, German, and Tolai languages (Mühlhäusler 1979a). In terms of the functional elements of the language G.Sankoff,

Romaine, and others (Sankoff, G. and Brown 1976; Sankoff, G. 1977; Romaine 1992) confirm the grammaticalization of adverbial phrases like *baimbai* as verbal markers, showing a sequence of borrowing of phrases from English (*bye-and-bye*) and their adaptation as obligatory markers (*bai*) in the grammar of Tok Pisin.

In his discussion of sociolects in TP, Mühlhäusler points to borrowing from English as the primary source of differences between 'Urban Pidgin' and the more conservative 'Rural Pidgin' (Mühlhäusler 1975). For Tok Pisin as a creole language, such borrowing has also been discussed by Bickerton as part of the processes of decreolization whereby language change in Tok Pisin reshapes it into a dialect of its superstrate, English (Bickerton 1975a).

In discussions of language mixing, borrowing (B) and code-switching (CS) are often contrasted. The concern is with evaluating when singly-occurring words from one language are to be considered a part of the standard vocabulary of another, and when they are spontaneous examples of language mixing. Poplack (1980, 1981) routinely disregards borrowed words, identified on the basis of criteria of integration, as outside the scope of code-switching analysis. The application of this rubric to English forms in Tok Pisin is discussed below. Myers-Scotton uses a "three occurrence metric" to distinguish borrowed (B) forms from novel CS Insertions in her corpus on the basis of frequency (Myers-Scotton 1993a: 68, 207). She hypothesises that borrowing occurs through the naturalization of forms introduced by CS (Myers-Scotton 1993a: 174). This is especially likely in the case of pidgin language development, because the reduced vocabulary of languages like Tok Pisin leaves open many situations where a word or expression from another language may fill a communicative gap. In time, occasional switches between languages become natural expressions in either. Following a similar tack, Romaine (1992) uses a single-occurrence metric to determine when an English word in her Tok Pisin corpus represents a 'nonce-loan', as opposed to a borrowing. The size of my corpus makes it inappropriate to use any such frequency-based

criterion, which is arbitrary in any case and subject to idiolectal variation and limitations on the scope of discourse.

Healy's (1975) criterion for determining when a word should be considered Tok Pisin (i.e. when it is included in a well-maintained dictionary) is clearly insufficient under present conditions: the benchmark Dictionary of Tok Pisin (Mihalic 1971) represents an account of Tok Pisin pre-Independence, and while many recent studies since its last edition have reported words in common use that are not featured in that text, a consideration of, for example, Baing et. al.'s (2008) updated work reveals the inclusion of words, e.g. *painapol* for 'pineapple'[54], that have obviously been directly adopted from English.

In order to distinguish borrowings from bilingual CS forms, G.Smith (2002) treats as Tok Pisin all lexical items that are phonologically, morphologically, and syntactically integrated into the text, effectively following Poplack's (1980, 1981) criteria for identifying borrowings. I prefer to base my judgements on observations of norms within the communities of speakers. However, as G.Smith explicitly highlights (2002: 205), determining whether a non-standard lexical item has entered the 'mental lexicon' of a speaker as a Tok Pisin word on the basis of frequency or use in monolingual communities requires a larger-scale survey than has yet been attempted.

In this section (and as discussed in s5.4) the presence in the corpus of all lexical items not in Mihalic's Dictionary (Mihalic 1971) are classified, where appropriate, as examples of CS of the Insertion type. My justification, following Myers-Scotton (1993a: 174) and Muysken

---

[54] Older, conservative, forms are *painap* or *ananas*.

(2000: 75), is that any English words that have become established in the Tok Pisin lexicon since Independence must have been initially introduced as code-switched elements by bilingual Tok Pisin/English speakers who used an English word in addition to, or in replacement of, elements of their established Tok Pisin vocabulary. Indeed, in the core corpus data are examples of the appearance of code-switched words alongside an explanation of their meaning, as in (178) (included as (156) in my discussion of alternation for reiteration in s6.3.1.1) :

(178) TiT#35b-202_L265

blong wanem yumi kolim ***budo planIng*** ?
butum a planing
statim lo has na i go antap lo hed blong en

PREP.POSS why 1pl.INCL call-TR bottom.up planning
bottom up planning
start-TR PREP bottom CONJ PM go above PREP head PREP.POSS 3sg.NEUT

bilong wanem yumi kolim ***bottom-up planning?***
bottom up planning
statim long as na i go antap long hed bilong en

'why do we(incl) call (it) 'budo planning'?
Bottom up planning.
Starts at the bottom and travels to its head'

Questions of integration are addressed in the context of an appreciation that Insertion of new material is the first step to its becoming established as a borrowing, and some examples of the kind of expressive use of new material that can allow people without an understanding of English to acquire new vocabulary by listening to code-switched speech are given. Flagged switches are discussed below in the context of distinguishing borrowed words from code-switching, though this is not a straightforward diagnostic.

7.2.1 Cultural Loans

Cultural loan words (Myers-Scotton 1993a) are very common English inclusions in Tok Pisin today. Singly-occurring words for new objects, concepts, or properties new to the developing language community are often drawn directly from English, as below:

(179) TiT#24b-15_L459

ol i ***siti*** bilo mipela nau

3pl PM city PREP-POSS 1pl-EXCL now

ol i ***city*** bilong mipela nau

'they are our city now'

(180) TiT#10b-05b_L299

yeah lo i no stret lo em lon ***pablik*** nogat trru

yes PREP PM NEG straight PREP 2sg PREP public NEG EMP

yes long i no stret long em long ***public*** nogat tru

'(it is) not right for (the) public at all'

(181) TiT#10b-07_L475

sambla kain ***risuhn***

some-ADJ type reason

sampela kain ***reason***

'some kind (of) reason'

(182) TiT#10b-05b_L169

lo gudpla ***helthi*** ol man kAn baim lo keke

PREP good-ADJ health PL man can buy-TR PREP eat

long gudpela ***healthy*** ol man ***can***/ken baim long kaikai

'in good health men can buy to eat'

New nouns such as *city*, *public*, *reason*, and *health* appear throughout the anglicised speech in the corpus, and can be considered as new lexical material in an expanded Tok Pisin vocabulary, and anglicised pronunciations of shared words like *can/ken* are heard. Conversely, Tok Pisin cultural loan words are used by Papua New Guineans in their English:

(183) TiT#24b-15_L820

es gona bi a big singsing
Pron.COP go-CONT COP DET big dance

there's gonna be a big singsing
'there's going to be a big dance'

This process of cultural borrowing is a typical case of expansion, and is to be expected in any language community undergoing cultural change. This type of conceptual expansion is licensed under the MLF model, where the Insertion of content morphemes from either language is possible.

However, Myers-Scotton (1993a) chooses to disregard all cultural B forms from her analysis of CS forms, regardless of frequency. She does this because her specific research objective is to investigate the grammatical constraints on switching within the sentence (intra-sentential CS), and observations of cultural B forms do not contribute to her analysis. This decision is reasonable for those cases where the cultural material has been completely integrated into the ML. The significance of integration for English words in Tok Pisin is discussed below (s7.2.3). Furthermore, in line with the research project of observing and accounting for the various interactions between Tok Pisin and English, cultural loans are a clearly valuable source of information on the level of cultural exposure experienced by speakers.

In the literature on the development of Tok Pisin 'cultural loans' are seen as a significant part of the development of the Urban Pidgin sociolect (Mühlhäusler 1975, 1985d). This is largely due to the recognition that the urban context in Papua New Guinea has previously been the site of greatest cultural contact between traditional communities and the metropolitan, English language, culture. This is particularly true of the texts of speeches made in the National Parliament. Here the context and topics of conversation encourage a large number of English words to be incorporated into Tok Pisin. A clear example of this was (176). In another example of political speech (184) all the English words are cultural loans related to the activity of government and administration:

(184) TiT#35b-202_L297

mipela sindaun na mikim ***desi*** na ***basiet*** blong wanwan # ***kounsela*** bilo yumi ol i wantaim ***presiden*** ol i ***pasing*** long en

2pl.excl sit CONJ make decision CONJ budget PREP each councillor PREP.POSS 2pl.incl 3pl PM with president 3pl PM pass-CONT PREP 3sg

mipela sindaun na mekim ***decision*** na ***budget*** bilong wanwan, ***councillor*** bilong yumi ol i wantaim ***president*** ol i ***passing*** long en

'We sit down and make (a) decision *and (a) budget for each one of our counsellors, they with the president approve them.'*

All these examples of Insertion are singly occurring content words, the clearest way in which borrowing presents itself.

The frequent use of new cultural words without offering a translation is one significant way in which Tok Pisin can lose its value as a lingua franca. Particularly when topics remote from the traditional life of rural people are discussed, many of the words used are integrated B forms taken from English. Without knowledge of either the EL donor language or the specifics of the culture that requires these concepts, many speakers fluent in the ML, Tok Pisin above, would become lost as to the meaning being conveyed. Though the data on multi-literacy, as discussed in s2.6 and s3.1.4, indicates knowledge of English is increasingly widespread, it is by no means the case that all Papua New Guineans use the language, and conservative, 'Rural', Tok Pisin remains the dominant lingua franca.

In other cases, cultural loans overlap with existing vocabulary. One example of this (185) comes from a Tok Pisin radio broadcast, where the DJ gives listeners the telephone number of the radio station:

(185) TiT#10b-05b_L8

thri tu zero wuhn nain faiv faiv o ***thri tu zero wuhn nain faiv thri*** suhpos yu laik ***ko***- ah- wokim totok tuh lo monin

three two zero one nine five five CONJ three two zero one nine five three COND 2sg MOD.Prox call ah- work-TR talk today PREP morning

three two zero one nine five five o/***or three two zero one nine five three*** sapos yu laik ***ko***- ah- wokim toktok tede long monin

'three two zero one nine five five or three two zero one nine five three if you want to ca- ah call this morning'

While Tok Pisin numerals follow a different paradigm (s5.2.2), the use of English numbering has been observed in increasing use since the 1930s (Reed 1943: 283) (see also Mühlhäusler's 'post pidgin' example s3.2.5). Telephone numbers are culturally new to the language community, and commonly follow the English form. The self-correction in the Tok Pisin part of this example *...sapos yu laik ko- ah wokim toktok* ... will be discussed below in terms of the Insertion of single words as example (214).

Similarly, the following examples have English-form numerals describing period of time in terms of years, again relatively new concepts to indigenous and rural culture:

(186) TIT#35b-202_L36

lo luluk lo mi long ## ***tuhtisiks yuhs*** blo mi nuh insaid lo ***paluhmuhn***...

PREP look PREP 1sg PREP ## thirty.six years PREP.POSS 1sg now inside PREP parliament

long lukluk long mi long, ***thirty-six years*** bilong mi nau insaid long ***parliament*** ...

'from what I have seen in, my thirty-six years within parliament ...'

(187) TiT#10b-07_L179

so wai yumi wet # ***nain yias*** nau ***an so*** suhpos ol i gat ...

DISC.CONJ why 1pl-INCL wait # nine year-PL now CONJ DISC COND 3pl PM have ...

so why yumi wet, ***nine years*** nau ***and so*** sapos ol i gat ...

'so why (did) we wait, nine years already and so if they have ...'

Though Tok Pisin would have as alternative constructions *tripela-ten six yia* and *ninpela yia,* it is likely that these cultural concepts (telephone numbers, years of age) would have been learnt through English-language education, and so are code-switched here rather than being expressed in the Tok Pisin paradigm. Further indications that these constructions are code-switched from English are the hesitations flagging the English-style numerals, hesitations that mark also the absence of a Tok Pisin plural marker to integrate the phrases into the language matrix. A more detailed discussion of this is presented below (s7.3.4.1).

While number systems exist in both languages, and there is evidence that under the influence of English language education the Tok Pisin paradigm is less and less used, there are other cases of English words replacing Tok Pisin vocabulary.

7.2.2 Core Loans

Language-mixing between a creole and its primary lexifier and putative superstrate make the subject of core loans a complex and significant one. Tok Pisin is a developing language that has grown in the last two generations from a second-language pidgin of inter-communication. Today it is a national language with large numbers of first-language speakers who use it alongside and in interaction with speakers whose fluency is not impaired by it being their second language. We can characterise it as a post-pidgin (Mühlhäusler 1985a, 1985d, 1985h; Siegel 1997) whose grammar and system is well established both in literature and in the language community. Such a characterisation, however, requires us to be careful in the definition of 'core' borrowing.

I consider G.Smith's list of "[b]orrowed items with referentially adequate alternatives" (Smith, G. 2002: 99) to be examples of 'core' borrowing. Two excellent examples from this short list are the borrowed words *araiv* = 'arrive' and *bihevia* = 'behaviour', which substitute for the pre-existing Tok Pisin words *kamap* (< 'come up') and *pasin* (< 'fashion').

Somewhat different are the words in the list of "[b]orrowed items of greater semantic specificity" (Smith, G. 2002: 99). While these are also 'core' B forms in that they substitute for existing words, if we take Myers-Scotton's definition of core borrowings as "items for which the ML always has viable equivalents" (Myers-Scotton 1993a: 169) then we are left with a dilemma. To what extent is *wokim* a viable equivalent for G.Smith's semantically more specific *bildim*? Similarly, the distinction between the new verb *fraim* and the more traditional *kukim*? Crucially, G.Smith's examples highlight the way in which 'core' borrowing is also a

process of expanding the referential capacity of a language when that language is evolving from a pidgin.

Reduction in the size and semantic specificity of the lexicon is one strategy in pidgin formation (Wurm and Mühlhäusler 1985b: 5; Siegel 2008: 1-3). In the case of borrowing into a pidgin or creole, then, we must appreciate that the 'one form, many meanings' tendency in such languages, including Tok Pisin, leaves open many cases where a 'viable equivalent' for a B form may exist. Borrowing words to increase semantic specificity is a significant process of lexical expansion in creoles, where 'core' borrowing is not necessarily anglicisation but rather a process of language expansion.

Certainly the presence of the new B forms in G.Smith's corpus is a sign of language change, however, G.Smith's data is not suitable for a specific analysis of mixed language behaviour as the language and genre of the monologues he recorded is consistently that of children's Tok Pisin narratives. In my core corpus dataset I consider most core B forms as examples of CS through bilingual production or as part of a special vocabulary for a particular style or discourse.

While Mihalic's dictionary (Mihalic 1971) provides us with a benchmark for considering whether a particular lexical item found in Tok Pisin speech represents a borrowing, it is not always clear that new vocabulary duplicating existing Tok Pisin expressions represents a replacement of core vocabulary.

Take the following example (188):

(188) TiT#24b-15_L785
bai gat ***komplen*** a?
FUT have complain Q.tag
bai gat ***complain*** a ?
'(they) will have a complaint, eh?'

Mihalic provides *tok* as Tok Pisin vocabulary for *complaint*, as in 'mi gat tok' = 'I have a complaint'. However, *tok* as 'talk' has a range of meanings, and the introduction of *complain* here expands the referential specificity of the language.

This can be contrasted with *concern* in (189):

(189) TiT#35b-203_L247
so disela i wanpela ***konsen***
DISC.CONJ this PM one-ADJ concern
so dispela i wanpela ***concern***
'so this (is) first concern'

Mihalic has two alternatives for this English word: *wok* and *wari*, as in *em i wok bilong ol* = 'that is their concern' or *em i gat wari* = 'they have a concern'. Again, using the English word increases referential specificity, as both *wok* and *wari* have many additional meanings relating to work and jobs generally or to problems and worries, respectively.

Both words can be considered to expand the capacity of the language to express fine-grained meanings. More traditional, un-anglicised, equivalent constructions such as illustrated in (190) and (191), below, would seem unspecific and overly old-fashioned to many Papua New Guineans today.

(190)
bai gat tok a?
'(they) will have a talk eh?'

(191)
dispela i wanpela wari
'this (is the) first worry'

In the case of *concern*, however, the replacement is not necessary. The use of the English word rather than use viable Tok Pisin alternatives represents a choice by the speaker to use an English word to indicate English-type distinctions. Cases such as *concern* are found

elsewhere in the corpus, with *idea* and *problems* in the examples below standing for acceptable Tok Pisin alternatives *tingting* and *wari*.

(192) TiT#35b-202_L618

ati ***aidia*** blo mi em nambawan

I.think idea PREP.POSS 1sg 2sg first

ating ***idea*** bilong mi em nambawan

'I think my idea is the best'

(193) TiT#10b-07_L442

ol ***problems*** ymi gat

PL problems 1pl.INCL have

ol ***problems*** yumi gat

'(the) problems we have'

Again, these words are used to increase the specificity of the language. As with *concern* in (189), *problems* in (193)[55] replaces *wari*.

Tok Pisin, with its narrow pidgin vocabulary and early reliance on internal resources to express a diversity of concepts, is particularly in need of vocabulary to expand its referential capacity. Indeed, *idea* may be considered a useful loan also, as *tingting* can represent English *thought*, *memory*, *mind*, *intellect*, or *opinion*. Tok Pisin words with multiple meanings are fertile ground for replacement with additional vocabulary.

Yet, the concern about the use of English words to replace Tok Pisin alternatives is valid, and justifications based on the need to expand the referential capacity of the language can only go so far. Assuming, as Myers-Scotton and Romaine do (implicitly with their frequency-based

---

[55] See also (232) at section 7.3.4.1 below for a discussion of this construction as an example of double-plural marking.

criteria), that with repeated use a code-switched word becomes an established borrowing in the matrix language, or, with G.Smith and Poplack, that the fluent integration of a word accompanies regular use and suggests it has been borrowed into the mental lexicon, then some words have the potential to push existing Tok Pisin vocabulary into obsolescence.

Consider the words *dress*, *clan* and *complain* in (194):

(194) TiT#24b-15_L717

na sapo yu yusim ***duhs drress*** na wanem bilo narabla ***klan*** na narapla man ba ol i ***komplen*** yu

CONJ if 2sg use-TR DET dress CONJ whatever PREP.POSS other clan now other man FUT 3pl PM complain 2sg

na sapos yu yusim ***this dress*** na wanem bilong narapela ***clan*** na narapela man bai ol i ***complain*** yu

'if you use this dress and whatever of another clan then another man will make a complaint (against) you'

All three words fit well into the utterance. However, though there are no signs of hesitation or sounds that seem out of place in a Tok Pisin sentence the word *dress* is preposed by a determiner without the conventional Tok Pisin form *dispela* but in a short form, *dis*, that resembles the PNG English pronunciation of English *this* (see s5.3.1). The resulting Det+N constituent may be considered an English EL island (see s7.3.5.2 below), though the analysis presented in s8.2.2 highlights the extent to which it reflects congruence between Tok Pisin and English.

Dress, *clan*, and *complain* in (194) are core loans of content. An Insertional analysis under the MLF model would view *dis* as a functional word inserted as part of an English EL island constituent with *dress*, rather than a new loan determiner. In (195) we can see a constructed example replacing *dis* with the Tok Pisin *dispela*, *dress* with the far more culturally specific word *bilas* 'decoration', the word *clan* with *lain*, and *complain* with the somewhat less specific *tokim* without any loss of meaning: underlined sections in (194) and (195) are equivalent:

(195)

na sapos yu yusim dispela bilas na wanem bilong narapela lain nau narapela man bai ol i tokim yu

CONJ if 2sg use-TR DET decoration CONJ whatever PREP.POSS other group now other man FUT 3pl PM talk-TR 2sg

'if you use decoration and whatever of another people then another man will speak to you'

The replacement of 'core' vocabulary in this way is a typical example of what has been criticised about 'anglicised' speech. The use of core B forms, for which there are pre-existing words in the vocabulary of the recipient language, is a possible sign of language shift: the replacement of words in one language with words from another in common use is a clear form of language change. Myers-Scotton clearly associates core B forms with the process of CS. The borrowing of 'cultural' lexemes, it is presumed, takes place to fill gaps in a language when cultural concepts codified in another language become significant to its speakers. As, Myers-Scotton reasserts "core B forms have a 'gratuitous' nature" (Myers-Scotton 1993a: 172). However, Chapter 8 shows how frequent replacement of 'core' vocabulary from different lexicons is a feature of Congruent Lexicalization: CS that indicates an underlying shared grammatical structure.

7.2.3 Integration

The adoption of English words as part of everyday vocabulary has been seen as a significant trend in the development of Tok Pisin. While Mühlhäusler describes the problems with borrowing from English (Mühlhäusler 1985d: 141-146) and though language-internal developments have been widely observed, a national policy of education in English combined with the widening social currency of Tok Pisin in Papua New Guinean society has led to a growing generation of effectively bilingual children. Census data (National Statistical Office 2002b) confirms that English is the most widely known language after TP. As G.Smith states: "the effect of English-medium education is likely to be the major force dictating the nature of

the lexicon for a newly emerging generation who are effectively bilingual in Tok Pisin and English" (Smith, G. 2002: 99).

In the absence of a revised and current dictionary of Tok Pisin, G.Smith accepts as Tok Pisin all words that appear to be well integrated phonologically, morphologically, and syntactically. In early CS research the same criteria of 'assimilation' was suggested to make the same distinction (Haugen 1950). Attempts to rely on partial assimilation, for example phonological, to categorise B forms are clearly flawed: strong first-language (L1) phonology can influence all production, regardless of whether an individual is speaking a second language (L2), mixing the two languages, or is using an established borrowed form.

As a consequence, Poplack (1980, 1981) essentially argues that we should consider as true borrowed (B) forms only those words that are integrated along all three criteria, and so excludes such integrated forms from her analysis of CS.

However, under Myers-Scotton's Matrix Language Frame (MLF) model (Myers-Scotton 1993a) the morpho-syntactic integration of code-switched forms with the matrix language is understood as characteristic of Insertion. The use of phonological integration as a diagnostic of borrowing in Tok Pisin is compromised by the observed variation in the sound system (s5.1) that has been elsewhere described as a consequence of expansion (Laycock 1985b; Romaine 1992; Smith, G. 2002). Therefore, in the absence of satisfactory grounds for distinguishing borrowings from code-switches, I differ from G.Smith in discussing all English forms not found in Mihalic (1971) as potential products of code-switching.

In this section the syntactic, phonological, and morphological integration of some English forms present in the Tok Pisin of the corpus are considered, and illustrate how code-switching analyses of Insertion motivate a better understanding of the core corpus data than can be obtained by assuming all well integrated forms to be borrowings.

7.2.3.1 Syntactical Integration

As discussed in Chapter 5, Tok Pisin and English have broadly similar word orders and constituent structure. Consequently syntactic integration of code-switched forms is regularly achieved. In (194) above, the English word *complain* in the VP *bai ol i complain yu* is well-integrated into Tok Pisin syntax. The system morphemes, in Myers-Scotton's (1993a) terminology, structuring the phrase all come from Tok Pisin.

Similarly, in (196) the English adjective+noun (Adj+N) phrase *management style* is fitted into the Tok Pisin possessive frame:

(196) TiT#10b-07_L62
ol ***manuhdjmuhn stail*** bl ol
PL management style PREP.POSS 3pl
ol ***management style*** bilong ol
'their management styles'

The English AdjP *management style* is marked as plural with Tok Pisin *ol*, and the possessive conveyed with the Tok Pisin preposition and pronoun construction *bilong ol* 'of them'.

Insertional code-switching under the MLF provides a strong explanation of syntactic integration that does not require an assumption that such integration is a marker of established borrowing.

7.2.3.2 Morphological Integration

Unlike the languages frequently studied for cases of Insertional code-switching (Azuma and Myers-Scotton 1990; Myers-Scotton 1993a; Myers-Scotton, Jake and Okasha 1996), the languages in this study, Tok Pisin and English, both have relatively limited affix systems. With its origins as a pidginized language, Tok Pisin has very little basic morphology, and while English has more obligatory morphology than Tok Pisin, particularly in the VP, PNG English shows adaptations that reduce the appearance of some morphemes (s5.2 s5.3). In the

corpus data this shows itself through the ready morphological integration of English words in Tok Pisin, but it also means that a general absence of morphological requirements leaves this as an unhelpful criterion for assessing integration.

As observed in Chapter 5, Tok Pisin affixes are limited to the transitive suffix on verbs *-im*, and an adjectival suffix *-pela*. Most other grammatical relations are indicated by pre-verbal particles, as in the TMA system, or through other unbound morphemes. As a result, Poplack's (1980, 1981) Free-Morpheme Constraint (Chapter 6) is rarely violated.

Because morphosyntactic integration is a fundamental of Myers-Scotton's MLF model (Myers-Scotton 1993a) of Insertional code-switching, this section discusses 'mixed constituents' and the issues around partially-integrated Insertions, which have been termed 'nonce' loans. These will be discussed in greater depth in the consideration of the various types of Insertions, below. Significantly here, I present some examples of un-integrated Insertions, and discuss their significance.

Within my data, English nouns and verbs are the most frequently inserted elements. As these are typical 'content' morphemes, under Myer-Scotton's schema they are prime candidates for Insertion into the Tok Pisin 'matrix'. While the integration of nouns is achieved syntactically, transitive verbs in Tok Pisin are marked with a distinctive suffix *-im*. To be morphologically integrated, all inserted polysyllabic transitive verbs (V-TR) should carry this suffix. Indeed, in the corpus large amounts of English vocabulary are used in Tok Pisin in this way, including many monosyllabic verbs. Table 7.1 contains some examples, with the English part represented in capitalised standard English orthography.

Under the MLF model this type of morphological integration is a requirement, and an indication of fluent code-switching. It does not imply that the resulting form is a borrowing. For Papua New Guinean bilingual Tok Pisin/English speakers, adding the *-im* suffix to an English verb is a convenient and highly productive way of augmenting the Tok Pisin

vocabulary from English sources, and there are many cultural loans in Table 7.1. However, established Tok Pisin verbs like *halavim* 'to help' and the more unusual *surukim* 'to move' are replaced here by English forms *helpim* and *moveim*, representing alternatives to existing, 'core', vocabulary.

Table 7.1 : Insertion : examples of V-TR morphological integration in the core corpus

AIRFREIGHT-im	AIRLIFT-im	AMALGAMATE-im	ASSIST-im
ATTEND-im	AUTHORIZE-im	BLOCK-im	BOOST-im
BUILD-im	CAUSE-im	CELERATE-im	CHECK-im
CLOSE-im	COLONISE-im	COMPARE-im	CONFIRM-im
CONTROL-im	CREATE-im	DECIDE-im	DEVELOP-im
DIRECT-im	DIVERT-im	EMPLOY-im	EXPECT-im
FACE-im	FINE-im	FUND-im	HELP-im
IMPLEMENT-im	IMPROVISE-im	INCREASE-im	INVITE-im
KEEP-im	LOAD-im	MAINTAIN-im	MANAGE-im
MOBILE-im	MONITOR-im	MOVE-im	OVERHAUL-im
OWN-im	PICK-im	PLAN-im	PRACTISE-im
PROVIDE-im	READ-im	RECOGNISE-im	REMEMBER-im
REPEAT-im	RING-im	SAVE-im	SEEKAUT-im
SELECT-im	SET-im	SPEND-im	STANDARDISE-im
SUGGEST-im	SUPPORT-im	TAKE-im	TRAIN-im
UPGRADE-im	USE-im	WANT-im	WEAR-im

The formation of Tok Pisin transitive verbs (V-TR) from the use of an English verb (or, in the case of *mobile*, adjective) root + the transitive suffix *-im* is considered in s7.3 as creating typical minimal mixed constituents. They indicate that the form of code-mixing considered by Myers-Scotton as the Insertion of components from one language into the Matrix Language Frame (Myers-Scotton 1993, 1995, 1997) set by another is occurring in Tok Pisin. As a

mechanism for innovation, this kind of mixing allows for the adoption of almost any English verb, radically complicating the potential task of dictionary-makers.[56]

Where G.Smith (2002: 205) observes that in the absence of an updated Tok Pisin dictionary well-integrated forms like this are candidates for inclusion in an expanded Tok Pisin lexicon, this analysis of code-switching allows us to consider these as CS forms until sufficient evidence of widespread use in the community is established.

7.2.3.3 Phonological Integration

While the close nature of the language pair this dissertation is considering and the open syntax of Tok Pisin make syntactic integration relatively straightforward, the status of the Tok Pisin phonological system provides some questions. As discussed in Chapter 5, there are some close similarities between the vowel sounds of Tok Pisin and PNG English, and some of the phonological rules regarding consonant clusters are shared. The expanded phonemic inventory of Tok Pisin includes many distinctively English phonemes, in particular fricatives, so it is quite possible for an English word in Papua New Guinea to already be well-integrated from the point of view of some Tok Pisin speakers.

In other code-switching contexts a lack of phonological integration is considered a diagnostic feature of the spontaneity with which such Insertions are made as new coinages from one

---

[56] It is likely, therefore, that a dictionary based on reference to published usage is to be preferred as the next step in the codification of Tok Pisin. Bilinguals may innovate a vast range of spontaneous verbs as well as nouns in their mixed speech. Even though any speakers of Tok Pisin unfamiliar with the English root would require a translation, the only current benchmark for widening acceptance is the appearance of such words in print media intended for monolinguals, such as Wantok Newspaper, though it is arguable that this too is effectively bilingual (Walczyński 2012).

language into another. However, the situation with Tok Pisin and English is complicated by their linguistic history (see s5.1 and s7.2.3, above).

The integration of English words into the grammatical framework of Tok Pisin is an accepted strategy of expansion, and has been consistently described in terms of the 'anglicisation' of the language. Hall (1943, 1955) associated anglicised varieties primarily with non-National speakers. Crucially, Laycock (1985b) identified certain phonemes as characteristics of National speakers with "anglicised" Tok Pisin dialects. Mühlhäusler (1975, 1985h) described such varieties as distinctively "Urban Pidgin". Since Mühlhäusler's (1979a) doctoral study of the Tok Pisin lexicon Urban Pidgin has been understood to be characterised by the appearance of an increasing proportion of English vocabulary, and Romaine (1992) and G.Smith (2002) observe that the extension of the set of phonemes used in Tok Pisin may be lexically driven.

In this way borrowing is explicitly associated with an expansion of Tok Pisin phonology, and integration becomes much easier for new vocabulary. The extended phonemic inventory provides circularity to criteria of phonological integration as a diagnostic of established vocabulary: borrowing is seen to introduce new sounds into the Tok Pisin inventory and words with those sounds are considered borrowings.

Within 'anglicised' Tok Pisin, sound changes to English lexemes acquired during the pidgin phase are sometimes reversed. Example (197) shows the use of an English pronunciation for 'thank' [θænk] as opposed to Tok Pisin [tenk].

(197) TiT#35b-202_L673
em taso ***thank*** yu
3sg only thank 2sg
em tasol ***thank***/tenk ***you***/yu
'that is all thank you.'

Using Poplack's criteria of integration *healthy* in (198) (part of (100) above) could be identified as a code-switch on the basis of the presence of /θ/, but the same speaker omits this

anglicised phoneme from the word *health* in another context, removing the final consonant from the cluster entirely (199):

(198) TiT#10b-07_L165
ol i ***helthi*** olsem
3pl PM healthy same
ol i ***healthy*** olsem
'they are as healthy'

(199) TiT#10b-05b_L270
tudei em ***wuhld hel dei*** a ?
today 3sg World Health Day Q.TAG
tede/***today*** em ***world health day***/dei a?
'today is World Health Day eh?'

The presence of /ə/ in [wəld hɛl dɛi], represents a phoneme outside the classical Tok Pisin inventory but, like /θ/, included in accounts of 'anglicised' varieties.

Taking the classical inventory as the system to which borrowings should comply, then both *World Health Day* and *healthy* are not borrowings, as defined by Poplack. However, as both these phonemes are part of the anglicised inventory, the English elements in (197-199) could all be considered to be well-integrated, and therefore to be borrowings.

Using phonological integration with an anglicised inventory as a criterion for distinguishing borrowings from code-switches is even more problematic when considering new words with new phonemes, such as *ginger*, which carries the phoneme /ʤ/. In (200) this word is used by two speakers in English, and in (201) in a Tok Pisin radio advertisement:

(200) TiT#24b-15_L254

MSM: ***dey mes it # mes it an den ey put in*** ol ah- #

TMS: ***djindja***

MSM: ***djindja an en ol sots of*** aah- # uh- # uh- # ***gu spaisis***

they mash it, mash it and then they put in all/ol ah-

ginger

ginger and then all/ol ***sorts of*** aah-, uh-, ***good spices***

'they mash it, mash it and then they put in all ah- '
'ginger'
'ginger and then all sorts of aah-, uh-, uh-, good spices'

(201) TiT#10b-03_L23

i gat ***kawn bi:f*** wantaim ***chili # kawn bi:f*** wantaim ***djindja*** # na ***kawn bi:f*** wantaim ***kuhri***

PM have corn beef with chilli # corn beef with ginger # CONJ corn beef with curry

i gat ***corn beef*** wantaim ***chilli, corn beef*** wantaim ***ginger***, na ***corn beef*** wantaim ***curry***

'(they) have corn beef with chilli, corn beef with ginger, and corn beef with curry'

The word *ginger*, like *chilli* and *curry* in (201), is a cultural loan word well established and widely used in Tok Pisin. It replaces the more traditional *kawawar*, and its lack of integration with the traditional Tok Pisin system marks it out as a new word but does not affect its status as a widespread borrowing.

The problem with identifying code-switches on the basis of their pronunciation is shown particularly well in (202), where a politician contrasts pronunciations of *manager* both with and without the English /ʤ/ phoneme. Consequently, on the basis of this example it is not plausible to ascribe *manager* as either an established borrowing or a code-switch on the criterion of phonological integration.

(202) TiT#35b-202_L245

... kisim gutpela ***manisa*** bilon mipla # bai mipela i ***manis***im ol ***de*** mi ***manidja***

... get-TR good-ADJ manager PREP.POSS 1pl.EXCL # FUT 1pl.excl PM manage-TR 3pl there 1sg manager

... kisim gudpela ***manager*** bilong mipela, bai mipela i ***manage***-im ol ***there*** mi ***manager***

'...get (a) good manager for us, we will manage them there I (am the) manager'

Significantly, the pronunciation on the transitive verb *manage-im* (*managim*) here [manisim] reflects thorough integration with classical Tok Pisin phonology. Here it may be a cultural loan, but (203) presents a more anglicised pronunciation on the radio:

(203) TiT#10b-07_L96

... yubela bin ***mAnedj***im gud ***fuhns*** lo mipla

... 2pl PAST manage-TR well funds PREP 1pl-excl

... yupela bin ***manage***-im good ***funds*** long mipela

'... you(pl) have managed funds well for us'

In all these cases *manager* and *managim* are syntactically and morphologically well integrated, but phonological integration does not help to determine the status of the words as either code-switches or established borrowings. In s7.3.4.3 below I discuss constructions featuring the root *manager* in Tok Pisin in terms of Myers-Scotton's 4M model.

The vowel system of Tok Pisin is also considered to be subject to some variation (see Chapter 5), with consequences for the form many shared words take in 'anglicised' speech. The presence of English /ə/ is a novel feature of the expanded inventory, and allows some speakers to realise English pronunciations on words that are characterised by vowel-differences in Tok Pisin and English. We have already seen the use of this centralised vowel in (199), and (204) shows its application to Tok Pisin *sapos* 'if' producing a form closer to the English root *suppose* [səpɒs], as well as on Tok Pisin *mani* 'money':

(204) TiT#10b-07_L180

... suhpos ol i got muhni ...

... COND 3pl have money ...

... ***suppose***/sapos ol i gat mani ...

'... if you have money ...'

Similarly, as discussed under variation in Chapter 5, the Alternation of /e/ and /a/ can lead to English-type production of shared words in Tok Pisin contexts such as [kæn] for Tok Pisin *ken*/English *can*, as in (205):

(205) TiT#35b-202_L538

na ba pile i kAn ***divelep***im wanwan peles bilo yumi wanwan

CONJ FUT 1pl.EXCL PM can develop-TR each place PREP.POSS 1pl.INCL each

na bai mipela ***can***/ken ***develop***-im wanwan peles bilong yumi wanwan

'and then we can develop each of our areas one by one'

Similar variation can be seen with *wok/work* in (206):

(206) TiT#35b-202_L161

bai yu lukim ol ***pabli suhvis*** i wawk[57]

FUT 2sg look-TR PL public service PM work

bai yu lukim ol ***public service*** i wok/***work***

'you will see (the) public services work'

With its origins with English *work*, Tok Pisin *wok* as [wok] is absolutely part of the established lexicon. While a pronunciation such as [kæn] in (205) and [wɔk] in (206) are marked in Tok Pisin, and are potentially conditioned by the use of English forms *develop* and *public service,* it should not be considered to indicate that *ken* or *wok* are borrowings, though their form may reflect style-switching in this context.[58]

In this way, the notion of an 'expanded' Tok Pisin phonemic inventory, augmented with English phonemes, means that applying Poplack's criteria of phonological integration to

---

[57] Note that my transcription orthography gives 'aw' for [ɔ])

[58] I have not considered style-switching analyses to date. This may be productive for deeper analysis at the level of discourse and register.

distinguish borrowings from code-switching is highly problematic. In this text, this same speaker uses [wɔk] six times and [wok] ten times, showing a preference for the standard Tok Pisin pronunciation, but also a clear variation that includes English usage. Such variation is also common for other speakers in the corpus.

7.2.3.3.1 Nonce Loans

Romaine (1992) touches on the debate over 'nonce' loans in relation to Tok Pisin. This category of borrowings was introduced by Poplack (1980) and her associates to account for counterexamples to the Free Morpheme Constraint in switching behaviour. Nonce B forms are defined as those that show morphological and syntactic integration, but maintain the phonological structure from their original language. English-derived elements in examples (205) and (206), under this analysis, could be considered 'nonce' loans, as would many of the verbs in Table 7.1.

Some of these are assimilated phonologically, as in (207). Others retain phonemes from English not found in traditional Tok Pisin, as in (208).

(207)

manisim

manage-TR

manage-im

'to manage'

(208)

cha:djim

charge-TR

charge-im

'to charge (with an offence)'

At (207), in line with phonological rules observed by Mihalic (1971) and by Laycock (1985b) the English phoneme /j/ ([ǰ] or [ʤ]) in "manager" [mænəʤə] is realised in Tok Pisin "manesim" [manisim] as /s/ [s], integrating the verb phonologically as well as

morphologically. In (208), however, the phoneme /ch/ ([č] or [ʧ]) is not changed to /s/ [s] as would be expected if it were integrated into the Tok Pisin system (see s5.1.2).

According to Poplack's (1981) criteria, then, "manesim" is a fully integrated form, while "chargim" is a nonce B form. However as we have already seen at (202) and (199), elsewhere *managim* is produced with the /ʤ/ phoneme, the data revealing that, while this interpretation may be correct, integration alone cannot be an acceptable criterion for establishing words' status. While the concept of nonce loans has currency within code-switching literature, the realities of variation and behaviour within the language community, and within the production of a single speaker, make it less valuable.

Indeed, as Muysken (2000) has asserted in a critique of the requirement for this category introduced by Poplack, D.Sankoff, and Vanniajaran (1990a) to account for seeming violations of the Equivalence Constraint, an Insertion analysis removes the requirement for a separate category of 'nonce' loans.

7.3 Insertion Analysis of 'Anglicised' Tok Pisin

The Matrix Language Frame model of Myers-Scotton is most clearly associated with accounts of Insertional code-switching (Myers-Scotton 1993a, 1995; Myers-Scotton, Jake and Okasha 1996; 1997; Myers-Scotton 1998a; Myers-Scotton and Jake 2000). As we have seen above, where Poplack (Poplack 1980, 1981; Sankoff, D. and Poplack 1981) established a basis for code-switching analysis based around Alternation between languages (see Chapter 6) by and large between sentences (inter-sentential code-switching), a study of switches within the sentence (intra-sentential code-switching) as Insertions allows us a better understanding of the 'anglicised' varieties of Tok Pisin.

The core corpus contains examples of the range of mixed language phenomena Myer-Scotton discusses as Insertion: lexical borrowing (as previously discussed); mixed constituents (her

category that includes Poplack et al.'s nonce loans); and EL islands (the Insertion of well-formed L2 constituents into an L1 sentence, or projection of complementizer (CP), discussed below).

Myers-Scotton's MLF model has the grammatical frame of the CP set by the language, termed the Matrix Language (ML), that provides all the system morphemes involved. Choice of the CP frame, the ML, is largely based on sociolinguistic and psycholinguistic criteria. However, it is at the conceptual level of lemma-selection that both languages are active. Content morphemes can be from any language the speaker knows, and can be inserted into the ML CP frame according to speaker intentions. Lemmas contain all the structural information regarding lexical-conceptual structure, predicate-argument structure, and morphological realisation patterns (Levelt 1989) needed to complete the constituent at the surface level as the maximal projection of the lemma (e.g. NP, PP). The selection of some lemmas indirectly elects (Bock and Levelt 1994) 'early' system morphemes (Myers-Scotton and Jake 2000) to complete the expression of a speaker's intentions (Jake and Myers-Scotton 1997). Under the 4M model (Myers-Scotton and Jake 2000), other classes of system morphemes ('late' and 'late outsider') are structurally assigned and language specific, and can only come from the ML (see s7.3.4.3). Significantly, however, as Myers-Scotton and Jake argue (2009) different categories of grammatical morphemes may function as different 4M morpheme classes in different contexts. Some examples from the Tok Pisin/English corpus are considered at s7.3.4.3, below.

The argument Jake and Myers-Scotton (1997) present is that speaker intentions at the conceptual level can place strain on the linguistic system when concepts associated with lemmas from a language that is not the ML are selected. In order to insert the Embedded Language (EL) content morpheme into the grammatical frame set by the ML in intra-CP CS, for Insertion to occur there must be congruence between the EL lemma and any indirectly elected system morphemes, and an ML counterpart. Where such congruence is lacking, bare

forms (where no grammatical marking, from either system, occurs) or EL Islands (constituents that entirely conform to the EL grammar) occur.

The absence of many bare forms in my data is evidence for two features of the languages involved in the core corpus.

The first clear feature is the simplicity of the morphology used by Tok Pisin and English. Recall that Myers-Scotton's original formulation of the MLF model (Myers-Scotton 1993a) was based on a corpus of data illustrating mixed language behaviour between two typologically very different languages. Swahili has complex agglutinative morphology that deploys system morphemes throughout the CP to set the frame. English has very few morphological affixes, Tok Pisin has even fewer.

As a consequence, the conditions under which an English bare form missing requisite Tok Pisin morphology might occur are infrequent. The only clear examples would be the absence of the transitive suffix *-im* on an English loan verb (s7.3.1), or the omission of the pronominal pluraliser *ol* where an English noun is used with plural meaning (s7.3.4.1). We saw some examples of the latter when considering numbers, at examples (184-186) above.

The second feature of the Tok Pisin/English language pair is their typological similarity. With a similar fixed word order and massively similar vocabulary, the morphosyntactic distinctions between the languages (Chapter 5) are limited.

Indeed, particularly bearing in mind the debate on continuum formation in Tok Pisin creole, in some ways the behaviour recorded in the corpus may be considered closer to dialect mixing than mixing of separate codes. Such similarity can only heighten congruence, and high levels of congruence are, under both the Equivalence Constraints and the MLF model, associated with ease of code-switching. However, in Chapter 8 they will also be shown to allow a third type of even more intimate mixing, described by Muysken (2000: Chapter 5) as connecting CS with dialect mixing and variation.

Here, let us consider the types of CS Insertions found in the corpus.

7.3.1 Single Words

Single-word Insertions have been discussed under Borrowing above (s7.2). In particular the ease with which English words are accommodated in the Tok Pisin framework has been highlighted, especially in the case of transitive verbs, as in Table 7.1, with the example *celebrate-im* presented in context as (209):

(209) TiT#10b-07_L455
klostu lo ***selebret***im ***silewa djubili*** bilo yumi ia
near PREP celebrate-TR silver jubilee PREP.POSS 1pl.INCL Q.TAG
klostu long ***celebrate***-im ***silver jubilee*** bilong yumi hia
'soon to celebrate our Silver Jubilee eh?'

In some cases however, a lack of integration marks the Insertion of a single word as a code-switch, as in (210):

(210) TiT#10b-07_L209
mi bin ***hop*** olsem ...
1sg PAST hope same ...
mi bin ***hope*** olsem ...
'I (have) hoped that ...'

Here *hope* is a 'bare form' (Myers-Scotton 1993a), an English verb inserted into a Tok Pisin matrix. The absence of the transitive suffix *-im* required by that language is licensed by the use of the complementiser *olsem*. However, the English past tense suffix *-ed* is omitted on the verb, with anteriority expressed in (210) by the Tok Pisin pre-verbal marker *bin*. Though Tok Pisin phonological rules restrict word-final consonant clusters (Chapter 5) and such -T/D deletion is common even in Papua New Guinean English (Smith, A.-M. 1986), in the absence of a transitive marker, this code-switch stands out as English. Grammatical roles, however, are entirely expressed in Tok Pisin.

In other cases, intransitive verbs fit neatly into Tok Pisin without any additional morphology:

(211) TiT#10b-07_L209

bai ol disela lain i ken ***ring***

FUT 3pl DET clan PM can ring

bai ol dispela lain i ken ***ring***

'these people can ring'

In particular, English nouns are inserted into Tok Pisin with ease:

(212) TiT#35b-203_L371

na mi luki mi no gat wanpluh ***prowguhres*** ## wanpela ***improwvment*** # insaid long ***polis faws*** na insaid long ami

CONJ 1sg look-TR 1sg NEG got one-ADJ progress ## one-ADJ improvement # inside PREP police force CONJ inside PREP army

na mi lukim mi no gat wanpela ***progress***, wanpela ***improvement***, insaid long ***police force*** na insaid long ami

'and I see I (have) not got any progress, any improvement, within (the) police force and within (the) army'

In some cases, however, hesitation, repetition, or hedges flag a particular Insertion as a code-switch:

(213) TiT#10b-05a_L164

yeah luk olsem ol ruhnaut long ol glol- gol- kiau wanem samting blum- wanem ***glob*** o wanem ***watsakolini # balb***

yes look same 3pl run.out PREP PL glol- gol- egg what something blum- what globe or what what.is.it.called # bulb

yeah luk olsem ol runaut long ol ***glol- gol-*** kiau wanem samting ***blum-*** wanem ***globe*** o wanem ***what's-it-called, bulb***

'yes look also they run out of gol- gol- eggs or what thing blum- what globe or what what it is called, bulb'

Example (214) (part of (185) discussed in s7.2.1 above) contains self-correction to replace the English *call* with a Tok Pisin phrase *wokim toktok*:

(214) TiT#10b-05b_L15

sapos yu laik ***ko***- ah- wokim toktok tuh lo monin

COND 2sg MOD.MOD.PROX call ah- work-TR talk today PREP morning

sapose yu laik ***ko***- ah- wokim toktok tede long moning

'if you want (to) call ah- make talk today this morning'

The flagging of such Insertion of content provides one argument, picked up on by G.Smith (2002) for the continuing differentiation of Tok Pisin and English for bilingual speakers: code-switching indicating that a continuum is not in place (Mühlhäusler 1975). However, the radio broadcaster in (214) and (215) could just as easily be making self-corrections to ensure his listeners follow his, elsewhere more heavily anglicised, speech. Indeed, there is evidence in support of this analysis in his use of English *call* elsewhere without hesitation:

(215) TiT#10b-07_L570

mi ksim wanpela ***kawl*** pastaim bai mi kisim ***kuhmuhsuhl breik***

1sg take-TR one-ADJ call first FUT 1sg take-TR commercial break

mi kisim wanpela ***call*** pastaim bai mi kisim ***commercial break***

'I will take one call first (then) I will take (a) commercial break'

7.3.2 Adjective + Nouns

As can be seen in (196), (209), (212) and (215), nominals like *management style, silver jubilee, police force,* and *commercial break* are sometimes inserted in clusters with their adjectives. Though these constructions do not include any English system morphemes, the absence of Tok Pisin morphology (the adjectival *-pela* suffix (s5.2.2) indicates they may be English EL islands. They are fixed collocations appearing in Tok Pisin as a chunk, and may be considered as compounds or single lexical items by monolinguals. Such multiple-word adjective+noun (Adj+N) Insertions are very common in the corpus data:

(216) TiT#10b-09

uh- sapos ol wok lo rausim ***asistid pesen*** blo eNPieF nau

uh- if 3pl work PREP out-TR assisted person PREP.POSS NPF now

uh- sapos ol wok long rausim ***assisted person*** bilong NPF nau

'who is it that is ousting NPF's assisted persons?'

(217) TiT#10b-05b_L146

at[59] taim yu lIk[60] ***autsaids kondisuhn*** ya #### em no gudpla

PREP when 2sg look-0 outside.POSS condition REL ### 3sg NEG good-ADJ

at taim yu luk ***outside's condition*** hia, em no gudpela

'at when you see (the) condition of the exterior, it (is) not good'

(218) TiT#35b-203_L429

olgeta ***singguhl bareks*** i ***klows***

all single barracks PM close

olgeta ***single barracks*** i ***close***

'all (the) single barracks (are) closed'

Examples such as (216-218) show the Insertion of English Adj+N constituents into the Tok Pisin sentence structure. These are not full English NP islands, because they do not include the English determiner. However, the absence of the Tok Pisin adjectival *-pela* suffix indicates a lack of integration with the Tok Pisin system: English grammatical rules apply. Examples (219) (a section of (176), above) and (220), based around a single speaker's use of the English noun *situation*, illustrate how such Insertions represent CS forms, with various adjectives being used to modify the noun:

(219) TiT#35b-203_L539

long ***sitiuaishuhn*** # blo ***law an awda sitiuaishuhn***

PREP situation PREP.POSS law CONJ order situation

long ***situation***, bilong ***law and order situation***

'... at (the) situation, of (the) law and order situation'

---

[59] See also discussion of this example as (303) s8.2.6 for the significance of the presence of the English preposition *at*.

[60] [lɪk] here is an unusual pronunciation of Tok Pisin *luk*, reflecting this speakers fronting of /o/ to /i/. Note also the absence of a transitive suffix here: the usually expected form would be *lukim* 'to see' or *lukluk long* 'to look at'

(220) TiT#35b-203_L609

lo faitim ***kraim sitiuaisuhn*** lo kanuhtri

PREP fight-TR crime situation PREP country

long paitim ***crime situation*** long kantri

'to fight the country's crime situation'

The intimate and regularised nature of the combinations in many cases, however, indicates that Adj+N Insertions are very close to being lexicalised. Lexicalisation of such multiple word Insertions would be required before they could be considered B forms. Examples (221) (again, from (176) above) and (222-224), from political discourse, include such Insertions as new vocabulary items:

(221) TiT#35b-203_L539

awlgeta ***disiplin fawses*** blo yumi

all discipline force-PL PREP.POSS 1pl.INCL

olgeta ***discipline(d) forces*** bilong yumi

'all our disciplined forces'

(222) TiT#35b-203_L570

bigpla ***kwestyen mak*** i stap lo hed bilo mi nau

big-ADJ question mark PM stay PREP head PREP.POSS 1sg DISC

bigpela ***question mark*** i stap long hed bilong mi nau

'(a) big question mark is over my head now'

(223) TiT#35b-203_L472

i gat awlguhtuh ***faiapowa kepabiliti*** balus wanem samthing # na yumi lukluk tasol long en

PM have all firepower capability airplane what thing # now 1pl.INCL look.into only PREP 2sg

i gat olgeta ***firepower capability*** balus wanem samting, na yumi lukluk tasol long en

'(they) have all (the) firepower capability aeroplanes whatever else, now we should look into that'

(224) TiT#35b-203_L310

nau ol ***sekuriti faws*** blo yumi

DISC PL security force PREP.POSS 1pl.INCL

nau ol ***security force*** bilong yumi

'now our security forces'

In Chapter 9, in a consideration of responses to mixed language data by two groups of UPNG students, some conclusions are drawn regarding the salience of these two-word English Adj+N combinations in Tok Pisin. Here, I note that they can be classed as nominal Insertions from English, and that they share many of the properties of single-word noun Insertions in this data.

7.3.3 Two-word Verbs

A.-M.Smith observed the presence of 'two word verbs' in PNG English (Smith, A.-M. 1978, 1988a), and these are inserted into Tok Pisin with a regularity that implies that they have become lexicalised for many Papua New Guineans. In (225-227) are examples of two-word verbs *shaping of, dress up*, and *on and off* used lexically in the Tok Pisin frame, the verb and associated preposition inserted together. In (225) *shaping of* appears as the verb in the PP following the preposition *long*:

(225) TiT#10b-07_L443

yumi ## mas lukluk lo ***seping of*** sampuhluh ***longtem kaunsel wuhk***

1pl.INCL ## must look PREP shaping of some long.term council work

yumi, mas lukluk long ***shaping of*** sampela ***long-term council work***/wok

'we, must look into (the) shaping of some long-term council work'

In (226) *dress up* acts as the main verb and is followed by a PP *long dispela*:

(226) TiT#24b-15_L813

an taim yu ***dress ap*** lo isela yu ***dres into dat kosyum an en # yu gonna gradueit***

and when 2sg dress up PREP this you dress into that costume and then # you going.to graduate

and taim yu ***dress up*** long dispela yu/***you dress into that costume and then, you***/yu ***gonna graduate***

'and when you dress up in this you dress into that costume and then, you (are) going to graduate'

In (227) the three words *on and off* together appear as the main verb with *sampela hap* 'some part' in subject position:

(227) TiT#10b-05a_L211

... *yeah* sampela hap ***awn an of*** ya

... yes some-ADJ part on and off LOC?

... ***yeah*** sampela hap ***on and off*** hia

'yes some parts (turn) on and off there'

Interestingly these Insertions take place in proximity to the Insertion of other English NPs as EL islands. Myers-Scotton's MLF model would resist analysing *shaping of sampela long-term council work* in (225) as whole VP Insertion due to the presence of the Tok Pisin determiner *sampela* as an active system morpheme.[61] Similar patterns of back-and-forth switching are discussed in Chapter 8 as indicative of Muysken's Congruent Lexicalization.

The role of PNG English two-word verbs in the EL vocabulary is indicated by the calque of *set up* in (228):

(228) TiT#35b-203_L223

ol israil i kam pinis ***set***im ***ap*** ol sampla ***beis kemp*** # long aah- # ***monitir***im ol ***aktiviti*** long ol solwara:

PL Israel PM come COMP set-TR up PL some-ADJ base camp # prep ah- # monitor-TR PL activity PREP PL sea

ol Israel i kam pinis ***set***-im ***up*** ol sampela ***base camp***, long aah-, ***monitor***-im ol ***activity*** long ol solwara

'Israeli's have already set up some base camps, to aah-, monitor activities in (the) seas'

---

[61] In the phrase *long-term council work* only the adjective *long-term*, modifying the Adj+N *council work,* is a recent English transfer, with both *kaunsil* and *wok* attested in Mihalic's (1971) dictionary. However, the conservative Tok Pisin *wok bilong kaunsil* has clearly been replaced by the lexicalised English *council work* here, and so I prefer the interpretation of the entire phrase as an EL constituent, modified by the Tok Pisin determiner, as described here.

The Tok Pisin construction *setim up* forms a mixed constituent with the verb root *set* accompanied by system morphemes from both English (the preposition *up*) and the Tok Pisin transitive *-im* suffix. It resembles earlier-attested forms such as *pulimapim*, suggesting that the strategy of creating mixed constitutents through Insertion into the Tok Pisin matrix grammar has been in use by Tok Pisin speakers for some time.

7.3.4 Mixed Constituents vs. Nonce Loans

As discussed in the brief consideration of Nonce loans (s7.2.3.3.1), above, it is in the analysis of mixed constituents, such as transitive verbs (V-TR) in Tok Pisin formed from English verb (V) roots with the addition of the *-im* suffix, that Myers-Scotton's MLF is the most useful. Where morphological integration is taken elsewhere (Poplack 1980), or within Tok Pisin speech (Smith, G. 2002), as indicative of borrowing, the status of such words as ML+EL mixed constituent Insertions under the MLF is a compelling reason to consider Insertional code-switching as an appropriate model for the analysis of anglicised Tok Pisin. In fact, most Papua New Guineans explicitly acknowledge the facility of such code-switching in expanding the referential capacity of Tok Pisin, though as observed by Mühlhäusler and others, this productive borrowing can also lead to the obsolescence of established Tok Pisin words.

In particular, two aspects of morphological integration under the MLF model are considered to explain data on pluralisation and aspect marking found in the data, and reported elsewhere (Romaine 1992; Verhaar 1995; Smith, G. 2002) as indicating the introduction of new morphology into Tok Pisin.

7.3.4.1 Double Marking Plurals

Consistently throughout the corpus English plural forms appear in Tok Pisin. Mühlhäusler (1979a) discusses the presence of unanalysed plural forms such as *anis* 'ant' from English *ants* in the early Tok Pisin lexicon. In the established, expanded, and creole varieties of the

post-Independence era English nouns with the plural *-s* suffix have often been found in Tok Pisin corpora (Romaine 1992; Verhaar 1995; Smith, G. 2002). However, the presence of double-marking on these nouns, where the Tok Pisin pluralising particle *ol* appears before the plural noun that includes the English *-s* morpheme, is understood under the MLF to indicate they are code-switched forms. There are many examples of this in the corpus, both with the English plural suffix on established English cultural loans (Mihalic 1971), *komiti* 'committee' (229) and *direkta* 'director' (230), as well as on more novel words (230-232):

(229) TiT#35b-202_L292

ol komiti***s*** blo mipela

PL committee-PL PREP.POSS 1pl-EXCL

ol komiti-***s*** bilong mipela

'our committees'

(230) TiT#35b-202_L187

ol ***ministuhs*** ## buhlong en # ol direkte***s***

PL minister-PL ## PREP.POSS 3sg # PL director-PL

ol ***ministers***, bilong en, ol direkta-***s***

'their ministers (are the) directors'

(231) TiT#10b-07_L424[62]

ol ***problems*** ymi gat

PL problems 1pl.INCL have

ol ***problems*** yumi gat

'(the) problems we have'

(232) TiT#10b-07_L40

lo ***sel***im ispluh l ol ***kontribjutuhs*** blong ol

---

[62] This is the first instance of this construction in the speech of one speaker, who repeats themself a little later. The later repetition is discussed at (193) in section 7.2.2 on core loans, above.

PREP sell-TR this PREP PL contributor-PL PREP.POSS 3pl
long ***sell***-im dispela long ol ***contributers*** bilong ol
'to sell things from their contributors'

All of the nouns *committee*, *minister*, *director*, *problem*, and *contributor*, both those in Mihalic (1971) and more recent, are integrated into the Tok Pisin matrix through the use of the plural *ol*, and the English pluralising suffix is therefore redundant, and is not active in the grammar.

Plurality is categorised by Myers-Scotton and Jake (1997; 2000, 2009) as an 'early' System morpheme, activated at the lemma-level by the content morpheme and so easily brought through in the realisation of an L2 form in bilingual production. The L1 plural is still required to express plurality within the matrix framework of the language, however.

Consider examples such as (233), where a Tok Pisin plural is formed from an English nominal without English plural marking:

(233) TiT#10b-05a_L66
I ol ***friwe*** blo yumi
PREP PL freeway PREP.POSS 1pl.INCL
long ol ***freeway*** bilong yumi
'for our freeways'

Again, it is the Tok Pisin marker that attributes plurality with the Tok Pisin matrix. As Romaine observes with reference to her corpora, such double marking is very much the norm where English nouns appear with their plural suffix, though G.Smith notes differences with different nouns and some underlying patterns of regional variation (Smith, G. 2002: 73-4). Under an Insertional analysis, the English *-s* suffix is not considered to be a part of the Tok Pisin grammar.

7.3.4.2 Double Marking Aspect

On verbs, double marking of continuous aspect is also observed in the corpus. Again, analyses of other corpora have concluded that the presence of the English *-ing* suffix on verbs can be considered a new part of Tok Pisin morphology. However, constructions such as (234) show that these forms can be considered in terms of Insertion, with the Tok Pisin aspectual phrase *i stap* + VP + *i go* conveying the continuous aspect in the matrix language, and the *-ing* suffix therefore being redundant:

(234) TiT#08-506_L293

em ***fes wan*** # em stap # ***bi:ting*** ol i go[63]

3sg first one # 3sg CONT # beat-CONT 3pl PM go

em ***first one***, em stap, ***beating*** ol i go

'him, the oldest, he was, beating and beating them'

In (235) aspect is conveyed by the Tok Pisin progressive *wok go long*:

(235) TiT#35b-202_L316

mipela sindaun na wok go long ***putting*** disela wan ***miliuhn*** ...

1pl.excl sit.down CONJ (work go PREP)CONT put-CONT DET one million

mipela sindaun na wok go long ***putting*** dispela wan ***million***

'we(excl.) sit and try putting this one million ...'

There are, however, other examples of verbs appearing with the English *-ing* suffix without Tok Pisin aspect marking:

(236) TiT#35b-202_L306

mi yet ***t bi siaman*** mipela i ***pasin***

---

[63] Hesitations such as the pauses here between constituents from different langauges are 'flagging', taken as one indication that Inserted EL material is not part of the monolingual lexicon.

1sg indeed to be chairman 1pl.EXCL PM pass-CONT
mi yet ***to be chairman*** mipela i ***passing***
'indeed we passed (the motion for) me to be chairman'

However, this use of English *passing* for *passed* in the context of a political vote is not expressed here with continuous aspect. The *-ing* suffix has come with the English verb due to activation as *passing a vote* at the lemma level.[64] The context makes clear that the verb is used here without continuous aspect, therefore requiring no additional Tok Pisin markers.

7.3.4.3 4M Model and L2 morphology

Under this code-switching analysis, English morphemes found in Tok Pisin corpora such as Verhaar (1995) and G.Smith (2002) can be considered as code-switched forms. Evidence for such morphology being applied to established Tok Pisin words would be required before the occurrence of English morphemes in such corpora can be considered evidence for language change.

Double marking phenomena such as these support an interpretation of language mixing as Insertion under terms of a Matrix Language Frame. The Tok Pisin grammatical frame also coveys the meanings of bound English morphemes *-s* and *-ing*. As a consequence, Tok Pisin monolinguals are exposed to English morphology, as English morphemes appear in a range of existing corpora.

---

[64] The quality of the nasal here distinguishes the alveolar pronunciation of English *-ing* /ɪŋ/ heard here as [in] on *passing* [pasin] from close homophony with the bilabial nasal on the transitive Tok Pisin *-im*.

In the corpus data are examples of English morphology forming a Tok Pisin verb *managim* the noun *manager* and the adjective *management* from the English root *manage*, for example (237-239)[65]:

(237) TiT#35b-202_L245

bai mipela yet sitdaun na kisim gutpela ***manisa*** bilon mipla # bai mipela i ***manisim*** ol ***de*** mi ***manidja***

FUT 1pl.EXCL yet sit.down CONJ get-TR good-ADJ manager PREP.POSS 1pl.EXCL # FUT 1sg PM manage-TR 3pl there 1pl.excl manager

bai mipela iet ***sit***-down/sindaun na kisim gutpela ***manager*** bilong mipela, bai mipela i ***manage***-im ol ***there*** mi ***manager***

'we will meet and get (a) good manager for us, we will manage them there I (am the) manager'

(238) TiT#10b-07_L97

luk olse yubela bin ***mAnedjim*** gud ***fuhns*** lo mipla.

look same 2pl PAST manage-TR good funds PREP 1pl.excl

luk olsem yupela bin ***manage***-im gud ***funds*** long mipela

'look that you(pl) have managed funds for us well'

(239) TiT#35b-202_L166

blon anem ***ledesip*** bilo ***wawk*** # ***ledesip*** bilo kuhntri ***manismuhn*** blo kuhntri i bagarap

PREP.POSS whatever leadership PREP.POSS work # leadership PREP.POSS country management PREP.POSS country PM broken

bilong wanem ***leadership*** bilong wok/***work***, ***leadership*** bilong kantri ***management*** bilong kantri i bagarap

'why (is) work leadership, country leadership, country management broken?'

---

[65] Examples (237-239) are also considered, in part, above as examples (202-204) in the discussion of phonological integration (s7.2.3.3).

A morphosyntactic Insertional analysis reveals that it is not the English morphology that distinguishes these different forms, for their constituency is signalled by the Tok Pisin frame they appear in: the adjective *gudpela* and the pronoun *mi* marks the nominals in (236), and the transitive verbal *-im* suffix marks the verbs in (237) (also partly discussed as (202), above) and (238), while the possessive preposition *bilong* associates *management* and *kantri* 'country' in (239). While Insertions like this represent an opportunity for monolinguals to be exposed to English system morphemes, as with double-marking, the Tok Pisin matrix precludes any requirement for the analysis of the internal structure of English morpho-syntactic structures.

English loans can become established in Tok Pisin in plural form, as observed by Romaine (1992). Similarly it is possible that the English morphology found by Verhaar in written Tok Pisin (1995) does not expand the grammar of the Creole, but remains unanalysed and unproductive for monolingual speakers who rely on existing Tok Pisin structures to nuance their meaning. Only bilinguals are necessarily aware that double marking takes place or of the regularity of sound changes that indicate word-class in English. More data is required before morphology carried with inserted English words can be truly considered productive in Tok Pisin.

Below, other examples of English system morphemes appearing in Tok Pisin are considered within the 4M model of bilingual production (Jake and Myers-Scotton 1997; 2000, 2009).

Content morphemes have already been considered as the most basic elements to be inserted, in terms of both cultural and core loans, and the expansion and integration of the vocabulary. English plural morphology forms are considered as 'early' system morphemes under the model, as are determiners, and Tok Pisin *-im* may also be considered as falling into this category (Harvey 2007: 81-82). Consequently, under the 4M model these morphemes may be present in the corpus data when they are activated with their content-morpheme lemma items.

7.3.4.3.1 Conjunctions

English conjunctions appear throughout our data, as in (240), where English *so* precedes a Tok Pisin VP, and where English *and* appears for Tok Pisin co-ordinating *na:*

(240) TiT#24b-15_L92

so em painim se ***thuhtin*** o o yu sa ***fotin*** o ***nain eit fiftin*** disla kain nabaut ### ***an*** yu la kamap disla ...

DISC.CONJ 2sg find-TR same thirteen CONJ CONJ 2sg MOD.HAB fourteen CONJ nine eight fifteen DET kind about ### and 2pl MOD.PROX come.up PRON

so em painim olsem ***thirteen*** o o yu save ***fourteen*** o ***nine eight fifteen*** dispela kain nabaut, ***and*** yu laik kamap dispela ...

'so they are nearly thirteen or or you are fourteen or nine eight fifteen around that (age), and (when) you are nearly this (old) ...'

Conjunctions are functional elements of a language, and may be considered as typical system morphemes, so their presence in mixed language poses a problem for the MLF model as initially formulated (Myers-Scotton 1993a). However, in their work on morpheme classes Myers-Scotton and Jake define conjunctions as either 'late bridge system morphemes', indicating that they can be inserted when they form bridges between elements that make up larger constituents, as between *tanim sa thirteen* and *yu laik kamap dispela* in (240), or else as content morphemes in that they assign thematic roles in terms of the relations between constituents (Myers-Scotton and Jake 2009). This argument provides one explanation for examples of the conjunctions *so*, *because*, *but*, as well as *and* in the corpus (as discussed also in s5.2.5):

(241) TiT#10b-05b_L166

mi # ***putin*** se liklik ***bikoz*** ensidisi mas lulu lo disela

1sg # put-CONT DET little because NCDC must look PREP this

mi, ***putting*** dispela liklik ***because*** NCDC mas lukluk long dispela

'I am saying this little thing because NCDC must look into this'

(242) TiT#10b-07_L578

luk ols em i go pinis ***so*** yi kisim ***komuhsuhl breik***

look same 3sg PM complete DISC.CONJ 1pl.INCL take-TR commercial break

luk[66] olsem em i go pinis ***so*** yumi kisim ***commercial break***

'look so they have gone so we (can) take (a) commercial break'

(243) TiT#24b-15_L674

so ## yu mas yusim disla ***drres***

DISC.CONJ ## 2sg must use-TR this dress

so, yu mas yusim dispela ***dress***

'so ... you must use this dress'

(244) TiT#35b-203_L57

(bifoa) ... i gat planti kainkain toktok insaid long ol ***ami*** yet lo ***polis*** yet lo ***an*** ol i gat ***politiks*** blong ol yet

PM have many kind talk inside PREP PL army REL PREP police REL PREP CONJ 3pl PM have politics PREP.Poss 2pl REL

... i gat planti kainkain toktok insaid long ol ami/***army*** iet long plis/***police*** iet long ***and*** ol i gat ***politics*** bilong ol iet

'... there was much talk within the soldiery and police on and they have their own politics ... '

(245) TiT#24b-15_L649

buht yu mas yusim disla kain ***drressin*** olsem we ## yu ***ansesas*** a ?

COND 2sg MOD use-TR DET kind dressing same where # 2pl ancestors Q.TAG

but yu mas yusim dispela kain ***dressing*** olsem we, yu ***ancestors*** a?

'but you must use the same kind (of style of) dress that ... your ancestors (used) eh?'

(246) TiT#35b-203_L222

---

[66] This use of *luk* 'looks, appearance' here does not follow conventional Tok Pisin usage, which, as at (214) above, would be *lukim* 'to see'. Here, the bare form *luk* may represent a calque from some registers of Australian English spoken discourse where *look* 'to see' can be used as a focusing connective.

bikaws long ol solwara bil yumi

CONJ PREP PL sea PREP.POSS 1pl.INCL

because long ol solwara bilong yumi

'because in our seas ...'

Muysken (2000) provides an alternative analysis of the frequent appearance of these elements in mixed language. He observes that these elements are typically inserted in sentence-peripheral positions, between other constituents (see also Chapter 6). As such, they are not embedded within a particular constituent structure, and are prime candidates for code-switching by Alternation. Essentially: these conjunctions are not inserted into the Tok Pisin frame, but are brief switches to English between Tok Pisin phrases. This may be indicated by salient pauses before a speaker continues. As a consequence they are widely used by bilinguals, and come to be a part of all the anglicised registers recorded in the corpus.

7.3.4.3.2 Late Outsider System Morphemes

There are also examples of 'late outsider system morphemes' from English within the Tok Pisin data. Under all the formulations of Myers-Scotton's MLF model, the Insertion of prepositions or pronouns is restricted by the System Morpheme Principal. However, examples such as (247) and (248) (also discussed as (217), above) are present in the data:

(247) TiT#35b-202_L486

yupela go ***tu***[67] toto lo rus ***kawpuhreishin***

2pl go to talk PREP Lus Corporation

yupela go ***to*** toktok long Lus ***Corporation***

'you go to speak with Lus Corporation'

---

[67] Conventional Tok Pisin (Mihalic 1971) has *tu* 'too, also', but prepositional *to* is English.

(248) TiT#10b-05b_L146

at taim yu lik ***autsaids kondisuhn*** ya # em no gudpla

PREP time 2sg look outside-POSS condition EMP # 3sg NEG good-ADJ

at taim yu luk ***outside's condition*** hia, em no gudpela

'when you see (the) outside's condition there, it is not good'

As we will see below (s7.3.5), in the consideration of larger constituent Insertion, Myers-Scotton's model excludes English prepositions like *to* in these examples from appearing in Tok Pisin outside of an 'EL Island': a complete constituent from English inserted into Tok Pisin frame, as in (249):

(249) TiT#08-506_L170

tupla man mi no tokim yu long ***kam on tu mi plein*** stori no

two man 1sg NEG talk-TR 2sg PREP come on to 1sg play-CONT story no

tupela man mi no tokim yu long ***come on to me playing*** stori no

'the two of you: I did not tell you to come on to me telling lies no'

Chapter 8 discusses the significance of the finding of English prepositions in Tok Pisin in contradiction of the MLF model (see s8.2.6 in particular).

The use of a Tok Pisin pronoun in English in (250) can be understood in terms of cultural borrowing because English does not retain an inclusive/exclusive distinction in its pronominal system (see Chapter 5) that is fundamental in Tok Pisin and many indigenous languages in Papua New Guinea:

(250) TiT#24b-15_L682

mipela gat ah- ## ***wi got trris we*** mipela ***livin***

1pl.EXCL have ah- ## 1pl got trees where 1pl.EXCL living

mipela gat ah-, ***we got trees where***/we mipela ***living***

'we have ah-, we have trees where we (are) living'

Myers-Scotton and Jake (2009) also provide an explanation that can be applied to an analysis of the presence of English pronouns *him*, *they*, and *it* in Tok Pisin, as in (251-255), as

examples of Insertion into the Tok Pisin frame by classing them as content morphemes, appearing in argument position and receiving thematic roles:

(251) TiT#35b-05_L90

Mi thank yu long ### ***im***

1sg thank you PREP ### 3sg

mi ***thank***/tenk yu/***you*** long, him

'I thank you for, him'

(252) TiT#35b-05_L95

behain ***im*** bai ***ma:k***im yu

FUT 3sg FUT mark-TR 2sg

behain ***him*** bai ***mark***-im yu

'later he will recognise you'

(253) TiT#08-507_L47

ohh- mi tokim ***im***

ooh- 1sg talk-TR 3sg

ooh- mi tokim ***him***

'ooh- I told him!'

(254) TiT#35b-202_L120

maski # lus ***it***[68]

no.matter # lose it

maski, lus ***it***

'forget it, let it go'

---

[68] This analysis is supported by the repetition that follows *lusim tinting* 'lose (the) thought': reiterating with the traditional Tok Pisin word the noun, 'thought', referred to with the English pronoun *it*.

(255) TiT#24b-15_L456

dei guht ah- # wuhnpluh kain ***li:v*** ia

3pl have ah- # one-ADJ type leaf LOC

they gat ah-, wanpela kain ***leaf*** hia

'they have ah-, one type of leaf there'

Indeed, such pronoun-switching is common in bilingual clauses cross-linguistically (see also examples in s5.2.3 and s8.2.6). The classification of pronouns as content morphemes in such positions is crucial to maintaining the integrity of the MLF model of Insertion, though under the 4M classification they remain closed-class system morphemes.

In Tok Pisin, therefore, the presence of English pronouns in bilingual production need not be considered indicative of an extension to the pronominal paradigm. However, as discussed in Chapter 8 (s8.2.6), the increasing frequency with which such switches occur may be taken to indicate the presence of a third, congruent, grammatical system in Tok Pisin.

Chapter 8 discusses the presence of other 'late outsider system morphemes' in both Tok Pisin and PNG English (including the use in PNGE of a pronoun form in a role congruent with the Tok Pisin predicate marker *i* and the use of the English copula in Tok Pisin) as indications of a shared, congruent structure (s8.2.6). Myers-Scotton (2002) describes the presence of such system morphemes as indicative of 'composite CS', Muysken (2000) as Congurent Lexicalization (s3.3.5).

7.3.5 Larger Constituents

In addition to the small Insertions of English already discussed, the corpus is filled with longer stretches of English appearing in Tok Pisin sentences, as at (176) above. Myers-Scotton's models of Insertional code-switching allows the analysis of larger segments of English inserted within Tok Pisin.

7.3.5.1 ML + EL Constituents

All the English verbs presented in Table 7.1 above represent mixed constituents under the Insertional model. English roots are morphologically integrated into the Tok Pisin grammar through the addition of the transitive *-im* suffix, and form Tok Pisin transitive verbs. Nouns marked for plurality by the Tok Pisin particle *ol* also form mixed constituents. Examples of both of these types of mixed constituents are in (256) (part of which has already been discussed as (224), in s7.3.2, above):

(256) TiT#35b-203_L310

ating yumi mas ah- # ***improvais***im nau ol ***sekuriti faws*** blo yumi

I.think 1pl.INCL must ah- # improvise-TR now PL security force PRE-POSS 1pl.INCL

ating yumi mas ah-, ***improvise***-im nau ol ***security force*** bilong yumi

'possibly we must ah-, improvise [sic] now our security forces'

In example (257) the possessive phrase *capability bilong defence force* is likewise a mixed constituent, with English (EL) nouns combined with a reduced form of the Tok Pisin (ML) preposition *bilong* to form the PP:

(257) TiT#35b-203_L292

mi no bagarapim ***keipabiliti*** blo ***defens fors*** # nogat #mi ***sapawt***im[69]

1sg NEG ruin-TRANS capability PRED.POSS defence force # NEG # 1sg support-TRANS

mi no bagarapim ***capability*** bilong ***defence force***, nogat, mi ***support***-im

'I (do) not ruin (the) capability of (the) defence force, no, I support (it)'

---

[69] As with many other examples of lexis analyzed here as resulting from language interaction (see s7.2 for discussion), *sapotim* 'to support' has widespread currency in Tok Pisin and is commonly found in Wantok newspaper, among other present day registers of spoken and written expanded Tok Pisin.

Such mixing within constituents by the Insertion of content morphemes is the route Myers-Scotton suggests for the introduction of B forms.

7.3.5.2 EL Islands

Under the MLF, English system morphemes should only appear in Tok Pisin within EL 'Islands' of English constituents. Where there are longer segments of English in the corpus, these can be understood as 'island' Insertion into the Tok Pisin matrix, but governed by English grammar internally.

In (180) *let us look* is exactly such an English island, the VP inserted into the Tok Pisin frame established by *olsem VP nau*. Similarly, *law and order situation* is an NP island, framed by the Tok Pisin preposition *long*. However, at this level of analysis the whole PP *long law and order situation* is a mixed constituent.

In (258), from the same text, this mixed PP can be contrasted with a PP island, the temporal phrase *in future*, introduced as the object of the verb *lukim*:

(258) TiT#35b-203_L298

tasol mi la luk lo em olse ***in fiutchuh*** long traim na lukluk lo ***law an oda sitiuashin*** blo yumi

CONJ 1sg MOD.PROX look PREP 2sg same in future PREP try-TR CONJ look PREP law and order situation PREP.POSS 1pl.INCL

tasol mi laik luk(luk) long em olsem ***in future*** long traim na lukluk lo ***law and order situation*** bilong yumi

'I will look in future to try and look into our law and order situation'

In (259) both the noun phrases *law and order problem* and *very hard economic crises* are English islands, following English grammar.

(259) TiT#35b-203_L92

yumi gat ***lo an oda probluhm*** lo kuhntri na kuhntri ***feis***im ***veri*** ah- ***had ikonomi***[70] ***kraisis*** we i no ga moni: lo lukautim ami na ***polis*** # disela taim

1pl-incl have law and order problem PREP country CONJ country face-TR very - hard economic crises where PM NEG have money PREP care-TRANS army CONJ police # DET time

yumi gat ***law and order problem*** long kantri na kantri ***face***-im ***very*** ah- ***hard economy crisis*** we i no gat mani long lukautim ami na plis/***police***, dispela taim

'aah, we have (a) law and order in (the) country and, (the) country faces very ah, hard economic crises where there is no money for maintaining (the) army and (the) police at this time'

Comparable Tok Pisin constructions, using the same nominals, would be as in (260-261):

(260)

problem bilong ***law*** na ***order***

problem PREP.POSS law CONJ order

'law and order problem'

(261)

krisis bilong ***economy*** i had moa

crisis PREP.POSS economy PM hard more

'very hard economic crisis'

Instead, in behaviour characteristic of the highly anglicised political register, political phrases are inserted entirely into the Tok Pisin matrix in preference to expressing the same conceptual relations using the grammar of Tok Pisin.

---

[70] The form here, which resembles the uninflected N form *economy* rather than the expected English adjective *economic*, may be explained by merging of the final consonant with the following consonant /k/ in *crisis*.

Such bilingual production is not always fluid. In (262) the inserted PP is marked by a hesitation before the English noun *months*, which replaces the Tok Pisin noun *mun*:

(262) TiT#35b-203_L57

lo harim blo mi ***in uh last kapol of*** ah- ***mans*** i go

PREP hear-TR PREP.Poss 1sg PREP DET last couple PREP ah- months PM go

long harim bilong mi ***in the last couple of*** ah- ***months*** i go

'in my hearing in the last couple of ah- months past'

The entire prepositional phrase *in the last couple of months* appears here in English. The fact that it is followed by the Tok Pisin adverbial phrase *i go*, indicating anteriority, clearly distinguishes it from Alternation: the PP is inserted into the Tok Pisin matrix, and integrated by the effective duplication of the meaning of English *last*. The whole utterance that includes the inserted "island" is consequently well-formed according to the Tok Pisin grammatical frame.

A comparable underlying Tok Pisin sentence structure indicating anteriority and featuring prepositions with any NP could be as (263):

(263)

long [NP] long [NP] i go

Such a structure would produce something like (264) in this case:

(264)

long [harim bilong mi] long [tupela mun] i go

An outright switch, Alternation (Chapter 6), from Tok Pisin to English would be something like (265):

(265)

long harim bilong mi ***in the last couple of months***

Significantly, in the recorded example (262) the English preposition *in* governs the whole EL island, and the anteriority indicated in the English PP is doubled, with the phrase *i go*, by the Tok Pisin ML.

English verb phrases can also be inserted as islands, as in (266), where *transfer out of New Ireland* is marked as a predicate by Tok Pisin *i* and the sentence completed with a subordinated Tok Pisin VP *i no gat haus bilong slip*:

(266) TiT#35b-203_L434

twenti plas puhlisman i ***transfe aut of niu ailan # bikawz*** i no guh haus bilo silip

twenty plus policeman PM transfer out of New Ireland # because PM NEG have house PREP.POSS sleep

twenty plus policeman i ***transfer out of New Ireland, because*** i no gat haus bilong slip

'twenty plus policemen transferred out of New Ireland, because (they) had no sleeping quarters'

The Tok Pisin matrix frame here is as in (267):

(267)

[NP] PM [VP] CONJ PM [VP]

The example from the core corpus (266) fits the structure of (267) as in (268):

(268)

NP[***twenty plus policemen***] PMi VP[***transfer out of New Ireland***] CONJ***because*** PMi VP[no gat haus bilong slip]

In addition, (268) illustrates that the English conjunction *because* is grammatically active in the structure of (266):[71] it takes the place of a conventional Tok Pisin conjunction such as *long wanem* (s5.2.5) and subordiantes the Tok Pisin explanatory VP constituent *no gat haus bilong slip* to the English NP constituent *twenty plus policemen. Because* is an 'anglicised' late bridge system morpheme alternative to *long wanem* in this example of mixed language structured by the Tok Pisin grammatical 'matrix': as well as showing the Insertion of well-

[71] see also discussions above at s5.2.5 and s7.3.4.3.1

formed EL constituents as 'islands', (266) illustrates the interoperability of EL system morphology within the ML grammatical frame. This indicates the operation of congruent grammatical structure, discussed further in Chapter 8.

7.4 Conclusions

This chapter has demonstrated that much of 'anglicised' Tok Pisin can be characterised as Tok Pisin featuring code-switching with English of the Insertional type. Single-word 'borrowings' can be considered as Insertions, both integrated and only partially integrated, and, consequently, code-switching theory throws a fresh light on the appearance in Tok Pisin corpora of English phonemes and morphology.

The Insertion of integrated English constituents here as EL Islands also paves the way for an understanding of the development of idiomatic phrases, and the acquisition of aspects of English grammar through lemma-transfer, under the lemma-based lexicalist production model inherent in the MLF.

An analysis of the most common forms of mixed language in 'anglicised' or 'Urban' Tok Pisin as Alternational or Insertional code-switching has been provided. In the next chapter, a third type of code-switching is discussed, and its significance for the development of an intermediate grammar between Tok Pisin and English.

8 Findings 4: Complex Behaviour – the Prospect of Congruent Lexicalization

Chapter 5 considered aspects of the grammar of Tok Pisin and illustrated the similarities and differences between Tok Pisin and English. In particular, the presence in the Papua New Guinean dialect of English of phonological, morphological, and syntactic characteristics that were more similar to Tok Pisin than their equivalents in Standard English, and the development of structures in expanded and first language, 'creole', Tok Pisin that increasingly resembled English patterns, were noted. These examples showed that there were high levels of congruence between these two languages in Papua New Guinea. Chapters 6 and 7 showed that mixed language and 'anglicised' Tok Pisin could be analysed in terms of the interaction of two separate systems, demonstrating the presence and applications of two types of code-switching: Alternation and Insertion.

However, the corpus also includes some behaviour that is unusual from the point of view of these models, and cannot easily be analysed in terms of either switches from one language to another, or the introduction of constituents from one language to fit within the system of another. In addition, the presence of such frequent Alternation and Insertional code-switching suggests the presence of a third type of behaviour underpinning the language recorded in this corpus.

As discussed above (s3.3), Muysken (2000) provides accounts of both insertion and alternation as different types of language mixing. In cases where the languages in question show high levels of congruence in their grammars, he discusses a third type of behaviour, described as 'congruent lexicalization': mixed language based upon a shared grammatical system. This is distinct from the 'mixed codes' discussed in Auer (1998a) or the regularised 'fused lects' (FLs) of Auer's (1999) typology. In Congruent Lexicalization morpho-lexical

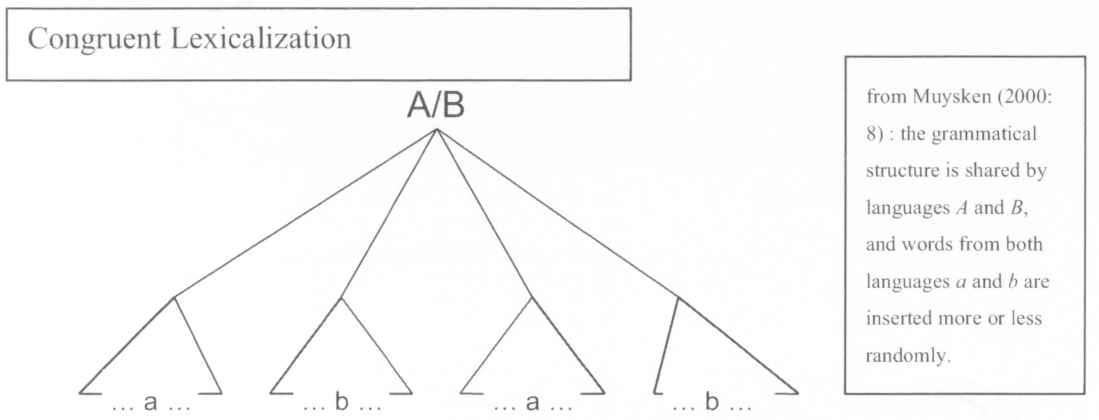
material from any of the languages participating in the mix can be used to fill out the constituents as required by the underlying mixed structure. Muysken (2000: Chapter 5) details some key indicators that Congruent Lexicalization is occurring.

This chapter presents an account of the applicability of this type of switching analysis to the situation in Papua New Guinea (s8.1), and provides an analysis of the data that shows that Muysken's indicators of Congruent Lexicalization are present in some Tok Pisin/English mixing, and that this type of code-switching provides an alternative account of some of the behaviour observed in previous chapters (s8.2). Following from this, the significance of these findings for the picture of language change in Tok Pisin and concepts of decreolization is discussed (s8.3).

8.1 Congruent Lexicalization and Variation

Where Alternation is concerned with the juxtaposition of two language systems (Chapter 6) and Insertion with the integration of material from one language into the system of another (Chapter 7), the third type of code-switching behaviour involves a convergence of the two systems. Structurally, Muysken (2000: Chapter 5) presents Congruent Lexicalization as occurring where material from either language is used to fill out a shared structure. An abstract representation of this structure is as in Figure 8.1

Figure 8.1 : Congruent Lexicalization - Structure (ref. Muysken 2000: 8)



Here there is no hierarchical relationship between a matrix language (ML) and an embedded language (EL). Instead, the languages share the grammatical structure of the sentence (A/B), and the major constituents are aligned in this shared structure, though the internal composition of constituents can be decided using a single grammar. The alternating (a,b,a,b) pattern illustrated in (273) is by no means the default: in some cases the structure may be entirely filled with words and morphemes from a single language (a,a,a,a), or a single morpheme from one language may appear in a grammatically active position in an utterance otherwise constructed using material from the other language (a,a,b,a). What distinguishes Congruent Lexicalization as a different type of language mixing from either Insertion or Alternation is that the underlying structure of the utterance represents a combination of requirements from both of the languages involved.

Muysken proposes that this structure has been developed by speakers 'ironing out' the differences between languages: they have established a new, congruent grammar, a single, shared system that underlies the mixed production. This is particularly frequent in bilingual communities, where structural convergence coincides with code-switching. Chapter 5 illustrated convergence between Tok Pisin and PNG English, and also defined some key areas of difference. We know from Census data (National Statistical Office 2002a; Paliwala 2012), discussed in s3.1.4, that there is a substantial community who are literate in both Tok Pisin and English, and levels of spoken language competence and exposure are predictably higher in this group. The discussion of findings in Chapters 5 and 6 has illustrated that the mixing of Tok Pisin and English is taking place in a range of patterns and contexts. The vital question, as Muysken observes is "Does convergence facilitate code-mixing or does code-mixing pave the way for convergence?" (Muysken 2000: 123). My discussion of the relationship between a creole and its superstrate poses the additional question: does convergence represent decreolization?

Muysken introduces his description of Congruent Lexicalization with examples of dialect switching. Different dialects of a language are usually based on the same grammatical structure, with phonetic and morphological differences predominating. The syntactic structure of different dialects is usually the same, and this allows for dialect switching to take place fluidly.

In his discussion, Muysken (2000) compares this code-switching analysis to earlier models of systemic variation in languages, arguing that such variation is in fact code-switching. While sociolinguistics makes a traditional distinction between 'style-shifting' as the expression of different options within a single system and 'code-switching' as activity involving two separate, invariant, systems, an understanding of the range of code-mixing behaviour begins to blur this distinction. Muysken's ideas provide a critical path into an understanding of variation in Tok Pisin through its dynamic relationship with English in the code-switching behaviour of bilinguals.

While variation studies under the Labovian paradigm look at the comparative frequencies of different variants in a language, patterns of covariation are not considered, and so the relationship between different elements is not considered. In research into code-mixing the linear relationship between elements viewed as being from different systems is the focus. While variation analyses assume a single variable system, code-switching analyses investigate the interaction of different systems.

In the case of pidgin and creole languages, these are radically different approaches to the same data: are the differences between 'Rural' and 'Urban' Tok Pisin, or between classical pidgin and expanded creole varieties, a result of variation within a single system (an assumption underlying prior analyses of corpora such as Verhaar's (1995)) or a product of the interaction of two separate systems (the influence of English described by G.Smith (2002) and earlier accounts of Anglicisation)? I have preferred a code-switching analysis for the

explicitly bilingual data of my corpus. However, individual uses of a particular code could, under a variationist analysis, be viewed as particular realisations of a particular variable.[72] For example, the production of the form [finiʃ] in place of an expected Tok Pisin marker of completion *pinis* [pinis] (269, below) could be viewed in one of two ways. Under a code-switching analysis as the lexical Insertion of the English stative verb *finish*; under a variationist analysis as the selection of two non-standard 'anglicised' (Laycock 1985b) phonemes (s5.1) from the 'expanded' Tok Pisin inventory (the variants /f/ for /p/ and /ʃ/ for /s/).

In general, studies of variation tend to focus on phonology, while those of code-switching focus on issues of morpho-syntax and discourse. Mühlhäusler's argument (1975) that code-switching between Tok Pisin and English indicated the two languages remained separate, was based on syntactic Insertion and Alternation, further exemplified in his UPNG examples (Mühlhäusler 1985a). Laycock's (1985b) account of the expanded Tok Pisin phonemic inventory is a model of variation: he does not investigate the connections between the sounds he records in Tok Pisin and other features of the speech acts that produced them.

The variationist paradigm precedes research into code-switching, which has shown that language mixtures that are individually unmarked and without apparent motivation are commonplace. While Alternations between systems may be highly and expressively marked, individual Insertions and many fluid switches may not be externally motivated, but may

---

[72] This approach was allowed for in the CLAN coding scheme, with variable:variant codes included on the %var tier of the database files. See Appendix III for details. As discussed in Chapter 4, a detailed analysis of variation across the database was not included as part of this project.

reflect the status of code-switching as the 'unmarked code' (Myers-Scotton 1993b) in a language community. In such contexts, meaning is transmitted through the selection of a particular code, just as it may be through the choice of a particular word (Gumperz and Hernandez-Chavez 1971a: 122), and it may be the frequency and timing of switches that are motivated and carry meanings. Poplack's work indicates that frequent switching bears meaning as a sign of social solidarity (Poplack 1980, 1981, 1987). In this, code-switching can resemble variable-selection in terms of motivation.

Patterns of covariation are hard to study (Muysken 2000: 126), and the consequent isolation of linguistic variables encourages a false distinction between variation and code-switching. In code-switching studies, the phonological form of a word can be understood as conditioning or triggering a subsequent switch (Clyne 1967, 1997), and G.Smith (2002: 205-209) discusses evidence from his data of hesitation and self-correction when his speakers used an English pronunciation on a word in their Tok Pisin. The appearance of such forms, viewed within the context of the whole discourse as switches under CS models, would be noted as the separate occurrence of a different variable under a variationist analysis. Congruent Lexicalization is not constrained in the same way as Alternation or Insertion, and so is a model for both complex mixed language and style-shifting or dialectal variation.

Under the MLF model (Myers-Scotton 1993a) the use of any functional elements from English in Tok Pisin, or vice-versa, is a clear sign that the mixing of the languages has gone beyond code-switching or Alternation to indicate the merging of two grammatical systems. As we have seen in the discussion of Insertion strategies in Chapter 7, the presence of functional elements alone is not enough to violate this model: these elements need to be active in the grammar, unsupported by any equivalents in the Matrix Language. The presence of peripheral elements, such as conjunctions and discourse markers, from a second language can be accounted for as a form of Alternation (Chapter 6), or are permitted under the 4M model (Myers-Scotton and Jake 2000, 2009). However, some examples of mixed language in my

corpus go beyond Insertion or Alternation, and provide evidence that a congruent grammatical framework underlies their production.

Due to the congruent features between PNG Enlgish and Tok Pisin observed in Chapter 5, Tok Pisin and English in Papua New Guinea are candidates for Congruent Lexicalization under Muysken's characterisation of the kind of bilingual settings in which it might be expected. In addition to the shared lexicon that is a common feature of a pidgin/superstrate pair, sound changes in PNG English have increased the number of homophonous words that can trigger a code mix. Both languages have, except in the formation of plurals, possessives and interrogatives, a general linear equivalence, and changes in the creole verb phrase have further increased similarities between the languages. There is also widespread categorical equivalence, except in the pronominal systems, where Tok Pisin distinguishes inclusive/exclusive and dual pronouns; while English has gendered 3rd person pronouns. In the roles of the predicate marker in Tok Pisin and the copula in English there exist grammatical elements specific to each language, though neither of these elements are mandatory for all speakers of these languages.

In the following section, examples are given from the core corpus of the types of mixed language behaviour Muysken characterises as indicative of Congruent Lexicalization. For at least some members of the national community, code-mixing appears to be based on a shared grammar, including categories from both English and Tok Pisin, moving one step beyond 'anglicised' Tok Pisin and towards a new, mixed, variety that could be considered intermediate between the two.

8.2 Examples of Tok Pisin/English Congruent Lexicalization

Congruent Lexicalization is predicated on and indicated by a range of features in Muysken's description (Muysken 2000). This section provides examples of these taken from my data.

8.2.1 Homophonous Diamorphs

Clyne (1967) suggested that the presence of homophonous diamorphs may facilitate code-switching. Where many language pairs share very few words with similar sounds and meanings, the history of English as the lexifier to Tok Pisin means that a large number of shared words exist in the standard vocabulary (Mühlhäusler 1979a). Sound changes in PNG English have increased homophony in their pronunciation for some speakers. A consequence of this is that words with similar sound shape are common in Tok Pisin and PNG English. Examples of these have been seen in Romaine's Unitech children's data (s3.4.3), and in many of the examples from the core corpus discussed above. Below, I discuss some of the more significant cases in my data.

Example (269) is a Tok Pisin utterance including what can be characterised as NP EL Islands from English *the pot* and *the man*. Between these, the Tok Pisin verb *pinis* is realised with an English form [finiʃ], showing a selection of /f/ over /p/ and the use of the fricative /ʃ/ word-finally in place of /s/.

(269) TiT#08-506_L81

so mi mi go rran ### kisim ***de pot*** anda # mi finish ***da man*** ia # mi go bak ***hom*** .

DISC.CONJ 1sg 1sg go run ### take-TR the pot above # 1sg finish the man LOC # 1sg go back home

so mi mi go ***ran***/ron, kisim ***the pot*** antap, mi ***finish***/pinis ***the man*** hia, mi go ***back***/bek ***home***

'so I I ran, took the pot there, dealt with that man, (and) I went back home'

The sentence *mi finish the man* is mixed, with the subject being the Tok Pisin 1sg pronoun *mi* and the rest of the sentence resembling English. However, the verb *finish* here is used with Tok Pisin semantics, meaning 'I dealt with the man', yet without transitive marking required by Tok Pisin grammar. Here, the similarity between pronunciations of English *finish* and Tok Pisin *pinis* characterises a mix that is not easily analysed as either Insertion or Alternation.[73]

Equally, the sentence *mi go back/bek home* is not easy to analyse as either language, since the pronunciation of the predicate places it indeterminately between English *go back* and the equivalent Tok Pisin *go bek*, while the pronoun subject is distinctively Tok Pisin, while *home* is English. The phonological congruence (s5.1) of Tok Pisin and PNG English and the interchangability of lexis and pronouns result in a particularly complex mixture, with the increased frequency of homophonous diamorphs at the centre of such congruence.

Example (270) contains a case in PNG English of a construction calqued from Tok Pisin:

(270) TiT#24b-15_L991

laik laik fuhr eksampuhl a ***if dhay wan fello go an*** ## ah- # ***wan marid man*** ah- ***wan marid wumehn an wuhn singuhl men*** a ?

like like for example a ***if they*** *one fellow* ***go and***, ah-, ***one married man*** ah- ***one married women and one single men*** a?

'like like for example eh if they go alone and, ah, one married man ah one married woman and one single man eh?'

Here *one fellow* reflects the etymology of the Tok Pisin numeral *wanpela* in its role as a determiner. The meaning 'if they go alone' is indicated by the flagged clarification *one married man*. In Tok Pisin, the selected *if they one fellow go* would be as in (271):

---

[73] We will return to this example later for additional indications of mixed language.

(271)

sapos em wanpela (i) go na ...
COND 3sg one-ADJ (PM) go CONJ …
'if they go alone and ...'

Examples such as this show Tok Pisin syntax and semantics applied to constructions apparently being formed in English. This is exactly what is predicted if the underlying structure is shared between the languages.

As observed earlier (s5.1), speakers show variation in the way they pronounce words in Tok Pisin and English that lead to the removal of distinctions established in standard orthographies.

In (272) a formal Tok Pisin phrase *tenk yu tru* 'thank you very much' is pronounced with /θ/ and with a low vowel /æ/, following standard English:

(272) TiT#35b-05_L77
thAnk yu trru ***suhr***
thank 2sg very sir
thank/tenk ***you***/yu tru ***sir***
'thank you very much sir'

In (273) only /θ/ is used, the Tok Pisin vowel form being retained:

(273) TiT#10b-05b_L334
thenk yu
thank/tenk ***you***/yu
'thank you'

The phonological variation observed in Chapter 5 allows for constructions such as (274), where the Tok Pisin pronoun precedes an English form of the verb *gat* 'have', possibly selected by the use of the English loan *idea*. The resulting construction *mi got idea* is not standard English, however, because the meaning of *got* is not past tense *get*, but Tok Pisin *gat* 'to have'. This present tense interpretation is supported by the repetition of the same statement using the traditional Tok Pisin word *driman* 'dream' that follows:

(274) TiT#35b-202_L400[74]

mi got ***aidia*** # mi gat drriman i stap

1sg have idea # 1sg have dream PM CONT

mi ***got***/gat ***idea***, mi gat driman i stap

'I have (an) idea, I have an idea'

The variation in the sound system that allows both *got* and *gat* in Tok Pisin and PNG English produces other homophonous diamorphs that can be seen to be involved in code-switching, as in (275) (also part of (182) in Chapter 7), where the modal *ken* is realised as [kæn]:

(275) TiT#10b-05b_L168

ol man kAn baim lo keke

PL man can buy-TR PREP eat

ol man ***can***/ken baim long kaikai

'men can buy to eat'

Similarities in the sound systems of PNG English and Tok Pisin lead to variation between the forms of words, so that they can resemble standard forms from either. Effectively, both pronunciations of *got/gat*, *kan/ken*, *this/dis* are available to some speakers in either language.

Consider the variation in consonants and phonemic reduction that produce a form of *dispela* as [θɛsɛla] in (276):

(276) TiT#35b-203_L181

bikos thesela bai ***ariv*** a- alvim tru long ***divelopmen*** blong kuhntri long ol ***rrural erias*** blo yumi

because this FUT arrive ah- help EMPH PREP development PREP country PREP PL rural area-PL PREP.POSS 1pl.INCL

because dispela/***this*** bai ***arrive*** a- halavim tru long ***development*** bilong kantri long ol ***rural areas*** bilong yumi

---

[74] This example was mentioned in Chapter 6 as a key example of repetition for translation purposes.

'because this will arrive ah- help very much with (the) development of (the) country in our rural areas'

In (277) the same speaker realises a less anglicised Tok Pisin pronunciation [disɛla] of the same words:

(277) TiT#35b-203_L177

ating -ah- desala # ol moni

I.think ah- this # PL money

ating ah- dispela, ol mani

'I think ah- these moneys'

This variation produces homophonous diamorphs that encourage constructions that would be analysed as either Tok Pisin or English, though with slightly different meanings. An example of this is (278), where [ɒl dis a] could be analysed as English *all this* with a Tok Pisin tag *a*, or as a switch to a reduced Tok Pisin *ol dispela:*

(278) TiT#24b-15_L1144

an den yu kaav yu kaav kunduus ***and*** ## garamuts ***an*** ol dis a ?

CONJ then 2sg carve 2sg carve Kundus CONJ ## Garamuts CONJ PL this Q.TAG

and then you carve you carve kundus ***and***, garamuts ***and*** ol dispela/***this*** a?

'and then you carve Kundus and, Garamuts and all this eh?'

The homophony here shows a congruence in this area of the grammar similar to Romaine (1992: 321) (s3.4.3) and G.Smith's (2002: 206-207) examples of constructions indicative of varieties indeterminate between English and Tok Pisin. Convergence of phonemic form coincides with convergence of meaning, preparing the ground for more complex mixing.

Homophony between the Tok Pisin plural marker *ol* and the English quantifier *all* encourages multiple glosses for constructions such as (279):

(279) TiT#24b-15_L208

an potin is ol ***wayl animuhls***

CONJ protein is PL wild animals

and protein is ol/***all wild animals***

'... and protein is wild animals'

The section [ɒl wajl animəls] could be glossed as a case of the Tok Pisin plural providing double-marking of the English plural NP 'wild animals', as in (279), or as an English quantifier, transforming the meaning to 'entirely wild animals', as in (280):

(280) TiT#24b-15_L208

an potin is ol wayl animuhls

CONJ protein is QUANT wild animals

and protein is all wild animals

'... and protein is entirely wild animals'

In (281) a similar utterance [ɒl a:ts] is amenable to an analysis as in (282):

(281) TiT#24b-15_L1124

yul luhn ol ***a:ts*** a ?

you.MOD learn PL art-PL Q.Tag

you'll learn ol ***arts a?***

'you will learn arts eh?'

(282)

yul leuhn ol a:ts a ?

you.MOD learn QUANT art-PL Q.Tag

you'll learn all arts a?

'you will learn all (the) arts eh?'

These examples illustrate how homophony in Tok Pisin and English can produce constructions that are open to different interpretations. The speaker in (279) and (281) indicates a pattern of marking English plural nouns with a pre-nominal [ɒl], which resembles Tok Pisin plural marking.

The homophony does not necessarily indicate that this is the meaning he intends, but it does indicate a fertile ground for code-mixing.

8.2.2 Morphological Integration

The discussion of Insertion in Chapter 7 provided an analysis of English morphemes on borrowed words that showed they were not grammatically active in the sentence. Tok Pisin has such limited morphology that there is little disjunction between the two systems. Example (270) showed the presence in PNG English of a form resembling the origin of Tok Pisin *wanpela*. In (283) the fusion of the verb *took* and its subject pronoun *him* in PNG English produces a form [tʊkɪm] that appears similar to a Tok Pisin transitive verb:

(283) TiT#08-507_L239

ba yu no hi tookim aldawe tu da beis of da mauntin

CONJ 2sg know 3sg.masc take.PAST-TR/take.3sg.masc all.DET.way PREP DET base PREP DET mountain

but you know he took-im/took him all-the-way to the base of the mountain

'but you know he took-him all the way to the base of the mountain'

The omission of the object of transitive verbs is an option in Tok Pisin (s5.2.1) and this may encourage the reanalysis of the *-im* suffix as a pronominal object. Here we see the inversion: a speaker pronouncing an English verb+object *took him* (V+Obj.Pron) as if it were a novel Tok Pisin transitive verb *took-im* (V-TR). Other examples are present in the data, as in (284), where a particular form [ɹəpɔtim] resembles a mixed constituent, English verb *report* and Tok Pisin transitive suffix *-im* as it does an English verb+object:

(284) TiT#08-506_L63

tell him de ai wan ruhpawtim nau

tell 3sg.masc(1) that 1sg report-TR/report.3sg.masc now

tell him that I want report-im/report-him now

'tell him that I want (to) report him now''

In (285) ['ʃɒw'him] takes the place of a transitive verb expected in Tok Pisin.

(285) TiT#08-507_L257

ol i go ## kam ***an show im spot***

3pl PM go # come CONJ show 3sg.masc ø spot

ol i go, kam ***and show him spot***

'they went and showed him the spot'

Here an analysis that sees *and show him spot* as an Alternation to English requires that we accept the absence of the determiner required for the nouns *spot* to be well-formed in English. This is possible in PNG English. However, (285) could equally be analysed as a Tok Pisin sentence with an English conjunction and well-integrated transitive verb, as in (286):

(286)

ol i go ## kam ***an showim spot***

3pl PM go # come CONJ show-TR spot

ol i go, kam ***and show-im spot***

'they went and showed (the) spot'

Homophony produces ambiguity, and allows for the appearance of forms resembling transitive marking in PNG English. Parallel interpretations of constructions such as (278) and (279) as reflecting both Tok Pisin and English grammars are examples of such convergence. Were constructions such as (286) on transitive verbs in PNG English to become consistent for an individual, or across the community, this would indicate a change in the requirements of the PNG English grammar in the direction of Tok Pisin.

Similarly, some nouns are marked with English determiners in Tok Pisin, with *the* as in (269) (discussed in s8.2.1 above), and also with *this* here in (287) (first introduced at s7.2.2 in the discussion of core loans as (194)):

(287) TiT#24b-15_L717

na sapo yu yusim ***duhs drress*** na wanem bilo narabla ***klan*** na narapla man ba ol i ***komplen*** yu

CONJ if 2sg use-TR DET dress CONJ whatever PREP.POSS another clan now another man FUT 3pl PM complain 2sg

na sapos yu yusim ***this dress*** na wanem bilong narapela ***clan*** na narapela man bai ol i ***complain*** yu

'if you use this dress and whatever of another clan then another man will make a complaint (against) you'

Along with, (269) above, (287) could be an example of English EL Det+N islands: *the man*, *the pot*, and *this dress* are all well-formed English constituents inserted into a Tok Pisin frame. However, as indicated in Chapter 5 and in Chapter 7, there is a suggestive similarity between the PNG English pronunciation of *this* as [dis] and the reduced, bi-syllabic, form of the Tok Pisin determiner *dispela* [disla] is suggestive. When English-form determiners appear in Tok Pisin optionally marking English nouns, as in (269) and (287), the distinction between the two grammars is further eroded.

The deletion of the determiner in (284) was permitted by PNG English, but the corpus also contains English-form determiners marking Tok Pisin nouns in English sentences, as in (288):

(288) TiT#24b-15_L148

an in dat haus tambaran ## **in** dat haus tambara ***# you don hev tu wuhk***

CONJ PREP DET house Tambaran ## PREP DET house Tambaran # 2sg do.NEG have PREP work

and in that haus tambaran, ***in that*** haus tambaran, ***you don't have to work***

'and in that spirit house, in that spirit house, you do not have to work'

Here, (288) shows that *haus tambaran* 'spirit house' is a well-integrated Insertion into the English system, being a traditional Tok Pisin phrase marked by English determiners. The fluid Insertions of (269) and (287) indicate a role for English-pattern determiners in some speakers' Tok Pisin. In (269) both of the marked nouns *pot* or *man* are well established in Tok Pisin, and so the Det+N constructions there are arguably not simple cases of Insertion. The equivalency-based substitution of English morphemes for conservative Tok Pisin forms was illustrated for (287) in Chapter 8 (194-195), where Myers-Scotton's categorisation of determiners as an 'early system morpheme' under the 4M exceptions to the MLF model was also indicated (s.7.3.4.3). Such flexible Insertions also fit a Congruent Lexicalization analysis.

At (289) is an example of a Tok Pisin determiner *narapela* being introduced into an English PP *in a narapela house* in place of the expected *another*, with homophony throughout the

sentence, on every constituent except the preposition *in* and determiner *a*, conditioning its role:

(289) TiT#35b-203_L42

...o in a narapela # ***haus bifaw*** ...

CONJ PREP DET other # house before

o/***or in a*** narapela, haus/***house before***/bipo

'... or in a another, house before ...'

These examples illustrate how congruence between Tok Pisin and PNG English encourages complex mixing, mixing that appears to be based on an interchangability of lexical material within a common grammatical framework.

8.2.3 Multi-Constituent Code Mixing

The mixing of multiple constituents within a single utterance is identified by Muysken as another indicator of Congruent Lexicalization. The core corpus data contains many examples, illustrated by (290) and (291), where sequences of noun phrases from English are inserted into Tok Pisin:

(290) TiT#10b-07_L364

evan kondision lo ***human welbein*** blo yumi go luk

even condition PREP human wellbeing PREP.POSS 1pl.INCL go look

even condition long ***human wellbeing*** bilong yumi go luk

'look into even (the) condition of our human wellbeing'

(291) TiT#10b-07_L516

ol i mekim samela kain ***longtem desisin*** we # ***filti yiuhs laita*** ## ol lain pikin blo yumi kan ***veluat***im ***eksakli*** # wanem kain ***ledesip eksuhsais*** ol lain li:duh bilo yu mekim

3pl PM make-TR some-ADJ kind long.term decision where # fifty years later ### PL group children PREP.POSS 1pl.INCL can evaluate-TR exactly # what kind leadership exercise PL group leader PREP.POSS 2sg make-TRANS

ol i mekim sampela kain/***kind long-term decision*** we/***where, fifty years later***, ol lain pikinini bilong yumi ***can***/ken ***evaluate***-im ***exactly***, wanem kain/***kind leadership exercise*** ol lain ***leader*** bilong yu mekim

'they make some kind (of) long-term decision where, fifty years later, our children can evaluate exactly, what kind (of) leadership exercise your leaders make'

The intensity with which constituents from English are introduced into Tok Pisin in such mixed language indicates a very close relationship between the underlying grammars. It is the frequency of the Insertions and the conditioning of the switches by homophonous diamorphs that, following Muysken's (2000) account, indicate the presence of Congruent Lexicalization.

8.2.4 Non-Constituent and Bi-Directional Switching

In contrast with the fluid Insertion of (290) and (291), however, are examples of non-constituent mixing that would not be expected under Insertional or Alternational models of mixed language. Frequently these are indicated by bi-directional mixtures, showing easy substitution of equivalent items from each language. In (292) *when* is used as a temporal marker in PNG English, and in (293) it is present in the same role in Tok Pisin.

(292) TiT#24b-15_L587

... ***wen yu wena finis doin at*** haus tambaran ...

TIME.ADV 2sg TIME.ADV.COP finish do-CONT DET house Tambaran

... ***when you when-are finish doing that*** haus tambaran ...

'when you when are finished doing that haus tambaran'

(293) TiT#24b-15_L431

... but wen yu antap lo wiwik ...

CONJ TIME.ADV 2sg there PREP Wewak

... ***but when you***/yu antap long Wewak ...

'but when you are in Wewak'

Conversely, in (294) the same speaker uses *taim* (the Tok Pisin equivalent of English *when*), in PNG English congruently in the same position:

(294) TiT#24b-15_L899

so taim ***yu kam autsait ## yu not gona rrelai on sambodi els ...***

DISC.CONJ TIME.ADV 2sg come outside ## 2sg NEG go-CONT.PREP rely PREP somebody else

so taim yu/***you*** kam/***come outside***/ausaid, yu/***you not gonna rely on somebody else*** ...

'so when you come outside, you (are) not going to rely on somebody else ...''

In (294) it is only *taim* that is clearly Tok Pisin, and the first phrase could in fact be either Tok Pisin or English: phrase structure and word order in both languages is congruent with only lexical choice and pronunciation differentiating them. Indeed (294) could represent either a straight switch from anglicised Tok Pisin (preposed with the English discourse marker *so* and characterised by the English consonant cluster /ts/) *taim yu kam ausaid* to English *you not gonna rely on somebody else*, or an English sentence with the Tok Pisin specifier replacing English *when*. In fact, (294) illustrates the characteristic form of Muysken's third 'type' of language mixing: the lexicalization of congruent grammatical structure using words from either language system.

Taken together (292)-(294) indicate a tendency for this speaker to use material from either language to fill the specifier position: English *when* and Tok Pisin *taim* are interchangeable.

8.2.5 Back and Forth Switching

Where free mixing takes place in either direction there is reason to expect that the underlying structure agrees. In (295) is a switch from Tok Pisin to English that resembles Alternation, and a switch to a distinctive Tok Pisin verb from the English:

(295) TiT#08-506_L158

wen tupla go ren antap lon haus ***dring mai biuh an*** kaikai

TIME.ADV two 3sg go round up PREP house drink 1sg beer CONJ eat

when tupela go raun antap long haus/***house drink my beer and*** kaikai

'when they two go about, up to (the) house drink my beer and eat'

In (296) (also (221) in Chapter 7, above) an English two-word verb appears following the preposition, and is itself followed by the Tok Pisin determiner *sampela* before the sentence concludes in English:

(296) TiT#10b-07_L443

yumi ## mas lukluk lo ***seping of*** sampuhluh ***longtem kaunsel wuhk***

1pl.INCL ## must look PREP shaping of some long.term council work

yumi, mas/***must*** lukluk long ***shaping of*** sampela ***long-term council work***

'we, must look into (the) shaping of some long-term council work'

In both these examples there has been a changeover of matrix language within the utterance, and in (296) this has taken place within a prepositional phrase. Such intimate mixing suggests that a single matrix language frame is not in operation, that both languages are actively involved in structuring the sentence.

This pattern is similarly evident in (297), which contains the construction discussed above at (289):

(297) TiT#35b-203_L41

mi laik tok olsem lo sampela taim i go pinis ***bifaw mai taim in poluhtiks*** lo disela hau:s # ***o in a*** narapela # ***haus bifaw:*** # ol i bin toktok lo ***amalgamaishuhn*** lo ami: na polis # o bungim ami: na polis wantim .

1sg like talk same PREP DET time PM go ComA TIME.past 1sg.POSS time PREP politics PREP DET house # CONJ Prep DET another # house TIME.past #

3pl PM PAST talk PREP amalgamation PREP army CONJ police # CONJ meet-TR army CONJ police with

mi laik tok olsem long sampela taim i go pinis bipo/***before my time in politics*** long dispela haus/***house***, o/***or in a*** narapela, haus/***house before***/bipo, ol i bin toktok long ***amalgamation*** long ami na plis/***police***, o bungim ami na plis/***police*** wantaim

'I would like to say: at some time in the past before my time in politics in this house, or in a another, house before, they talked of amalgamation of (the) army and (the) police, or bringing (the) army and (the) police together'

There are repetitions here, following an Alternational pattern, but the phrase *in a narapela house before* is conjoined to *long dispela haus/house*, which is itself part of the NP *my time in politics*. The complexity of this mixture is not amenable to analysis as Insertions or Alternations, but is a mixture of constituents drawn evenly from both the languages involved.

In the context of widespread Insertion of English content, as in (298), back and forth switching takes place with ease:

(298) TiT#10b-07_L478[75]

ol ***lida*** bilo yumi i gat sambla kain ***risuhn*** # lo ol ***pipol*** bilong ol oa # ol i lai ***takim*** samela kain ***souht*** o ***souhtem mesuhs*** lo # luki ***veri kwik*** # thuh resals ova thuh neks faiv yias taim ol i stap lo ***palimant***

ol ***leader*** bilong yumi i gat sampela kain ***reason***, long ol ***people*** bilong ol o, ol i laik **take**-im sampela kain ***short*** o ***short-term measures*** long, lukim ***very quick, the results over the next five years' time***/taim ol i stap long ***parliament***

'... our leaders have got some reason, from their people or, or electing in some short or short-term measures to, see very quick, the results over the next five years when they are in Parliament'

The last VP here represents a switch to English following the preposition *long*, and the shared homophone [taɪm] *taim/time*, Tok Pisin 'when' English 'time', triggers a switch within the phrase back to Tok Pisin. These switches are neither simple Insertion or Alternation, but

[75] Part of this utterance has been discussed separately in Chapter 7 as (177)

represent a free combination of constituents based on homophony and a high level of congruence. Such behaviour is a characteristic of Congruent Lexicalization.

8.2.6 'Function Words': System Morphemes and a Congruent System

Under the MLF and other models of Insertional switching, active function (system) morphemes can only come from a single language except where they trigger an EL island. This is the system morpheme principle of the MLF model (Myers-Scotton 1993a). In Congruent Lexicalization, however, functional material can be drawn from either language to fit the shared structure. This provides an explanation for the presence of typical system morphemes such as pronouns and conjunctions that is an alternative to arguments for peripheral Alternation (Chapter 6 and s7.3.4.3.1) and morpheme-class distinctions such as the 4M model[76] (Chapter 7 and s3.2.3.4). It also explains the presence of switched prepositions in mixed language production.

As observed above (s5.2.3, s7.3.4.3.2), we find pronouns switched between Tok Pisin and English in the data. These can be utilized to convey different distinctions – such as gender, single/dual, or inclusive/exclusive – to those available from the pronominal system of the main language of the utterance.

In (299) the English 3sg masculine pronoun *him* is used in subject position, following a preposed possessive noun phrase object:

---

[76] On this model and its account of exceptions to the MLF see Myers-Scotton (Myers-Scotton and Jake 2000; Myers-Scotton 2002; Myers-Scotton and Jake 2009) and the discussions noted above.

(299) TiT#10b-05b_L223

siuht blong en ***im*** no wasim long ***twenni yiuhs***

shirt PREP.POSS 3sg 3sg.masc NEG wash-TR PREP twenty year-PL

shirt/siot bilong em ***him*** no wasim long ***twenty years***

'he hasn't washed his shirt in twenty years'

As observed, Tok Pisin 1sg pronouns are often used in PNG English by some speakers. Example (299) indicates a reanalysis of English *him* as an optional Tok Pisin pronoun, used contrastively here for emphasis by a speaker for whom *en*, *em*, and *im* 'him' are considered equal alternatives.

In (300) the Tok Pisin 1sg pronoun is used in an English sentence, along with back-and-forth switching indicative of Congruent Lexicalization:

(300) TiT#08-507_L399

mi ***wawk aut bikos dis kidz hu o- in*** aus wantim Dino ol kam bek nau ***an*** wok lo singaut na

1sg walk out CONJ DET kid-PL o- PREP house CONJ Dino 3pl come back REL CONJ work PREP call.out DISC

mi ***walk out because these kids who*** o- ***in house***/haus wantaim Dino ol kam/***come back***/bek nau/***now and*** wok/***work*** long singaut nau

'I walk(ed) out because these kids o- in (the) house with Dino they came back now and the four of them called out now'

In (301) (also discussed as (217) and (248) in s7.3) an English preposition is used in Tok Pisin, in violation of constraints on Insertional code-switching under the MLF and 4M models:

(301) TiT#10b-05b_L146

at taim yu lik ***autsaids kondisuhn*** ya # em no gudpla

PREP time 2sg look outside-POSS condition EMP # 3sg NEG good-ADJ

at time/taim yu/***you*** luk(luk)/***look outside's condition*** hia, em no gudpela

'when you see (the) outside's condition there, it is not good'

The preposition *at* is used here exactly as Tok Pisin *long* would be, as in the construction *long taim* 'when'. Note too that possession on the NP *outside's condition* is marked in the English pattern, indicating that this NP is following English grammatical rules without any support

from Tok Pisin possessive structures. The introduction of English lexical material in a functional role in Tok Pisin indicates that the grammatical framework is not language specific: there is no matrix language.

Also in the data is evidence of the copula in Tok Pisin utterances, with English *be* in use in (302):

(302) TiT#08-506_L147

klan blo mi ***is*** tokim yu: longtaim

Clan PREP.POSS 1sg is talk-TR 2sg long-time

clan bilong mi ***is*** tokim yu longtaim

'my clan is telling you many times ...'

This cannot be characterised as an Insertion, and the copula appears here precisely in its role in English grammar: joining the possessive NP and the verb. The presence of the copula here indicates that both languages are involved in structuring this sentence.

Further evidence for a congruent grammar underlying the production of some of the data comes from the presence of Tok Pisin functional morphology in PNG English. Consider the constructions underlined in (303-305):

(303) TiT#08-507_L114

yu hev tu weit uhntil dhe perent of dhe gai hi kuhmz to get dere de chail

2sg have PREP wait until DET parent PREP DET guy 3sg.masc/PM come-CONT PREP get there DET child

you have to wait until the parent of the guy he/i comes to get there the child

'you have to wait until the parent of the guy he comes to get there the child'

(304) TiT#08-506_L172

na ***dis pesen*** ### ***hi*** kam na ***tuk***im em

CONJ this person ### 3sg.MASC/PM come CONJ took-TR 3sg

na ***this person***, ***he***/i kam/***come*** na ***took***-im em

'and this person, he/0 came and took him'

(305) TiT#08-507_L328

an den it wosn long aftuh dat dat insiden hi hapen

CONJ then 3sg was.NEG long after PRON DET incident 3sg.MASC/PM happen-Ø

and then it wasn't long after that that incident he/i happened

'and then it was not long after that that (this) incident happened'

In each of these a word resembling the English 3sg masculine pronoun *he* is present between the subject and the verb in the underlined VPs. In (303) the sentence is otherwise entirely English, and the pronoun is reduplicated redundantly. In (304) the subject is an English NP, but the VP is well-formed, according to the Tok Pisin matrix, if [hi] is taken to be a variant form of the predicate marker *i*. In (305) there can be no question of a standard English process of pronoun reduplication because the Subject NP *that incident* is abstract and ungendered. In all three of these examples *he* is used in a position that marks the predicate in a manner consistent with the conventional role of Tok Pisin *i*. These examples are of particular interest in the context of Keesing's argument on the origins of *i* as a subject-referencing pronoun (Keesing 1988), which is the role of *he* in the examples above.

The examples of active EL system morphemes in mixed Tok Pisin/PNG English speech, represent remarkable findings: illustrating the interchangeable usage of pronouns and prepositions in Tok Pisin and PNG English, and suggesting a congruent grammatical frame allowing active roles for determiners, the copula, and predicate marking in either language.

8.2.7 Violation of some Constraints on Code-switching

In addition to the unusual mixing observed above, my corpus also includes evidence of some direct violations of constraints proposed on code-switching behaviour by DiSciullo, Muysken, and Singh (1986). Muysken (2000: Chapter 5) has argued that such counter-examples are indicative of a congruent structure into which material from either language can be inserted.

The first case is switching between the subject pronoun and the main verb. Examples of this can be seen above in (299), between an English pronoun and Tok Pisin verb in Tok Pisin sentences, and in (300) between the Tok Pisin 1sg pronoun and an English verb in an English sentence. In (306) the verb phrase is inflected with Tok Pisin habitual marker *save* and the transitive *-im* suffix, but the pronoun is English *we*:

(306) TiT#10b-07_L234

ol i kisim athin ***neks dei*** o # samthin ***eventu*** ***wi*** sa ***kondaks***im

3pl PM take-TR perhaps next day or # something eventually we MOD.HAB contact-TR

ol i kisim ating ***next day or***/o, samting/***something eventually we*** save ***contact***-im

'they got (it) I think (the) next day or, something eventually we will catch up with (them)'

Indeed, examples of switching between subject pronouns and main verbs are found throughout our dataset, and occur in many of the examples discussed here.

Example (307) illustrates similar switching in a Tok Pisin sentence, between the English 1sg pronoun *I* and the Tok Pisin verb *bekim* 'to answer/respond':

(307) TiT#10b-05b _L309.

yu go ***an*** ***ai*** bekuhm wai yu mekim pundaun ah se suhmptin aselum asd aas # o pike o samting ?

2sg go CONJ 1sg answer Q.Tag 2sg make-TR go.down PL DET something same ash ash # CONJ gum CONJ something

yu go ***and I*** bekim wai yu mekim pundaun ol dispela samting olsem as as, o PK o samting?

'you go (and do something) and I call out 'why are you throwing things like ash or chewing gum or suchlike on the ground''

Such patterns of pronoun/verb switching are specifically associated by Muysken with Congruent Lexicalization.

There are also examples of switching between the verb and the negator. In (308) (also (119) above) the verb *undermine* is given a past tense meaning by the Tok Pisin auxiliary *bin*. It is

stripped of English morphology and thoroughly integrated into the Tok Pisin frame. Yet the negator *never* is from English:

(308) TiT#35b-02_L171
paluhment mu^s ***neva*** bin ***u^nduhmain***
parliament MOD.obl NEG PAST undermine
parliament mas/***must never*** bin/***been undermine(d)***
'parliament must not be undermined'

As in other examples above, reanalysis by the speaker of *bin* here on the basis of its etymological associations with English *be* may be a factor increasing the congruence of these constructions, and enabling this kind of lexical mixing. Certainly, this passive voice construction is novel in Tok Pisin, and is dependet on congruence with PNG English.

In (309) (introduced in s5.2.1 as (52) to illustrate the presence of the Tok Pisin transitive verbal suffix in PNG English sentences) a Tok Pisin verbal suffix marks *kolim*, but it is negated by English *never*:

(309) TiT#35b-04_L75
yu: toktok pinis ***an hi neva kor***im ***korom***
2sg talk COMP CONJ 3sg.masc NEG call-TR quorum
you/yu toktok pinis ***and he never call-***im ***quorum***
'you gave a speech and he never called quorum'

This is also an example of switching between the pronoun and the main verb.

Another proposed constraint (DiSciullo, Muysken and Singh 1986) is on switching between the subject and the copula. This takes place in (302) above, with a switch between the Tok Pisin possessive noun phrase subject *klan bilong mi* and the English copula present there. Another suggestive example is (310) where Tok Pisin *wara* 'river' (in this usage) is juxtaposed with the English copula *is*:

(310) TiT#08-507_L215
nau wara ***is*** # yu: now # ***firs*** lo ***dis said***
DISC water COP # you know # first PREP DET side

now/nau wara/***water is***, yu/***you know, first*** long ***this side***

'now water is, you know, first on this side'

Appearing alone, (310) might not be convincing evidence, due to the similarities in semantics and pronunciation possible between PNG English *water* and Tok Pisin *wara*. Indeed, alternative analyses of several of the examples in this section could be made that propose that the verbs in (283-285) and the verbs *contactim* and *callim* in (306) and (309) are actually English. However, these verbs are produced with Tok Pisin morphology and are well-formed in that language, and not so in English. Similarly, the noun *klan* in (302) could be proposed as English, yet it is integrated into a Tok Pisin possessive construction. So too following the phonology of PNG English *wara* in (310) could be considered an English form, albeit reflecting Tok Pisin semantics, since it is used to indicate a river in this context, but this speaker has deliberately chosen this form, and the rest of the utterance is in Tok Pisin.

Taken together, the evidence from our dataset indicates that the constraints proposed by DiSciullo et. al. (1986) are not active in these cases. These counter-examples provide a persuasive argument for a form of language mixing that is without constraint, which Muysken suggests is only made possible by an underlying congruent grammar.

8.3 Congruent Lexicalization in a Creole/Superstrate pair

While not all the features of Congruent Lexicalization as described by Muysken (2000) have been identified in my data, this corpus does include a range of indicators that a congruent grammatical structure explains some speech acts by some speakers in Papua New Guinea.

In the Papua New Guinean dialect of English, Tok Pisin bilinguals have a language system whose phonology is highly congruent with Tok Pisin. Both languages have limited morphology. In 'creole' Tok Pisin structures showing congruence with English are found (see Chapter 5). This chapter has shown how homophonous diamorphs play a role in triggering code-switching in the corpus data, and how the frequency and complexity of some mixed Tok

Pisin and English language is not amenable to analysis as Alternation or Insertion. Function words are switched between the languages, and a number of proposed constraints on code-switching are violated. In particular, there is some evidence of the use of the copula in Tok Pisin, and of an analogous form of Tok Pisin predicate marking in PNG English.

While Tok Pisin and English are highly related languages, there are significant differences between them. Congruent Lexicalization, as proposed by Muysken, is a form of language mixing that underlies style shifting and dialect switching among closely related varieties, and switching between highly congruent languages. In the data there is evidence of behaviour that indicates that both languages are able to provide material equally when Papua New Guineans speak. Individual switches are not motivated by local meaning-making, and there is no clear hierarchy between the languages.

Theories of decreolization (Bickerton 1975b, 1975a) suggest that the prolonged interaction between Tok Pisin and English in Papua New Guinea will lead to the development of a creole continuum and a merging of Tok Pisin with English. Among the community of people multilingual with Tok Pisin and English, code-switching does not always seem to depend on separate systems or a hierarchical relationship. Instead there is evidence for an exchange of material between the languages, and I have presented data in this section that shows some highly complex mixing behaviour.

With copula deletion a feature of PNG English, and predicate marking a variable feature of Tok Pisin, evidence of these grammatical roles underlying production in either language provides a first indication of an intermediate grammar developing. As the way people speak these languages becomes more similar, and as their tendency to mix them freely increases, it may yet be possible to find conclusive evidence of a mesolectal grammar.

It is in the speech acts of bilinguals such as my speakers that this grammar will develop. Congruent Lexicalization is, conceptually speaking, evidence of the merging of codes. There

will need to be compromise in areas of conflict, but between Tok Pisin and PNG English these are few. The emergence of indications of a congruent grammar underlying bilingual speech is a particularly fascinating feature of this dataset, because the theoretical concepts of pidginization, creolization, and decreolization presume that structural development in Tok Pisin will ultimately reflect interference from English. The concept of shared bilingual grammars allows us to understand how structural convergence develops.

The data from my corpus indicate, however, that this structural development will not be one-way. Tok Pisin and English multilingualism is not only encouraging the Tok Pisin grammar to be more congruent with English, but it is also adjusting the English spoken by Papua New Guineans to accommodate Tok Pisin. The congruent grammar evolves not through the development (expansion or reduction) of one language but through a fluid comparison of both, and it is this bilingual process that may produce the mesolect.

Additional work and data collection will be necessary before it is possible to conclude that the observed patterns are becoming widely established. The core corpus data was recorded from adult speakers a decade after G.Smith's schoolchildren's creole study. Tok Pisin and English in Papua New Guinea are in a continuous dynamic relationship. With substantial knowledge and ability in English spreading throughout the national community in PNG it is probable that any eventual continuum will be patterned through bilinguals. To what extent their innovations will be carried through into the Tok Pisin of monolinguals is yet to be seen.

The findings of three different types of code-switching behaviour provide not only evidence for the separation of Tok Pisin and English, as Mühlhäusler argued (s2.6), but conversely also evidence that this behaviour may be bringing them closer together, a convergence that may pave the way for a continuum. Only by appreciating the tri-partite characterisation of CS behaviour outlined here can we unpack the interaction of separate grammatical systems from

the indications of a new, intermediate system underlying some speech acts. Further research is necessary before any conclusive findings on a continuum can be made.

9 Findings 5 : Speaker's Perceptions – the UPNG Responses

The findings of Chapters 5, 6, 7, and 8 have shown that English and Tok Pisin exist in a complex and intimate relationship in Papua New Guinea. However, vital questions remain about the significance and salience of these different types of behaviour in the language community.

This chapter moves from the previous discussion of structural types and meaning-making in mixed language to sociolinguistic issues, presenting issues of language ideology alongside analyses of speakers' quantitative responses to samples of mixed language. In this way it balances the structural findings with findings of both qualitative and quantitative feedback from bilingual speakers. It shows that some different types of mixing between Tok Pisin and English are readily recognised, while others are less salient, and discusses the significance of the less-salient language mixing for our picture of the relationship between Tok Pisin and English.

Initially, in order to provide some additional context to the findings presented so far, some comments from Papua New Guineans at the University of Papua New Guinea (UPNG) on the relationship between Tok Pisin and English are introduced alongside my own observations of behaviour. This offers some context: indicating the attitudes underlying the speech behaviour (s9.1). Field journals, containing many notes from interviews and discussions, are the data source.

The main part of this chapter (s9.2) discusses a pilot 'Listening Exercise', '*Skelim mix lo tok Inglis na tok Pisin',* conducted at UPNG in 2003 as a novel attempt to get quantitative feedback on the mixed language discussed in the chapters above. The aim of the exercise and the experimental method is described (s9.2.1) and three different analyses of the data

collected from UPNG students are presented (s9.2.2) to illustrate the value of this approach. The first of these looks at the range and pattern of responses to specific key samples (s9.2.2.1). The second compares the two groups' average scores for the samples, and considers whether differences between those responses-scores are significant (s9.2.2.2). Most interesting are the findings of the third set of analyses (s9.2.2.3), which orders the samples from 'most English' to 'most Tok Pisin' by their mode and median average of scores and discusses the place of each sample on that ordered scale in terms of their linguistic features.

The final section discusses the extent to which this novel approach to getting feedback on mixed language from speakers throws light on their perceptions of types of CS and the development of mixed varieties of Tok Pisin (s9.3).

9.1 Diglossia and Language Choice

As discussed in Chapters 2, 3 and 6, early descriptions of diglossia in the Papua New Guinean language community are not strictly applicable today. Nonetheless, some comments on languages associated with particular contexts of use, or with particular topics, can be made.

The situation described in the work of G.Sankoff (1980d) and in Kulick's study (Kulick 1992) is typical: where a community retains a Tok Ples, this is assumed as the community language, with Tok Pisin featuring as a language of wider communication. Kulick also provides a record of language shift from Taiap to Tok Pisin in Gapun, explained by the cultural associations of the different languages and by a process of generational change. During my fieldwork, speakers in semi-urban villages often commented on the necessity of making an effort to teach the younger generations their *tokples*, and so avoid it being taken over by Tok Pisin, and there is a national consciousness of the importance of retaining traditional languages.

In urban areas, however, two factors combine to reduce the currency and usage of *tokples*. Primarily, the diversity of *tokples* languages and the overall usage of Tok Pisin as the

language of wider communication means that, except where language groups form close, insular, settlements, *tokples* is rarely spoken. Even in Port Moresby and the Motuan villages of the Papuan region children are increasingly speaking Tok Pisin in place of Motuan dialects, or Hiri Motu, despite the latter's status as a national language. Secondarily, the great proportion of children growing up in urban areas are born there and operate principally in Tok Pisin up to the point at which they begin to attend school, when English is taught as the language of wider communication. This encourages a shift in the status of Tok Pisin in urban areas: from a language of out-group communication to the language of in-group communication.

Various adult speakers observed that in town they would use English and Tok Pisin, but in the village they would speak Tok Pisin and Tok Ples. For the most educated, English was reported to be a language they would use in town. For people without the benefit of an extended English-education, they would use Tok Pisin.

While signs in school remind pupils that English must be used in the classroom (Smith, G. 2002: 31), schoolchildren in Port Moresby, Madang, and Vanimo all described using Tok Pisin consistently except when directly addressing the teacher, and several noted that the teachers would sometimes use Tok Pisin in class if they felt they needed to. An older speaker described to me the negative consequences of using *tokples* in school:

(311) TiT#25b-T@3.30

na ***edukeisuhn*** tu yu tokples long # skul yu ***panis*** .

... na ***education*** tu yu tokples long, skul yu ***punish***

'also [when you are in] education if you speak indigenous language at school you are punished'

At the University of Papua New Guinea I observed classes where tutors would use Tok Pisin in their introductory comments, or to bring the class to order, or to provide a humorous comment or aside. However, in multilingual PNG, *tokples* do not generally have the kind of range that enables them to play a significant part in urban life.

Speakers interviewed about the role of Tok Pisin in school were often concerned that its use would negatively impact on the development of English skills. In this their attitudes were nothing new, and echoed concerns widely discussed at the time of Independence. However, the persistence of such concerns and their spread beyond academic contexts to the wider community is notable.

Tok Pisin/English mixing in PNG is prolific, and is recognised as a characteristic of certain speakers. This reduces the value of Tok Pisin as a language of wider communication. Wurm's chapter (1985) highlights exactly this issue, while Nekitel quotes a gardener at UPNG:

> Mipela i no save wanem ol memba blong Palimen o Haus Asembli i toktok long em – ol save yusim Inglish na abusim a miks nabaut na mipela i paul
>
> *'We do not know what the members of Parliament or House of Assembly were talking about. They usually use English and mix that with Tokpisin thus leaving us very confused'*
>
> (Nekitel 1998: 153) (reformatted)[77]

A series of interviews with language academics and local language experts at UPNG indicated that code-switching was perceived as commonplace, both functionally, in response to factors such as topic, speakers, environment, and audience (Nekitel 2000: pers. comm.), and in order to make Tok Pisin "more productive" (Thomas 2000: pers. comm.).

Dicks Thomas (2000: pers. comm.) identified CS as occurring not only to incorporate new terminology for new cultural domains ('cultural loans' s7.2.1), but also as a result of the "cultural load" when people do not realise it (for example, the 'core loans' of s7.2.2). His

---

[77] Nekitel, as a native speaker of Tok Pisin, several *tokples* and a fluent speaker of English discusses the presence of English in the Parliament extensively, though primarily applying the methods of discourse analysis. Following his indication of established Tok Pisin political terminology, no language formatting is applied for the cultural loans *parliament* or *house assembly* here, though I do so elsewhere (see s5.4 and s7.2 for discussion).

estimate was that 10-12% of Tok Pisin words had developed, post-Independence, to express more specific meanings, or were new borrowings for new concepts. Alongside this process of borrowing from English, however, Thomas observed an acceptance of standards in Tok Pisin spelling, and an increasingly wide use of Tok Pisin, including by NGOs working in PNG.

Nicholas Faraclas (2000: pers. comm.) observed that English-language was used by only a low percentage of Papua New Guineans either formally, in education or the armed forces, for example, or informally.

Kenneth Sumbuk (2002: pers. comm.) explained that for English bilinguals, the ability to joke in Tok Pisin remains an important in-group value, with side-comments, encouragements, and familiar remarks all commonly expressed in Tok Pisin. He also noted the possibility of generational differences in the patterns and perceptions of language combinations.

In the family groups I worked with, parents were observed using Tok Pisin with the youngest children, and Tok Pisin and English interchangeably with older children. *Tok ples* languages were used sporadically, if at all. Often, discourse markers or conjunctions from Tok Pisin were used in English speech, just as English discourse markers like *so* and *I mean* were used in Tok Pisin. The effect was of consistently indexing a bilingual identity (s3.2.4.3).

For the post-independence urban generation, it is commonplace for Tok Pisin or English to be used in almost any situation, with language choice largely dependent on the identity of the person being spoken to. With a substantial population (>1.2 million) (National Statistical Office 2002a, 2002b; Paliwala 2011, 2012) of Tok Pisin/English bilinguals in the country today, the data presented in earlier chapters shows that choosing to use both languages is often an option.

9.2 Establishing Support for a Hierarchy of Salience

My findings establish that a range of code-switching behaviour takes place between Tok Pisin and English in PNG. This data does not permit broad statements about the role of socio-economic or generational factors, or specific details of educational attainment, in patterning the distribution of types of code-switching in the community as a whole (s4.2.4.1, see also s10.4). However, the analysis presented so far makes a significant contribution to understanding and appreciation of the range of types of behaviour involved in 'Urban' or 'anglicised' Tok Pisin, and their increasing impact on the development of the language.

Following on from the data collection and analyses of code-switching characteristics in Tok Pisin/English speech, a further question presented itself. The approach to the field had been founded on the principle of participant observation (see Chapter 4), of creating and sustaining intimacy with speakers and their context, yet the processes of transcription, coding, and analysis took place for the most part at the University of Sydney campus, at a remove of some sixteen hundred miles from Papua New Guinea. The generous and detailed assistance of the Port Moresby speakers in reviewing transcripts, and of the University of Papua New Guinea staff and students in considering conclusions drawn from the subsequent theoretically-informed analyses had been an integral part of the project. However, no matter how successfully a fieldworker integrates with a community or consults with local experts, the assessments and judgements of an abstracted science remain a peculiarity without some form of additional corroboration from an entirely 'insider' perspective.

This question of corroboration was, in the context of my study, how native speakers of mixed Tok Pisin and English varieties themselves assessed the kind of feature-filled mixed language that had been found in abundance in Papua New Guinea. Many people in PNG had spoken with me about a 'mix' of Tok Pisin and English, and commented on its significance in social interaction and differentiation, but, in the light of the varied types of CS found through my

analysis I was interested to discover how people might evaluate these different types of mixed language.

Specifically: might people identify the linguistic discontinuity of a switch from Tok Pisin to English, language Alternation, as making an utterance sound 'less' like Tok Pisin than the more tight incorporation of English words and phrases of 'island'-type Insertion, and what might be made of language containing indications of Congruent Lexicalization?

The Listening Exercise experiment '*Skelim mix lo tok Inglis na tok Pisin'* was designed to collect quantitative data on the perception of mixed language by Papua New Guineans. The comparative analyses of the collected respondents' scores of different selected samples of naturally occurring language investigated patterns in the way UPNG students from different disciplines evaluated mixed language. The aim was to assess the connection between the way informants had scored different samples and the linguistic features of those samples in terms of 'types' of code-switching. Informants' scores were taken to indicate the distinctions between types of mixed language in the samples of speech they heard during the exercise, rather than judgements about the speakers of those samples, in line with the experimental design (s9.2.1 and also see s9.2.2.1). The initial trial of the exercise, described below, produced data that indicates the validity of this approach. It establishes a methodology for getting quantitative feedback on types of mixed language and their salience for speakers.

9.2.1 Method

To investigate this question of the way people might differently perceive mixed language, and to consider what, if any, impact particular linguistic features have on people's evaluation of language as being 'more' or 'less' English or Tok Pisin, I created in the field a simple listening exercise '*Skelim Mix lo tok Inglis na tok Pisin'* in which participants were invited to listen to a series of short audio samples extracted from the substantial core corpus dataset. Each participant was asked to rate the language of each sample by giving it a score on a scale

from '1' to '10'. The objective was to collect quantitative 'scores' for each sample with which different linguistic analyses could be compared. Subsequent analyses would consider any indication of relationships between the 'score' of a sample and the type of CS it contained in relationship to other samples with similar and different types of CS.

Working in the field this experiment was made as simple as possible. Essentially, this was an adaptation of the informant response gathering of Lambert et al.'s 'Matched Guise' experiments (Lambert, Giles and Picard 1975) with a number of significant differences.

Primary among these was the nature of the research question: Lambert et al's methodology was designed to investigate the social meanings of different language varieties by collecting informants' attitudes to speakers who used these varieties in terms of stigma and prestige. This was done by asking informants to rate a speaker in terms of various 'moods', using a sliding scale, for example: from 'reliable' to 'unreliable'. Other researchers, for example Bender (2004) extended this technique to investigate the social associations of specific linguistic features. In contrast, I intended to gauge the perception of different types of mixed language, and used a single scale: informants were asked to rate the language they heard, rather than the speaker.

A second difference was in the nature of the samples and the experimental design. The Matched Guise methodology uses specially recorded examples of speech specifically composed to include particular features of interest. Significantly, in investigating the sociolinguistic evaluation of the varieties/features presented through informants' assessments of the character of the speakers, the Matched Guise method uses the same speakers to record different variant samples. This step means that the true identity of the speaker will be the same for different samples of speech containing different speakers, and so the informants' assessments of 'their' character will be based on the nature of the language itself. Informants are not made aware that differing samples may have been produced by the same speaker. In

my experiment, however, I did not use specially recorded speech samples. Instead I selected from my field recordings 25 short sections that included what I had analysed as different types of language mixing. One sample (samples 2 & 11) was repeated as a check against the consistency of responses. This design focussed on the research question as to how Tok Pisin/English speakers would assess the language of the samples as opposed to the identity of the sample speakers.

In some cases the samples were from the same speaker. However, informants in my experiment were directed to grade the language they heard, and not the speakers, in terms of its 'mix': true Tok Pisin was to be given a score of 10, true English a score of 1, a half-half mix a score of 5, and other numbers on the scale used to grade any intermediate mixtures.[78] As a result, the quantitative data collected was intended to assess the language directly, rather than any social values.

In the workshops where this experiment was conducted there was a period of discussion before and after the samples were played and scored. Before the experiment informants were encouraged to talk about their background, the languages they used, and their impressions of the national language culture in Papua New Guinea. In this way a general awareness of language variation was established. All the informants expressed a keen awareness of variation in Tok Pisin, and largely defined it in terms of mixing with English. After the experiment informants were given the opportunity to discuss any impressions they had formed about the speakers whose speech samples they had heard. While scoring each sample,

---

[78] The terms *tru tru Tok Pisin*, *tru tru Tok Inglis*, and (*hap hap*) *mix* were Tok Pisin phrases often used by Papua New Guineans when discussing varieties of language.

some informants had made notes on the identity of the speakers they were hearing: their regional background, educational level, or job. However, these impressions were not consistently recorded by all informants, and where discussed below are only introduced as background to the analysis. It is worth observing that many informants were able to identify particular radio personalities or national politicians where their speech was included in the samples.

Following on from widespread community perceptions that some speakers', for example politicians', varieties of Tok Pisin contain a level of English that makes them unintelligible to non-English speakers, the aim of the experiment was to make a survey of how some people evaluated the samples' linguistic content: Do some types of language mixing have a more salient impact on peoples' perceptions that a given utterance is 'in' a particular language?

What follows is an account of the exercise as it was conducted with UPNG students in 2003. The different linguistic features identified in Chapters 4-7 are discussed in a characterisation of the samples (s9.2.1.1). The composition of 2 groups of informants is outlined (s9.2.1.2). The informants' responses to the samples are then presented in quantitative terms (s9.2.2) considering: the spread of responses to each sample (s9.2.2.1); some apparent differences between the median averages of the two student groups' scores of the samples (s9.2.2.2); and, most significantly, connections between the types of CS in each sample and the mode and median of scores given to those samples (s9.2.2.3). These quantitative analyses of responses show patterns that relate to the types of CS in the samples.

In the concluding section (s9.3), the extent to which the UPNG informants' responses to the audio samples indicates a hierarchy of salience for different types of Tok Pisin and English mixing is discussed, with some conclusions on the significance of different types of mixing for the students' assessment of language.

9.2.1.1 Selection of speech samples

From the field recordings of conversation, monologues, political speeches, and radio broadcasts that comprised the corpus data, a selection of twenty-five samples of spoken mixed language, featuring the voices of eighteen Papua New Guinean speakers and one non-National English speaking researcher, were transferred to cassette tape. The mean length of these samples was 22.11 seconds, but they ranged from 6 to 39 seconds in length (sd 8.72). The median length was 20 seconds.

The samples were selected as showing a range of mixed language. Alternation (Chapter 6) was present in six samples. A further six samples showed patterns that could be associated with Congruent Lexicalization (Chapter 8). Samples involving Insertion were chosen that showed the inclusion of content morphemes (s7.3.1), discourse particles (s5.4.1), system (function) morphemes (s7.3.4.3), mixed constituents V-TR (s7.3.4) e.g. *tenderim,* and English Adj+N pairs (s7.3.2) e.g. *leadership exercise*, and larger islands (s7.3.5) e.g. *up until now*, for a total of fourteen samples. Several samples featuring Insertion included different types of inserted material.

Three typical examples will be referred to throughout this chapter:

9.2.1.1.1 Example of Alternation: Sample 1

Sample 1 (312) includes a single clear marked Alternation from Tok Pisin to English to focus attention on a question (s6.3.1.5):

(312) TiT#25b@5.55 = **sample 1**

yu yu yu olsem yu stap ... long disela . y yu stap insaid steret lo disela kain ***stadi*** a ? au yu thing ols anem , ***ts gud tu go bek an len tokpisin in klasrum o tokpisin s oredi dhe, ts mater of yusing it*** ?

yu yu yu olsem yu stap, long dispel, y- yu stap insaid stret long dispela kain ***study*** a? hau yu ting olsem wanem, ***it's good to go back and learn Tok Pisin in classroom or Tok Pisin is already there, it's a matter of using it***?

'you you you are you are, in this, you are involved in this kind of study eh? How do you think: it's good to go back and learn, Tok Pisin in classroom or Tok Pisin is already there, it's a matter of using it?'

9.2.1.1.2 Example of Insertion: Sample 9

Sample 9 (313) includes clear Insertions from English, featuring cultural nouns (s7.2.1), Adj+N combinations, (s7.3.2) mixed V-TR constituents (s7.2.3.2 and s7.3.4)), and double-marked plural NPs (s7.3.4.1):

(313) TiT#10b-07 @ 01.28 = **sample 9**

SSS: ***las wik*** mi bin arim long, ol i gavman ol i
putim isela kain lo blong olim ol ***prais*** lo ol samting
ia . ... beain ***aselum*** o samthin . ...
nau lo tinting blo mi gavaman ogata taim a thin sa se
yu gavman, ol i got nau wantaim ol ***veri siotem***
plans a lo ***teigim*** samuhla kain ah& .

UET: ***sotkat*** .

SSS: ***priemtiv mesuhs*** lo ol ***problems*** yumi gat .

last week mi bin harim long, ol i gavman ol i putim dispel kain lo bilong holim ol ***price*** long ol samting hia, behain ***asylum*** o samting. Nau long tingting bilong mi gavman olgeta taim ating save olsem yu gavman, ol i gat nau wantaim ol ***very short-term plans*** a? long ***take***-im sampela kain ah-

short-cut

pre-emptive measures long ol ***problems*** yumi gat.

'last week I heard the government made a law to fix the price of goods' ... 'something to do with asylum' ... 'now my thought is that government have got very short-term plans eh, in taking some kind of ah-...'
'shortcut'
'...pre-emptive measures for our problems'

9.2.1.1.3 Example of Congruent Lexicalization: Sample 22

Sample 22 (314) from the same speaker, exhibits similar Insertions, but also the English discourse conjunction *and also* in prominent position, substantial idiomatic phrases as EL islands (s7.3.5.2), mixed constituents (s7.3.4), and rapid back and forth switching (s8.2.5) involving active English grammatical morphology (s8.2.6). It was included as an example of behaviour indicating Congruent Lexicalization (Chapter 8).

(314) TiT#10b-07 @2.20 = **sample 22**

athin ***dhuh oul liduhsip eksesais*** blo ol ***lida*** blo yumi lo ***Parluhmuhn*** nau ya .

ating ***yeah*** yumi gat pinis a em wok long ol lo Giad Ga kamuhpim gudpla sindaun blo man lo ples ia ***evuhn kondisuhn*** lo ***yuman welbeing*** blo yumi luk .
ol man na meri blong isuhla kuhnri .
n a tin ***whateva da plein on da kard*** ol mas lulu lo ***longtem va den***, ah& ***an olsouh*** ol ma:s uh& traim ol ting samla ***proaktiv mezuhs*** .
la ***wotim*** ol ***problems*** umi gat e umi , ***aksperimentim*** na umi ***ekspiriensim*** awgeta ***laas twenifaiv yiuhs*** .
ol ***probuhms*** yumi gat athing ***its uhbuht taim*** yumi # mas lukluk lo ***seipin ov*** sampla ***lontem # kaunsil wuhk den yu wen se*** .

ating/***I think the whole leadership exercise*** bilong ol ***leader*** bilong yumi long Palmen/***Parliament*** nau hia, ating ***yes*** yumi gat pinis a? em wok long ol long Giad-Ga- kamapim gudpela sindaun bilong man long ples hia ***even condition*** bilong ***human wellbeing*** bilong yumi luk.
ol man na meri bilong dispela kantri.
na ating ***whatever the playing on the card*** ol mas lukluk long ***longterm rather than***, ah- ***and also*** ol mas uh- traim ol ting sampela ***proactive measures***
laik ***vote***-im ol ***problems*** yumi gat and yumi ***experiment***-im na yumi ***experience***-im olgeta ***last twenty-five years***.
ol ***problems*** yumi gat ating/***I think it's about time*** yumi mas lukluk long ***shaping of*** sampela ***longterm, council work then you when say***.

'I think the whole leadership exercise of our leaders in Parliament today, I think yes, those we have now eh? they work for those Giad- Ga- increase the wellbeing of the rural population and oversee even our condition of human wellbeing (for) this country's men and women.
And I think whatever the game they are playing/is on the cards they must look to (the) longterm rather than, ah- and also they must uh- think of some proactive measures, as to decide and solve (those) problems we have and we experimented and we experienced throughout (the) last twenty-five years.
(Those) problems we have I think it's about time we must look into (the) shaping of some longterm, council work then you will see.'

9.2.1.1.4 Examples of 'English' and 'Tok Pisin'

Characterising the samples by 'matrix' language, six samples were in English. Four of these included Alternation to Tok Pisin and of those two (samples 6, 17) showed some characteristics of Congruent Lexicalization, another (sample 5) Tok Pisin discourse marking with the question tag *a* on the English part and the Insertion of English nouns and verbs in the Tok Pisin part, and the fourth (sample 26) Alternation for repetition of a sentence. Of the remaining two English samples, one (sample 14) showed the influence of Tok Pisin in the use of the Tok Pisin verb *stap*, while the other (sample 20) (315) a non-standard construction *for many times* characteristic of PNG English:

(315) TiT#35b-02 = **sample 20**

dhi:s pipuhl hav fo meni taims #### trai tu krusifai dhuh liiduhs uhv dhis neisuhn dhuh paluhmuhn mas asuhet its othoriti an bring dem bifo duh privluhdgis komiti: ## weduh it was duh intensin tu unduhmain palamuhnt !

these people have for many times, try to crucify the leaders of this nation the parliament must assert its authority and bring them before the privileges committee, whether it was the intention to undermine parliament !

'These people have tried many times to crucify the leaders of this nation, the parliament must assert its authority and bring them before the Privileges Committee, (to discover) whether it was their intention to undermine Parliament!'

Of the twenty samples primarily in Tok Pisin, two (sample 2/11) were actually the same sample, included as a repetition to assess whether the sample order might affect responses.[79] Four samples (samples 1, 10, 22, 23) featured Alternations. Of these, once again, one (sample 1 (312)) featured a single Alternation, two (samples 10, 23) included English features in the

[79] This duplication of sample 2/11 brought the total number of samples provided to informants participating in the exercise to 26.

Tok Pisin part and Tok Pisin islands inserted in the English part. This kind of complex mixing was a stronger feature of the fourth sample (sample 22 (314)), which was selected specifically as showing multiple types of mixing and indications of Congruent Lexicalization.

One short Tok Pisin sample (sample 21) was selected as showing the Insertion of English function morphemes and patterns of Congruent Lexicalization, and its language shared similarities with the Tok Pisin of one of the Alternation examples (sample 10).

The remaining fifteen samples were Tok Pisin featuring different types and combinations of Insertions. Seven of these featured English language islands and also included English cultural loans as single-word nouns. Of these, six (samples 2/11, 3, 4, 7, 12, 18) also included Adj+N phrases from English. Sample 24 (316) features cultural loans and a single idiomatic prepositional phrase *up in government*, and was included as an example of cultural/idiomatic 'anglicised' Tok Pisin:

(316) TiT#35b-202 @01.20 = **sample 24**

mi ting olosem bai PapuaNuGini # bai fait lo narapla ***independen*** .
blo wanem mogetuh ***stouhs*** ## olgeta liklik samtin oguhtuh pi:puhl blo narapla ausaid i kan kisim ## .
nau olosem ### mi laik # tok olosem # mi sori tumas lo lukim #### .
a tin plani man ol i toktok pinis long ol ***benks*** # blong yumi lo insaid lo PapuNuGini .
yumi ***uhp in guhvman*** yu wok lo kirapim awlguhtuh kainkaind ***inityushin*** long # halefim ol pi:pol ## tasol yumi n puhl inuhp lo halivim #### hau m uhlipl i halivim #### .
eh ! no got ***ansa*** blong en .

mi ting olsem bai Papua New Guinea bai paitim long narapela ***independence*** . long wanem olgeta ***stores***, olgeta liklik samting olgeta pipel/***people*** bilong narapela autsaid/***outside*** i ken kisim . nau olsem, mi laik, tok olsem, mi sori tumas long lukim . ating planti man ol i toktok pinis long ol ***banks***, bilong yumi long insaid long Papua New Guinea. yumi ***up in government*** yu wok long kirapim olgeta kainkain ***institution*** long, halavim ol pipel/***people***, tasol yumi no pilim inap long halavim, hau mipela i halavim ? eh! no gat ***ansa*** bilong en .

'I think that Papua New Guinea will fight for another Independence, because all the stores, all little things will be taken by other people from overseas. So now, I want, (to) say, I (am) very sorry to see (this). I think many people have already spoken about our banks in Papua New Guinea. We up in government work to start all types of institution to help people, but we don't feel able to help, how (can) we help? Eh! There's no answer for that.'

In addition, two (samples 4, 12) also included English cultural loans as single-word verbs. Five of the Tok Pisin samples (samples 8, 9 (314), 2/11, 15, 16) contained mixed constituent transitive verbs (V-TR) from English roots with the Tok Pisin transitive *-im* suffix, as well as cultural loan nouns. Four of these also included cultural adjective+noun (Adj+N) combinations. One (sample 15) included an English adverb. A further two samples (samples 13, 25) included English discourse markers and also features of Congruent Lexicalization. The final sample (sample 19) was characterised by English verbs: a single-word verb as a cultural loan and also a full VP inserted as an island.

The samples are listed by number in Table 9.1 alongside these observations on the matrix language, the types of CS they include, and the range of constituents involved in CS. The mode and median response scores (as discussed in s9.2.2.3 below) are also given. The full text of the samples can be found in Appendix VII.

Together these samples represented the full range of code-switching types I had identified, with both English and Tok Pisin matrix languages. The most English were sample 14 (the researcher in conversation) and sample 20 (a national politician speaking to Parliament in English), though both showed non-standard English features. The sample set was weighted towards Tok Pisin language, and the Tok Pisin included was in every case inclusive of at least one or two words that would have been characterised as English pre-Independence. The most uniquely Tok Pisin sample was sample 18, taken from a Yumi FM radio news broadcast. Consequently, no samples were included that were expected to receive an uncontroversial 'pure' language score of 1 (for English) or 10 (for Tok Pisin) from all informants.

Table 9.1 : Table of numbered Listening Exercise samples with 'average' scores, language, and CS types

						CS constituent types				
Sample Number	**Median Score**	**IQR**	**Mode Score**	**"Matrix" language**	**CS type**	**Content morphemes**	**System morphemes**	**Mixed Constituents**	**Islands**	**Alternations**
1	5	0.75	5	TP	Alt	1 Cultural N				1 to English
2	7.5	4	10	TP	Ins	1 Cultural N		1 integrated V (translated)	3 Adj+N	
3	7	3	7	TP	Ins	4 Cultural N (1 repeated 3 times); 1 Core N			1 Adj+N; 1 Two-Word V	
4	8	2	8	TP	Ins	3 Cultural N; 2 Core V, Numbers	Disc.Tag; Plurals		2 Adj+N	
5	3	2.25	4	PNGE	Alt / Ins	TP : 1 Cultural N; 1 Core V	PNGE: Qtag, Neg; TP: Disc.Tag;		1 Discourse	2 : Tok Pisin sentence with Insertions, and back
6	5	3	5	PNGE / TP	Alt / CLx		PNGE : Plurals (dbl marked); Q.Tag; Conj ; TP : Conj			4 : back and forth
7	8	2	9	TP	Ins	3 Cultural N; 1 Pronoun	Disc.Tag; plurals		1 EL Island VP; 2 Adj+N	
8	7	3	9	TP	Ins	2 Cultural N; 2 Core N		1 integrated V; 1 integrated Two-Word V	1 Adj+N,; 1 EL island (idiom) PP; 1 EL Island NP	
9	8	2.5	10	TP	Ins	1 Cultural N; 1 Core N	Plurals (dbl marked)	1 integrated V	1 Adv.Time; 1 EL island AdjP; 1 Adj+N	
10	7	2.75	5	TP / PNGE	Clx / Alt		PNGE : Conj; TP : Prep			4 : back and forth
11	8	2.5	10	TP	Ins	1 Cultural N		1 integrated V (translated)	3 Adj+N	
12	7	2.5	7	TP	Ins / CLx	2 Cultural N; 2 Core N (1 repeated)	Conj; Adv		3 Adj+N (1 repeated); 1 VP; 1 PP; 1 NP	2 : flagged PP and back
13	9	2	10	TP	Ins	1 Cultural N	Conj; Pron; Cop; Adv		1 Disc	
14	1	1	1	Eng	CLx	1 Core V				
15	6	3	6	TP	Ins	1 Cultural N; 1 Adv		2 integrated V	1 Pron+V (discourse)	
16	6.5	2	6	TP	Ins / CLx	1 Core N	Conj.Disc	1 mixed VP; 3 integrated V	1 Adj+N	1 : ragged
17	5.5	2.75	5	PNGE	Alt / CLx		Pron			5 : back and forth
18	9.5	1.5	10	TP	Ins				1 EL Island NP (repeated); 1 Adj+N; 1 Proper N (repeated)	
19	9	1	9	TP	Ins	1 cultural V			1 EL Island VP - idiom	
20	1	0	1	PNGE	CLx					
21	8	1.5	8	TP	CLx	3 Cultural N; 1 Core N; 1 core V (unintegrated)	Pron; Conj;Plurals (dbl marked)		1 CondP	3 : back and forth (incl. ragged CondP)
22	6	2.125	6	TP	CLx / Ins	1 Core N	Plurals; Conj; Adv	3 integrated V	3 Adj+N; 1 Two-Word V; 1 Adv+N; 3 EL Island NP	5 : back and forth (ragged)
23	6	2	7	TP / PNGE	Alt / Ins	TP = 1 Cultural N			PNGE = Insertion TP island	3 : back and forth
24	9	1	9	TP	Ins	6 Cultural N	Plurals (dbl marked)		1 EL Island PP - idiom	
25	8	2.75	8	TP	Ins / CLx	4 Cultural N; 1 core N; 1 V-Cont	Conj; Prep; Spec; Cont		1 EL? Island VP (CLx)	
26	3.5	3.5	1	PNGE	Alt					2 : repetition and back

9.2.1.2 Informant Groups

The participants in this experiment were students at UPNG in 2003. In addition to being convenient to access,[80] this community was selected as a consequence of Mühlhäusler's observations of code-switching in the community of students attending UPNG in the 1970s and early 1980s (Mühlhäusler 1985d).

These students could not be said to represent the majority of Papua New Guineans. On the contrary: they could best be characterised as the youngest cohorts of the educated élite. However, they represented an accessible and responsive group of informants for this pilot exercise, and their responses could be used as a benchmark against future administrations of the exercise. In addition, selecting the informants from contemporary students satisfied a number of criteria.

Firstly, this group of informants were all well-educated in English and fluent in Tok Pisin. They were part of a multilingual university community within which traditional vernaculars are a more highly marked choice of language than either Tok Pisin or English, and within which code-switching was an accepted part of daily communication.

Secondly, the students were part of the growing demographic of speakers of Tok Pisin: the post-Independence generation of people largely educated in English and accustomed to using Tok Pisin as a language of wider communication.

---

[80] At the time I designed the experiment I was working in Waigani, Port Moresby, and in close contact with staff and students at UPNG.

Thirdly, undergraduate students in particular could be considered part of the same diachronic community of youngsters studied by Suzanne Romaine (Romaine and Wright 1987; Romaine 1992, 1993, 1994, 1999) and Geoff Smith (Smith, G. 1990, 2000, 2002). As such, their assessments could be considered representative of developing norms in the national community.

Fourthly, access to groups of students and their consent to participate was straightforward, and they were easy informants to trial the format of the experiment with, as well as to provide feedback on the exercise. Volunteers responded to a notice put up in the university buildings, or by word of mouth, and participated in the listening exercise in sessions arranged outside of class time.

Finally, comparison could be drawn between students on the basis of their area of specialisation, and so the experiment was conducted with two groups of informants. Group 1 was a self-selecting group of ten male Economics and Law students between the age of 20 and 23 years old. One was from a highlands province, two from the Papuan region, two from the Sepik region, and five from Morobe and Madang. Group 2 was a volunteer group of thirteen (6 female and 7 male) students from the Language and Literature department aged between 20 and 29 years old. Two were from highlands provinces, three from the Papuan Region, two from Manus, one from New Britain, one from the Western Province, one from the Sepik Region, and three did not give their region of origin.

The additional research question that could therefore be addressed was: would students studying economics or the law (Group 1) give scores that could be assessed as significantly different to language and literature focused students (Group 2). That is: did the subject of study impact on informant's responses?

9.2.1.3 Administering the Exercise

With each group the session was held on campus but outside of scheduled classes, and began by sharing basic personal information about age, background, education, and language experience. We then continued with a group discussion about language issues, and I made note of observations on individuals' opinions that might affect assessments of mixed language. A key question asked of the groups was what they thought the role of Tok Pisin was in their community and in the nation as a whole. The second core question was what the informants thought the relationship was, if any, between Tok Pisin and English in the lives of Papua New Guineans. The groups were specifically directed to the concept of the "mix" between Tok Pisin and English that some people had said took place. I explained that this project was concerned with understanding what this meant and what kinds and quantities of mixing were taking place. Their responses informed aspects of the discussion of community impressions of language mixing (Chapter 1 and s9.1).

The central part of the sessions introduced the listening-and-assessment exercise '*Skelim mix lo tok Inglis na tok Pisin'* with a pre-recorded audio passage on the cassette (see Appendix VI). Paper and pens were distributed to the informants, onto which they wrote their name, age, and place of origin. It was explained that every sample would be played twice and that they could ask to listen to the sample as many times as they liked. Time would be given for people to write down their 'score', ratings of the language of the sample, before the group moved onto the next sample of audio. Informants were asked to write a single number between 1 and 10 next to each sample on their sheet to indicate how they rated the language of the sample: these numbers were the sample response scores. In this experiment, the number "1" was to indicate 'true' English, and "10" to indicate 'true' Tok Pisin. It was explained that all the intermediate numbers were allowed for people to indicate different amounts of mixing between the two languages that they might hear in the samples. Informants were requested to

make their own answers and not to be tempted to discuss the samples or to copy other peoples' scores. It was explained that there would be time for a group discussion at the end of the exercise.

The score sheets were collected following the exercise and prior to further discussion.

9.2.1.4 Data Analysis Techniques

The raw sample response score data from the listening exercise was pooled and compared, with the ordinal score data subjected to a series of basic statistical analyses. The intention was to investigate patterns in the feedback and see if this could be related to types of CS in the primary data. The sample size and experimental design did not support rigorous statistical analyses.

All the individual scores were entered in Excel spread sheets, along with the informants' personal information and any qualitative observations they had made on their papers.[81]

From this score data frequency tables were created and used to produce bar charts for each sample (see discussion s9.2.2.1 below).

The data was then assessed to identify the mode score for each sample, the median scores, the quartile figures, and the interquartile range. These were calculated for each of the two groups of informants separately, and for all informants' responses as a whole. A range of visual

---

[81] One informant had inverted the scale and used "10" for English and "1" for Tok Pisin, so their responses were re-entered to reflect the standard scale. Also, four informants gave some samples scores with decimal values (e.g. '5.5'). The original values were retained when finding the mode and median average scores (s9.2.2.2 and s9.2.2.3) to reflect the intermediate responses those informants intended. These figures were not included on the frequency tables or the bar charts discussed in s9.2.2.1 as they did not fit into the integer-value bins.

figures and tabulated reports were developed and are used to illustrate the comparative analyses presented below.

These analyses revealed agreement by the informants on the most Tok Pisin and most English samples, which matched the sample selection (s9.2.1.1.4). The analyses also gave clear values for some other samples (s9.2.2.1). However there was greater variation in responses to some types of mixing. The median and mode scores of the two groups were compared, and showed some differences, these differences, though interesting, were not found to be statistically significant (s9.2.2.2). All responses to the different samples were assessed, and general groups of samples by mode and by median response were considered (s9.2.2.3). The range of scores for individual samples was considered by calculating the median with the interquartile values.

These patterns are discussed below for what they indicate about the salience for these informants of different forms of mixed language behaviour (s9.2.2) and for issues regarding the development of language more generally (s9.3).

9.2.2 Responses to the Listening Exercise

The 'qualitative' component of the exercise data was limited to descriptive responses to the samples: informants were free to comment on individual samples or the set as a whole. This data was consequently unstructured, though it did provide some useful insights into the language attitudes of these students. In Group 1, the Economics and Law students, only one informant commented on the language of the samples, observing that the samples were mostly pidgin with some English, with the English used to "emphasise" the speech. In Group 2, the Language and Literature students, seven out of thirteen provided comments on the quality and nature of the language in the samples. One drew attention to the increased quantity of English in what they perceived as "formal" speech. Several others drew attention to the role of English borrowings as shorter ways of expressing cultural concepts that would require careful

explanation in Tok Pisin. Generally, it was agreed that for speakers of Tok Pisin unfamiliar with English most of these samples would not have been well understood. However, two students indicated that English language skills were commonplace in towns, with one from Manus estimating that 90% of city households had English as a first language at home. A second informant from Manus suggested that social change was leading to changes in local languages, and that even in Tok Pisin mixed language was now the norm.

While these comments are not statistically correct or generalisable, and are presented here as indicative rather than structured qualitative data, they do provide some support for the insights G.Smith drew from his own work on Manus (Smith, G. 1994, 1998), and the consensus of the group suggested that these observations applied across the home regions of these students.

The following sections consider the quantitative responses in detail.

9.2.2.1 Quantitative Responses to the Listening Exercise: Frequency

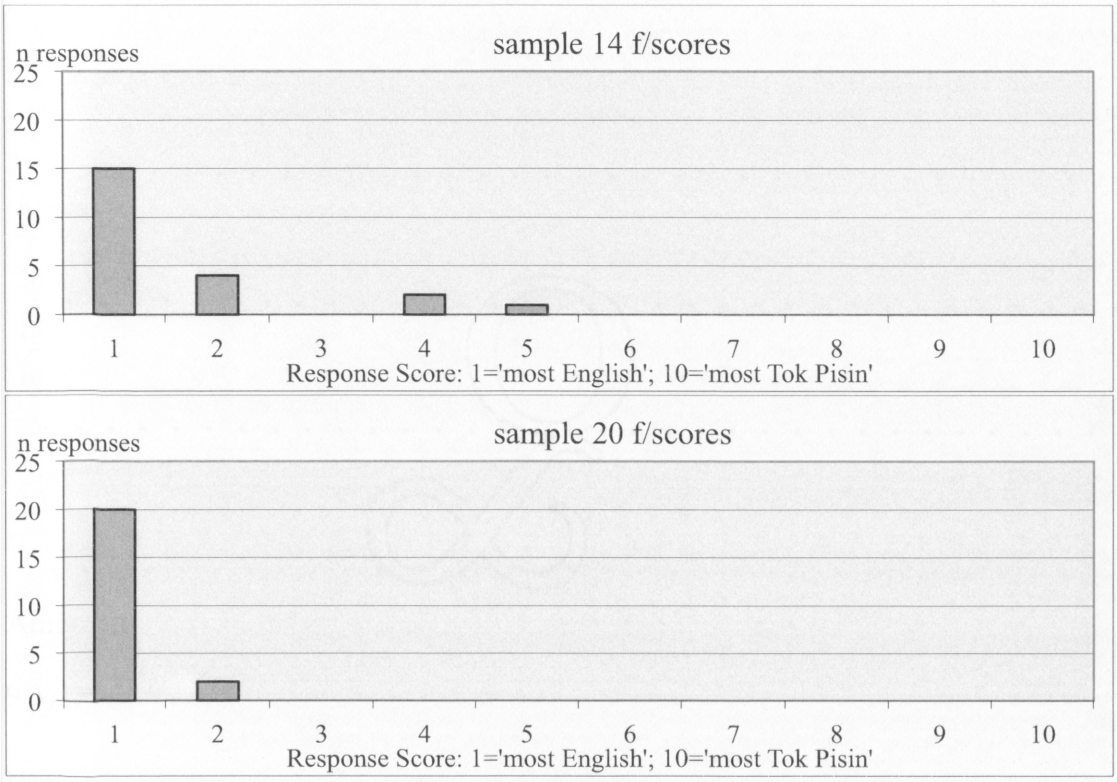
With the individual scores given by each informant to each sample tabulated, frequency tables and bar charts were produced to illustrate the range of responses.[82] From these it was clear which samples had been scored 'more Tok Pisin' or 'more English', which showed a pattern around a central value, and which showed the greatest range of responses.

---

[82] As noted above, decimal scores were not included in the frequency tables or the bar charts produced.

From these basic charts samples 14 ((317), below) and 20 (315) were identified as being scored 'most English'. Over half of all informants gave these samples a score of 1 (Figure 9.1).

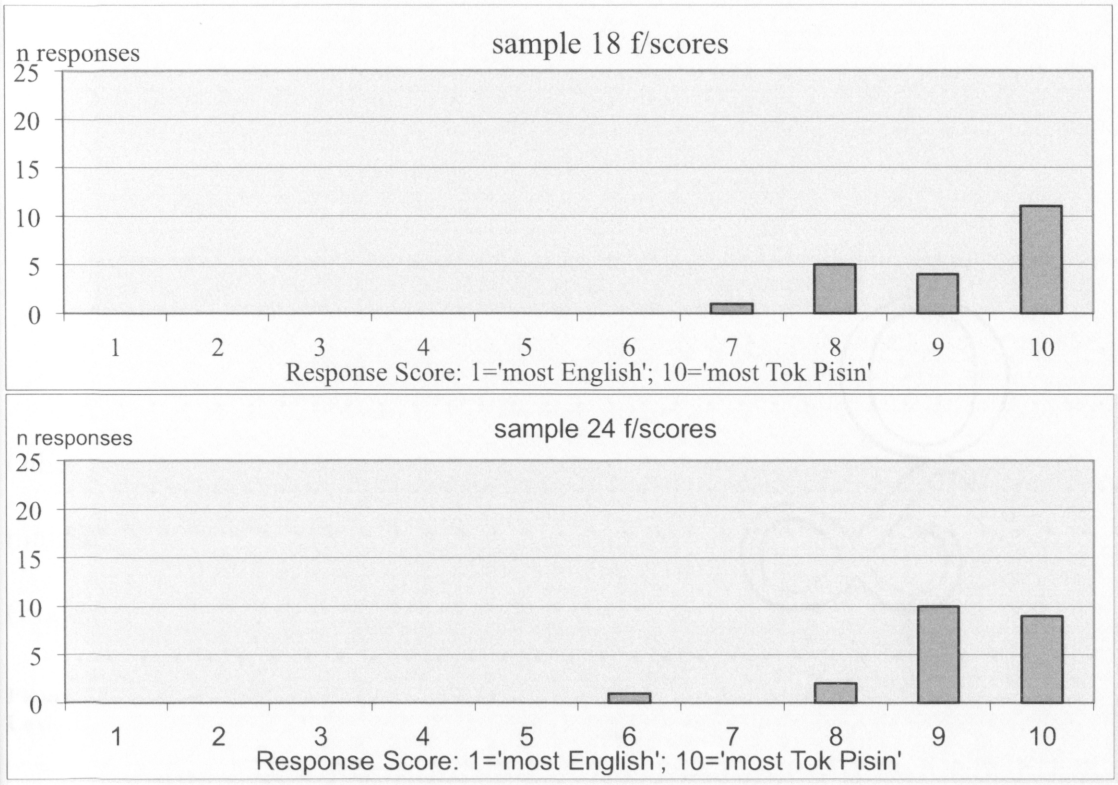
Figure 9.1 : Listening Exercise Responses : Bar Charts : 'most English' samples



These responses supported my sample selection criteria (section 9.2.1.1 above): samples 14 and 20 were included as being structurally English, albeit with some unusual constructions.

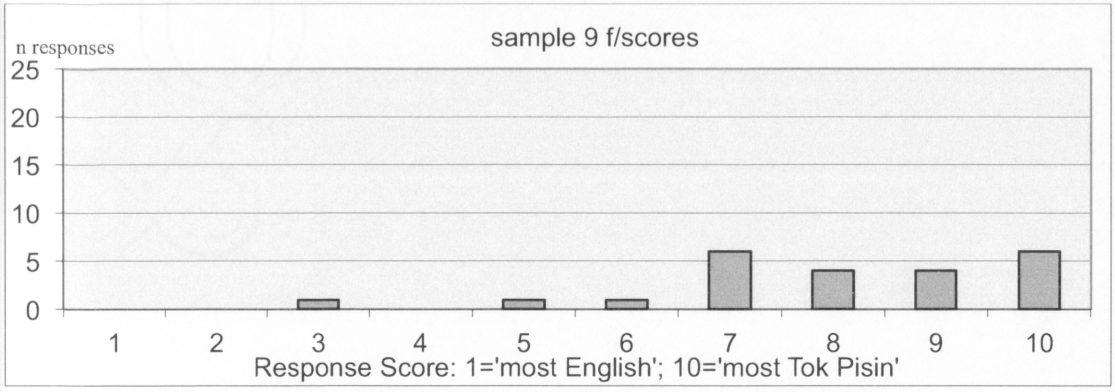
At the other end of the scale, no samples received such similar responses from these informants. This reflected the mixed language of these Tok Pisin samples. The samples scored 'most Tok Pisin' were items 18 ((321), below) and 24 (316). Figure 9.2 shows that most informants scored these samples above 8.

Figure 9.2 : Listening Exercise Responses : Bar Charts : 'most Tok Pisin' samples



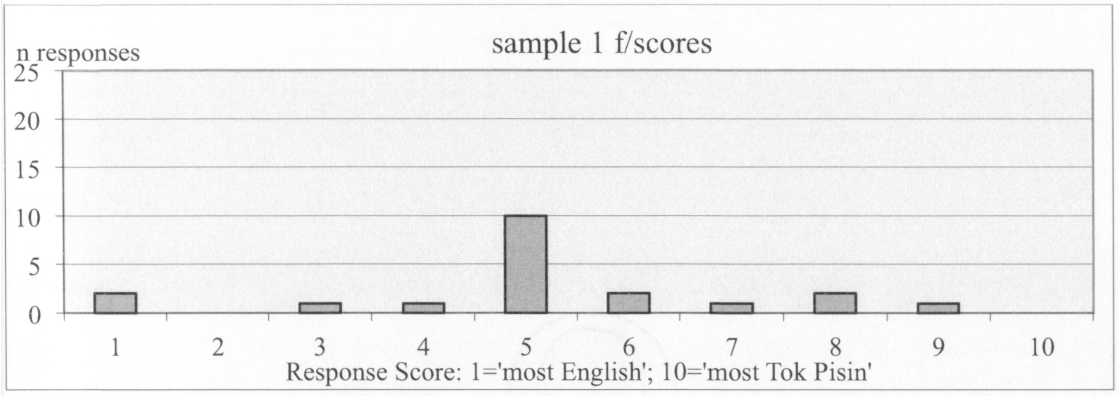
Among the samples that consistently received higher-end scores was sample 9 (313) (Figure 9.3), our example of Tok Pisin with English Insertion (s9.2.1.1.2).

Figure 9.3 : Listening Exercise Responses : Bar Charts : (313) Insertion = Sample 9



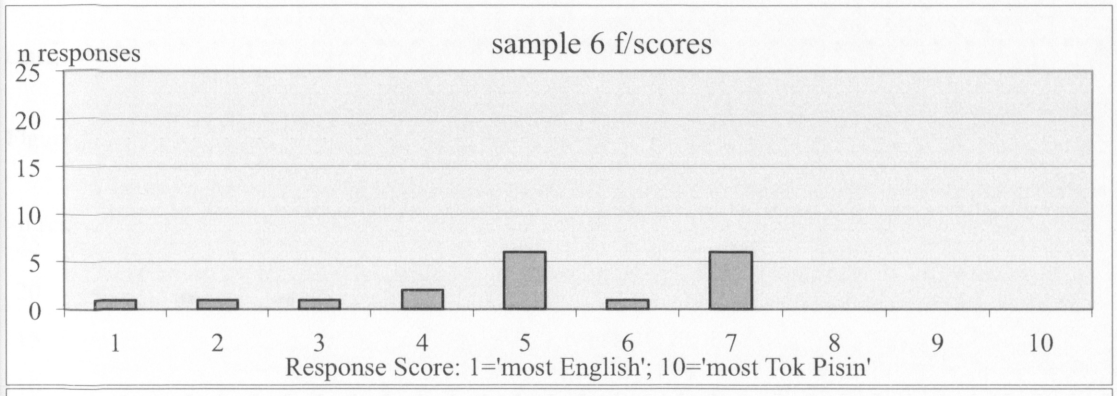
A third pattern of responses to samples could also be identified: those with scores ranged around a centre point. Of these the clearest example was sample 1 (312) (Figure 9.4), our example of Tok Pisin/English Alternation (s9.2.1.1.1). Despite a large overall range of responses, this marked Alternation was clearly placed mid-scale between 'most Tok Pisin' and 'most English' by the UPNG students.

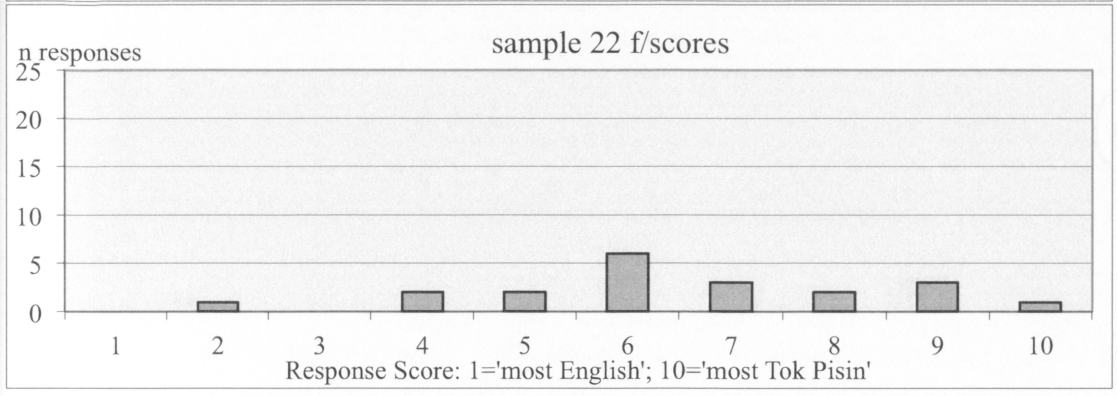
Figure 9.4 : Listening Exercise Responses : Bar Charts : (312) Alternation = Sample 1



Other samples that elicited similar mid-scale scores had more scores above and below the middle of the scale. Typical of these were sample 6 ((318), below) and sample 22 (314) (Figure 9.5)

Figure 9.5 : Listening Responses Bar Charts : (318) back-and-forth Alternation = Sample 6 and (314) Congruent Lexicalization = Sample 22





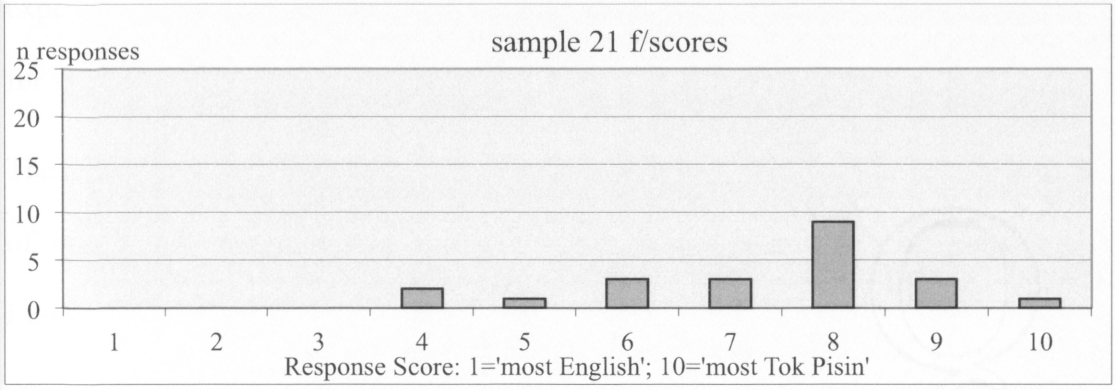
Again, these patterns matched with the types of code-switching in the selected samples. Sample 6 (318) featured repeated Alternations, and signs of Congruent Lexicalization in both English and Tok Pisin segments. Figure 9.5 shows that the spread of scores was 'more English'. Sample 22 (314) had been selected as showing the greatest indications of Congruent

Lexicalization in Tok Pisin.[83] Figure 9.5 shows clearly that the spread of scores for this sample was 'more Tok Pisin'.

Comparing the frequency of different responses to each of the samples it appeared that different types of CS could receive similar ranges of scores. This suggested that a comparison of samples by 'average' scores is necessary. It also indicted that Alternation between Tok Pisin and English was more salient for informants as making a sample 'more English' than either Insertion or Congruent Lexicalization.

In some significant cases, however, linguistically significant use of English function words in Tok Pisin, indicating Congruent Lexicalization, received a higher score. Sample 21 had a large range of scores, but these were clustered around a 'more Tok Pisin' value. This sample of speech showed signs of Congruent Lexicalization in the use of English pronouns, conjunctions, and conditionals (Figure 9.6).

Figure 9.6 : Listening Exercise Responses : Bar Charts : Insertions with ragged Alternation switching = Sample 21



[83] Sample 6, with switches to full-English sentences with Tok Pisin Insertions as well as Tok Pisin sentences with English Insertions received some lower scores than Sample 22, our example (316) of complex Insertion and back and forth switching in Tok Pisin indicative of Congruent Lexicalization. However, the bar charts in Figure 9.5 clearly show a cluster of '7' responses for sample 6, indicating that it was consistently assessed as 'more Tok Pisin' than a mid-value by around 20% of informants.

This initial analysis of responses to each sample strongly indicated that the informants' judgements of the 'most English' and 'most Tok Pisin' samples agreed with the selection criteria. It also indicated that Alternation was associated with scores around the mid-point values of 5 and 6. The majority of samples, those featuring Insertion of English elements in Tok Pisin, received scores in the upper part of the scale. It appeared that samples containing Congruent Lexicalization received a large range of scores, indicating differing perceptions of this type of CS: in some cases more mid-scale, in some cases 'more Tok Pisin'.

The analysis of the frequency of responses also showed that these informants were responding to the language of the samples rather than making judgements about the speakers. Responses to samples of speech by the same speakers received different scores: compare the bar chart for sample 9 (313) (Insertion) (Figure 9.3) with the chart for sample 22 (Congruent Lexicalization) (314) (Figure 9.5). Words spoken by the same person receive different scores when the type of CS involved differs. This finding proves the effectiveness of the experimental design (s9.2.1), and indicates that different types of CS have different salience for speakers.

The analysis of the frequency of different scores indicated that there was often a large range of responses, and that different examples of CS were scored in different ways. Consequently, I undertook an analysis of the 'average' responses to each sample. The following sections describe the patterns revealed by these analyses.

9.2.2.2 Quantitative responses to the Listening Exercise: comparing the two student Groups

As analysis of the frequency tables of scores given to each sample revealed, the range of responses for most of the samples was large, with only four out of twenty-six samples being given fewer than five different scores by the twenty-three informants. Having looked at the

distribution of response scores for each sample individually, 'average' responses were compared between samples.

Working with a small number of informants' responses in this initial exercise meant that I had no expectation of finding patterns behind individual differences in responses. However, the listening exercise had been designed in the field so that it could be administered to any group of informants with access to a cassette player. The intention was to develop a research tool that could extend to a much larger number of informants selected for what could be significant variables such as age, region of origin, gender, and level of education. For this pilot I compared the students by subject group: certainly, the Language and Literature students in Group 2 had made more comments on language issues in discussion, and on their answer papers (s9.2.2).

In order to generalise from individual patterns of responses within the smaller pilot groups and make some observations about the general assessment of these samples the mode, median, 1st quartile (Q1) and 3rd quartile (Q3) scores and the interquartile range (IQR) were found.[84] The median score was considered as a measure of the 'average' response to the samples' language. These average scores were calculated for the whole set of twenty-three informants, and for those in Group 1 and Group 2 separately. In this way the two groups could be compared to each other.

---

[84] These have been discussed in relation to similarities between samples in terms of the frequency of responses above (s9.2.2.1), and are considered in more detail here and particularly in s9.2.2.3 below.

Considering the median average of response scores showed high levels of agreement, with the two groups' median scores agreeing for thirteen out of twenty-six items (50%),[85] including our examples of Alternation (312) as well as examples of PNG English (315) and culturally and idiomatically inflected Tok Pisin (316). The median score of Group 1 (Economics and Law students) informants were higher ('more Tok Pisin') for nine samples,[86] including both our examples of Insertion (313) and Congruent Lexicalization (314). The Group 2 informants (Language and Literature students) gave four samples higher median scores.

The differences between the median averages of scores assigned to the samples by the two groups were interesting in the light of the different types of code-switching they featured. I hypothesised that the language specialisation of study of the Group 2 students meant that they recognised the L2 elements in some samples more readily than the Group 1 informants.

This hypothesis was tested using a Mann-Whitney non-parametric statistical test. This found nothing exceptional about the differences between their median average scores for each sample. As a consequence, it is most probable that the observed differences, though fascinating, are conincidental: it is likely that any division of the students into two groups could have produced similar differences.

Based on this initial experiment, there is no statistically significant difference between the responses of the two subject groups. Future applications of the listening exercise may include other variables and find significant differences. The indications from the pilot exercise with

---

[85] Two Groups agreement by median score: samples 1 (312), 3, 4, 6 (318), 10 (310), 11, 12, 17 (319), 19, 20 (315), 24 (316), 25)

[86] Group 1 (Economics and Law) 'more Tok Pisin' than Group 2 (Language and Literature), by median score: samples 2, 7, 8, 9 (313), 13, 15, 16, 18, 22 (314)

two groups of UPNG students are that, despite apparent differences in their responses, specific subject specialisation does not have an impact on the median response score to these samples.

9.2.2.3 Quantitative responses to the listening exercise: scaling the samples

For a final analysis, informants' responses to each sample were considered together, and 'average' scores for each sample identified. This made it possible to consider the samples of mixed language in an ordering that related to the scores they were given and to see what patterns emerged.

As discussed in terms of the frequency and range of responses above (s9.2.2.1), many of the samples received a wide range of response scores in the UPNG listening exercise sessions. Differences between the most-frequently-occurring (mode) score and the mid-point (median) score for each sample reflect this, as do larger interquartile ranges (IQR) (see Table 9.1).

The choice of four informants to use decimalised scores (e.g. '5.5') while others stuck to integer values had some affect on the mode score. These decimalised scores, if rounded into integer 'bins' could combine with other scores to increase their frequency. The decision to retain the informants' original scores meant that, in some cases, the most-frequently-occurring (mode) score would not include decimalised responses. Consequently, the median score, calculated as the mid-point of all the informants' responses, was used to compare the samples.

The range of scores for each sample made it clear that mixed language was not consistently salient for all informants. The key question addressed by this comparison was: were different types of CS generally receiving particular scores?

In order to address this question the samples were ordered by the 'average' of all response scores: detailed comparisons were made of samples with similar median and mode scores, and

in terms of the IQR of responses. The key question addressed by this analysis was whether these groupings and the ordering of samples could be related to any observed linguistic features of CS in the samples: how did the scale match the selection criteria?

Figure 9.7 illustrates the scale of samples by median score. It provides a consolidated view of the findings from the frequency tables discussed with bar charts in s9.2.2.1: that the samples with the most agreement in scores (the smallest IQR, indicated by the size of the 'whiskers' (lines above and below the marker indicating the range of responses) on Figure 9.7) are those ranked 'most English' (samples 20, 14), 'most Tok Pisin' (samples 24, 18), our example of a single marked Alternation (313) sample 1. All these samples have very similar median and mode scores (indicated by the boxes and crosses on Figure 9.7 respectively). With a similarly close range of responses is sample 19, which is an example of Tok Pisin with a single, idiomatic, EL island (*come live on TV*) and an Inserted cultural noun (*hope*). This indicates a similar salience for the type of language involved in these samples for most informants.).

Figure 9.7 : Samples Ordered by Median Score, showing IQR whiskers and Mode

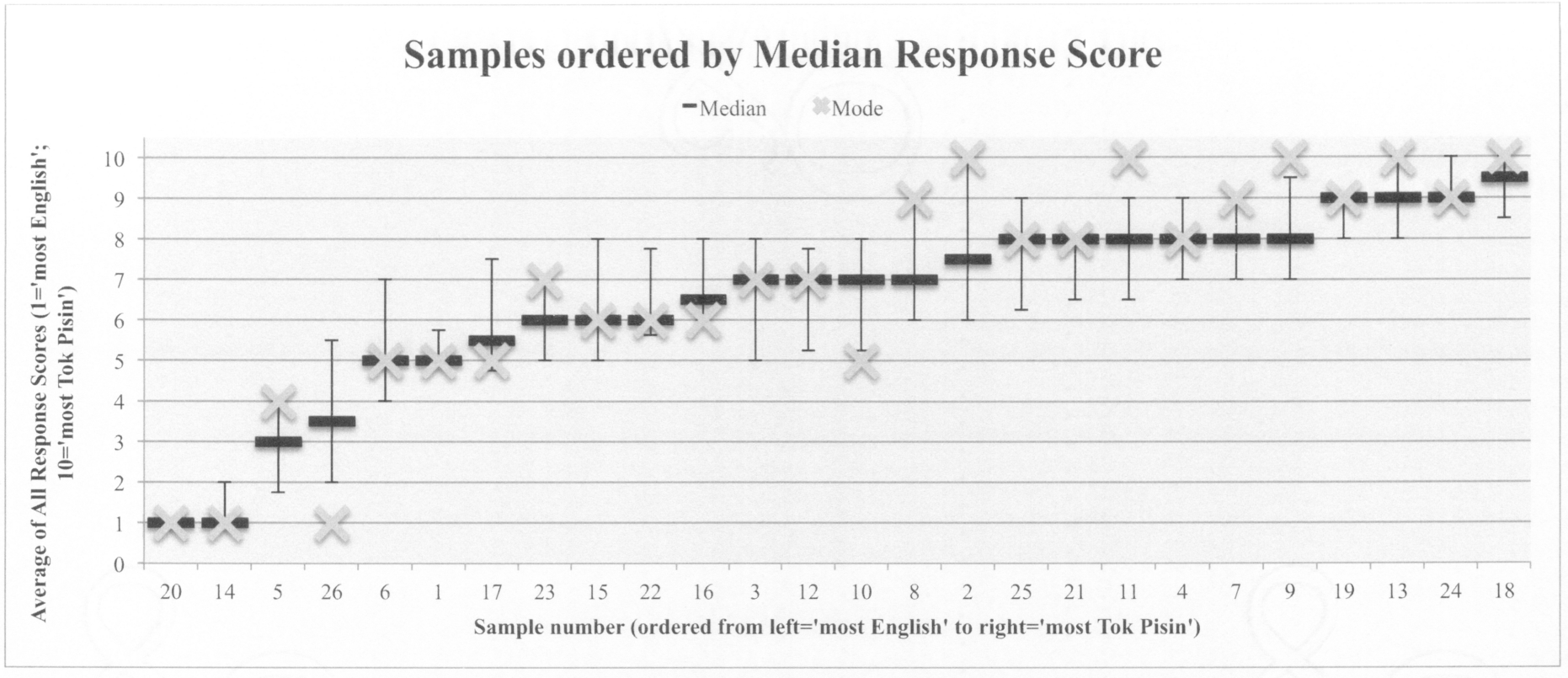


Figure 9.7 presents an ordering of samples (along the x-axis) by median score. Where the same median score is indicated, the quartile (Q1 and Q3) values (indicated by the whiskers Figure 9.7), are used to differentiate them. This scale from 'most English' (left) to 'most Tok Pisin' (right) is discussed below. The mode score is also indicated for comparison.

For samples with a median score of 7 or less, with three clear exceptions the median and mode scores generally agree. Of these exceptions, samples 5 and 26 were PNG English with Alternations to Tok Pisin, and these had similar median scores. The third exception was sample 10, which featured back and forth Alternations between PNG English and Tok Pisin that were noted as indicating Congruent Lexicalization.

Among the samples scored 'more Tok Pisin' however, with median scores of 8 and above, the median and mode scores often were different. This indicates that there were many different judgements of these samples by these informants.

Table 9.2 reproduces the earlier list of samples (Table 9.1) ordered by ascending score. It reveals clear associations between particular scores and different types of CS. With the top of Table 9.2 indicating the 'most English' samples and the bottom of the table the 'most Tok Pisin' samples, as ranked by the median of informants' response scores, it can be seen that the selection criteria were reflected in these responses.

We can read Table 9.2 column by column and see clear patterns in line with the selction criteria (s9.2.1.1): samples with "Matrix Language", English were ranked 'more English' than those with ML Tok Pisin. In terms of CS type, samples with Alternation appear in the middle of the table, generally associated with median values below 6. Insertions dominate the lower, Tok Pisin, part of the table, though samples with mixed constituents are towards the middle. Samples marked with Congruent Lexicalization appear towards the middle of the table. However, most interestingly, some of the samples ranked 'most Tok Pisin' include English system morphemes, including a pronoun, adverb, and copula in sample 13.

The following sections characterise the samples by median score.

Table 9.2 : Table of numbered samples - ordered by Median score and quartiles

						CS constituent types				
Sample Number	**Median Score**	**IQR**	**Mode Score**	**"Matrix" language**	**CS type**	**Content morphemes**	**System morphemes**	**Mixed Constituents**	**Islands**	**Alternations**
20	1	0	1	PNGE	CLx					
14	1	1	1	Eng	CLx	1 Core V				
5	3	2.25	4	PNGE	Alt / Ins	TP : 1 Cultural N; 1 Core V	PNGE: Qtag, Neg; TP: Disc.Tag;		1 Discourse	2 : Tok Pisin sentence with Insertions, and back
26	3.5	3.5	1	PNGE	Alt					2 : repetition and back
6	5	3	5	PNGE / TP	Alt / CLx		PNGE : Plurals (dbl marked); Q.Tag; Conj ; TP : Conj			4 : back and forth
1	5	0.75	5	TP	Alt	1 Cultural N				1 to English
17	5.5	2.75	5	PNGE	Alt / CLx		Pron			5 : back and forth
23	6	2	7	TP / PNGE	Alt / Ins	TP = 1 Cultural N			PNGE = Insertion TP island	3 : back and forth
15	6	3	6	TP	Ins	1 Cultural N; 1 Adv		2 integrated V	1 Pron+V (discourse)	
22	6	2.125	6	TP	CLx / Ins	1 Core N	Plurals; Conj; Adv	3 integrated V	3 Adj+N; 1 Two-Word V; 1 Adv+N; 3 EL Island NP	5 : back and forth (ragged)
16	6.5	2	6	TP	Ins / CLx	1 Core N	Conj.Disc	1 mixed VP; 3 integrated V	1 Adj+N	1 : ragged
3	7	3	7	TP	Ins	4 Cultural N (1 repeated 3 times); 1 Core N			1 Adj+N; 1 Two-Word V	
12	7	2.5	7	TP	Ins / CLx	2 Cultural N; 2 Core N (1 repeated)	Conj; Adv		3 Adj+N (1 repeated); 1 VP; 1 PP; 1 NP	2 : flagged PP and back
10	7	2.75	5	TP / PNGE	Clx / Alt		PNGE : Conj; TP : Prep			4 : back and forth
8	7	3	9	TP	Ins	2 Cultural N; 2 Core N		1 integrated V; 1 integrated Two-Word V	1 Adj+N,; 1 EL island (idiom) PP; 1 EL Island NP	
2	7.75	4	10	TP	Ins	1 Cultural N		1 integrated V (translated)	3 Adj+N	
25	8	2.75	8	TP	Ins / CLx	4 Cultural N; 1 core N; 1 V-Cont	Conj; Prep; Spec; Cont		1 EL? Island VP (CLx)	
21	8	1.5	8	TP	CLx	3 Cultural N; 1 Core N; 1 core V (unintegrated)	Pron; Conj;Plurals (dbl marked)		1 CondP	3 : back and forth (incl. ragged CondP)
11	8	2.5	10	TP	Ins	1 Cultural N		1 integrated V (translated)	3 Adj+N	
4	8	2	8	TP	Ins	3 Cultural N; 2 Core V, Numbers	Disc.Tag; Plurals		2 Adj+N	
7	8	2	9	TP	Ins	3 Cultural N; 1 Pronoun	Disc.Tag; plurals		1 EL Island VP; 2 Adj+N	
9	8	2.5	10	TP	Ins	1 Cultural N; 1 Core N	Plurals (dbl marked)	1 integrated V	1 Adv.Time; 1 EL island AdjP; 1 Adj+N	
19	9	1	9	TP	Ins	1 cultural V			1 EL Island VP - idiom	
13	9	2	10	TP	Ins	1 Cultural N	Conj; Pron; Cop; Adv		1 Disc	
24	9	1	9	TP	Ins	6 Cultural N	Plurals (dbl marked)		1 EL Island PP - idiom	
18	9.75	1.5	10	TP	Ins				1 EL Island NP (repeated); 1 Adj+N; 1 Proper N (repeated)	

Median Scores below '5': 'more English'

As discussed above, the UPNG students agreed with the selection process on the 'most English' samples. However, for the 'most English' samples, the lower IQR of 0 on Table 9.2 (and the whiskers on Figure 9.7) shows us that most informants responded to sample 20 with a 'more English' score than sample 14. This result is interesting, because 20 is a national politician while 14 is a sample of the researcher's speech. As the researcher's speech was included as a control, it might be expected that 14 would be given a 'more English' score, as the accent was not a Papua New Guinean English one. It seems clear, however, that it is the structural characteristics of sample 14 (317) that were identified by informants as more 'like Tok Pisin', in particular the constructions using the word *stop* in English sentences in a position and role equivalent to Tok Pisin *stap* :

(317) TiT#25b@7:30 = **sample 14**

... A thIng Iss wuhn sIksth ## ov uhl dhuh lAngwuhdjIs ## In dhuh wuhl dhei stuhp *uhp ## In Is kuhntrI ### . nau wuhn sIksth sIkstIn pesent uh aul dhuh wuhld lAngwIdjIs dhuh* stuhp *In* PapuNuGInI

... I think it's one sixth, of all the languages, in the world they stop/stap ***up, in this country.*** Nau/***Now one sixth sixteen per cent of all the world languages they stop***/stap ***in Papua New Guinea.***

'... I think it's one sixth, of all the languages, in the world they stop up, in this country. Now one sixth sixteen percent of all the world languages they stop in Papua New Guinea'

This marked calque from Tok Pisin could explain why this sample received a higher mean score than sample 20, spoken by a National politician.

With higher median scores than the 'most English' samples are samples 5 and 26: English featuring Alternation to Tok Pisin. Sample 5, with its Tok Pisin discourse marker *a* on the English segment had a higher mode yet a lower median score than sample 26, where the Alternation to Tok Pisin was temporary, and served to repeat a phrase from English. The differences in mode and median scores indicate some confusion between these samples. By mode, the Alternation in sample 26 is grouped with the 'most English' samples, and the Tok

Pisin discourse marking on sample 5 was clearly salient for many speakers, though it had a slightly lower median score.

Median Score '5': Alternation

The next group to be considered are characterised by Alternation.

Showing the greatest agreement in score (an IQR of 0), with a median score of '5', was sample 1, which was standard, Rural Tok Pisin except for a single cultural loan, followed by a monodirectional Alternation to English for a question. The English samples on either side of it on the scale, samples 6 (318) and 17 (319) featured both Alternation and characteristics of Congruent Lexicalization, and both have median scores of '5' or above.

The larger IQR allows us to appreciate that responses to sample 6, English with multiple Alternational switches to Tok Pisin, including the switching of conjunctions and Congruent Lexicalizations, as discussed in Chapter 8, were more varied than for sample 1. This shows that these informants had more agreement about the rating of the simple Alternation of sample 1 (312) than the mixed language of sample 6 (318):

(318) TiT#24b-15@1:40 = **sample 6**

yu slip dea an deyl kuk fud fo yu # so yu kaikai kaikai yu stap tasol ### *an den # dey a pipuhl hu wil gow an* ah- ## *luk fo yo ## fo yo potin ## . an potin is ol wayl animuhls ## dhem kil flain foks dheyl kil pig waly pigs* uhm- uh- *bandicoot an watnot* a ? *wayl buh:ds #####* bai ol i kisim kam ol i kukim lo yu . *an* ah- ##### nau bai ol i kukim nuh ## bai i ### kukim saksak tuhnim saksak . na sup # *deyl mek sup yuno* uh- *get* ah- ### *taro an yam an en dhey mes it* a ?

yu/***you sleep***/slip ***there, and they'll cook food for you. so*** yu kaikai kaikai yu stap tasol, ***and then, they are people who will go and ah-, look for your, for your protein. and protein is*** ol/***all wild animals, them kill flying fox they'll kill pig wild pigs*** uhm- uh- ***bandicoot and whatnot*** a? ***wild birds***, bai ol i kisim kam ol i kukim long yu. ***and*** ah-, nau bai ol i kukim nau, bai i kukim saksak tanim saksak. na sup, ***they'll make soup you-know*** uh- ***get*** ah-, ***taro and yam and then they mash it*** a?

'You sleep there, and they will cook food for you. So you eat and eat, you just eat, and then, there are people who will go and kill flying fox they'll kill wild pigs uhm, uh bandicoot and what-not eh? Wild birds, they will bring them and they cook for you. And ah, they will cook, they will, cook sago pull sago. And soup, they'll make soup you know uh het ah, taro and yam and they mash it eh?'

Sample 17 (319) received a higher median score. Although like sample 6 it features five Alternations it has no obvious Tok Pisin discourse marking on the English. However, it does include an English pronoun (s8.2.6) in the final sentence:

(319) TiT#08-507 = **sample 17**

firs ting thed in tuhms of yuu pipuhl sed eit tuhti and wi astatinga miting at ten oklok yu kan savi olsem wuhnem mipla gat ***ada things tu du*** ***###*** . aam no isi long apim ol nau # ***gimmi pon ##*** . ***yuu tel mi dhis lausi ## grr tradisuhn fain mi fo twentieit kina an # dhis chip shuht det ai am werin fo twenti kina*** em wanem ? oh- mi tokim ***im***.

first thing first in terms of you people said eight thirty and we are starting the meeting at ten o'clock you/yu ***can***/kan save olsem wanem mipela gat ***other things to do***. em no isi long hapim ol nau, ***gimmie point. you tell me this lausy, grr-tradition fined me for twenty-eight Kina and, this cheap shirt that I am wearing for twenty Kina*** em wanem? oh- mi tokim ***him***.

'first things first in terms of you people said eight thirty and we are starting the meeting at ten o 'clock you must know that we have other things to do. They weren't able to argue with them then, gave me a point. You tell me this lousy grr-tradition fined me for twenty-eight kina and, this cheap shirt that I am wearing for twenty kina what of that? Oh, I told him!'

Significantly, then, it appears that the kind of rapid alterntions including inserted EL pronouns in sample 17 may be viewed as 'more Tok Pisin' than the language in sample 6: a combination of sentence Alternations with mixed discourse marking and conjunctions. Though the lower quartile (Q1) figures (indicated by the whisker lines below the median marker on Figure 9.7) for both of these samples are the same, sample 17 has greater range of scores (IQR) and an upper quartile (Q3) scoreof 8 compared to 6 for sample 6 (indicated by the upper whisker lines on Figure 9.7).

Median Score '6': large amounts of English

Comparing the features of the samples in each group by median score allows us to analyse the bulk of the samples characterised as Tok Pisin with various types of Inserted material.

With median scores of 6 are samples containing different behaviour. Sample 23 has lower quartile scores, which indicate responses tending to 'more English', but a higher mode than samples 15 and 22. Sample 23 features an Alternation to an English segment including a Tok

Pisin VP. Sample 15 was selected for the prominent placement of mixed verbs and an English adverb *indirectly*. Sample 22 also includes mixed verbs, along with inserted idiomatic phrases such as *whatever the playing on the cards*.

With the same mode as sample 22 but a slightly 'more Tok Pisin' median score of 6.5 is sample 16, with a mixed verb phrase in addition to V-TR mixed constituents. It overlaps with the next grouping.

As a group these samples have large proportions of English integrated with their Tok Pisin.

Median Score '7': mixed language

The samples grouped together with a median score of 7 are definitely mixed language, and are characterised by the inclusion of longer stretches of English into Tok Pisin, principally by Insertion. Reflecting the overlapping patterns of responses discussed in s9.2.2.1 above, most of the samples in this grouping[87] have quartiles that match those of samples 15 and 22. Sample 8, however, has higher quartiles, indicating some higher response-scores. Sample 16 features the mixed verb *recognisim* and the use of English *still* for Tok Pisin *yet*, as well as the continuing conjunction *and also* from English. Its quartiles of 6 and 8 suggest it may be considered with the median 7 group, but it is also an excellent example of the fuzziness of the rankings on this scale.

On linguistic features alone it is difficult to assess this grouping, with different types of Insertions receiving similar median scores. Samples 3, 12, and 10 have similar quartile scores, but different types of behaviour. Sample 3 is characterised only by the Insertion of cultural

[87] Median scores of '7': samples 3, 10, 12, 16

loan nouns. Sample 12 has long stretches of mixed language, which may be responsible for quartiles of 5 and 8, a spread similar to that of sample 10.

Sample 10 features back and forth switching and English prepositions, indicating Congruent Lexicalization. An impression that this is a high median score for apparent Alternation is moderated by an awareness of the mode score of 5, similar to other Alternations. The difference may be a consequence of the repetition of English constituents with Tok Pisin equivalents, rather than the languages being used to express different thoughts. This can be seen in two places (320) where English NPs *him* and *the problem* are repeated with Tok Pisin *em*:

(320) TiT#25b @3.10 = **sample 10 (extract)**

laik ***him # is*** em malgebuk a ### ? ...
... ***thuh problem*** em sap lo mipluh uhm paba ...

like him, is em marit-Gembog a ? ...
... ***the problem*** em stap long mipela em papa ...

'like him, is he married a Gembog woman eh? ...
... the problem it is our responsibility the fathers ...

In the 'more Tok Pisin' part of this group, Sample 8 includes a range of different Insertions, including the whole English island NP *almost twelve thousand pounds of materials*. However, it has a 'more Tok Pisin' mode of 9, indicating, once again, a large difference of opinion.

Median Score '8': 'Urban Tok Pisin'

Samples with a median score of 8[88] could be characterised as Urban Tok Pisin, with some loan phrases, such as *outside's condition* in sample 25, and discourse markers, but minimal English structural elements.

In this group our duplicate sample 2/11 occurs. With a median score mid-way between 7 and 8, a consequence of one informant using fractional scores, sample 2 has an identical mode score to sample 11. This sample, from a radio news broadcast, features Inserted cultural loans: nouns, Adj+N, and a single V-TR mixed constituent that is translated by the speaker. The differences in median scores indicate that the ordering of samples may have some bearing on the informants' responses.

As a group none of these samples show the same large islands of English or complex mixtures of the samples with 'more English' scores.

Three different samples in this group have 'more English' quartiles. One of these is sample 2/11. Sample 25 featured cultural loans and discourse elements such as the continuing phrase *so that*. Sample 21 was discussed above as being scored differently by the two informant groups due to characteristics of the replacement of some Tok Pisin words by English alternatives: for example *alone* is used for Tok Pisin *tasol* 'only'. Samples 21 and 25 are distinguished by the presence of some English phrases.

With 'more Tok Pisin' quartiles, samples 4, 7, and 9 have the same Q1 figures of 7 but different mode and Q3 scores. Sample 4 includes core loans replacing Tok Pisin verbs, and

---

[88] Median scores of '8': samples 4, 9, 11, 21, 25

English-style numbers. Sample 7 was spoken by a radio broadcaster easily recognised by the student informants, and the language featured English discourse markers *well* and the pronoun *we*, as well as cultural nouns related to talkback radio: *calls*, *lines*, and *company*. Notably, the difference in mode and median scores indicates the EL island *try me again tomorrow morning*, spoken with a PNG English accent, was not salient for all informants. Sample 9, which has a high Q3 of '10', features only the short idiomatic phrases *very short-term plans* and cultural loans, including Adj+N islands, *last week*, *pre-emptive measures*, and a single noun: *problems*. These are all well-integrated into the speaker's Tok Pisin, which was recognised by most informants as educated urban language.

Median Scores over '9': 'most Tok Pisin'

Receiving the highest median scores of 9 and above were four samples[89]. Samples 19 and 13 similarly featured only cultural loans and idiomatic phrases. Sample 19 was mentioned above as having a low IQR. However, its 'most Tok Pisin' score is interesting in that it contains the English conjunctions *because* and the adverbial construction *that's why*, which contains an English pronoun and copula. Sample 24 (315), our example of idiomatic and culturally-inflected Tok Pisin, spoken by a recognised national politician, had the 'most Tok Pisin' lower quartile of all samples: informants consistently placing it at the top of the scale.

The 'most Tok Pisin' score overall was given to Sample 18 (321), which, while it included EL islands, was from a radio announcement specifically explaining the meaning of those English NPs in Tok Pisin:

---

[89] Median score '9': samples 7, 13, 19, 24.

(321) TiT#10b-02 = **sample 18**

yes wan lili isin tokim mi tetei olsem em i ***wuhld hel dei*** # em i taim we ***lai neishins*** i ma:kim olsem taim bai yumi ma:kim # ol wanem kain kain ah& ah& sait blong uh& ***wawk helt*** # yumi saw lukim luh isela isa graun so tetei # ah& em i ***wuhl hel dei*** emm ah& ***ilait neishuhns*** i maakim long em .

yes one liklik pisin tokim mi tedei olsem em i ***world health day***, em i taim we ***United Nations*** i makim olsem taim bai yumi makim, ol wanem kain kain ah- ah- sait bilong uh- ***work health***, yumi save lukim long dispela dispela graun ***so*** tedei, ah- em i ***world health day*** em ah- ***United Nations*** i makim long em .

'yes I have some pidgin to say: today is 'World Health Day', it is recognised by the United Nations and will be recognised by us, it is like ah- ah- the same side as uh- Work Health, we understand what that is so today, ah- is 'World Health Day' it, ah- United Nations have established it.'

9.2.2.4 Summary of Responses and Analysis

The responses to the Listening Exercise appeared to confirm the salience of certain types of language mixing for the UPNG informants. Alternations were scored 'more English' than Insertions, yet the mode and median of response score often differed, indicating a range of scores from informants. EL island Insertions encouraged 'more English' scores than simple content, discourse, cultural loans, or idiomatic phrases. Congruent Lexicalization features were variably scored: median scores of samples with back-and-forth switching and the presence of English system morphemes such as pronouns, conjunctions, and prepositions outside of EL islands were not placed particularly 'like English' or 'like Tok Pisin' on the scale.

A relationship between types of language mixing and responses to the samples seemed a coherent explanation.

As part of the coding of the core corpus (s4.2.5.2) I had applied language-tag codes to the texts from which the samples were drawn. Calculating the proportion of Tok Pisin words in each sample allowed the comparison of the median response scale and the proportion of languages in each sample. This reflected the frequency-based criteria for determining the ML of bilingual speech suggested by Myers-Scotton (Myers-Scotton 1993a), though this was only

intended to apply at the level of the discourse-sample as a whole, rather than smaller sections, as was the case in the Listening Exercise.

Figure 9.8 overlays markers (the circles) indicating this proportion on the scale of samples by 'average' score and shows they follow a similar trend from 'most English' to 'most Tok Pisin': the median of scores given to each sample increased along with the proportion of Tok Pisin words in the sample. Was it possible that the informants had scaled the samples on the basis of rapid calculations of similar proportions? Or can variations from the overall trend best be understood in terms of the types of CS in specific samples?

Once again, examining the data in detail provided answers to these questions. Sample 17 (319) has a low proportion of Tok Pisin words (indicated by the red dots (@1_x10) on Figure 9.8) that is similar to sample 5, yet its median score was significantly higher. However, sample 17 (319) features its short switches to Tok Pisin in prominent topic-comment relations. The prominence of the switches would appear to make them more salient, and sample 17 is scored with examples of Alternation.

In sample 15, where the median score seems low for such a high proportion of Tok Pisin words, the English nouns, discourse tags, and integrated verbs and adverbs are very prominent.The same may be said for the few English cultural loans in sample 3, though these are embedded in Tok Pisin. This prominence may account for the lower median score compared to the proportion of Tok Pisin words in these samples: the English more clearly perceived by informants, the CS consequently more salient, These two samples have the same proportion of Tok Pisin words as sample 24, with cultural loans and idiomatic phrases, but a lower median score, indicating that the informants' responses related to the type of language mixing.

Figure 9.8 : Samples Ordered by Median Score with Proportion Tok Pisin Words indicated

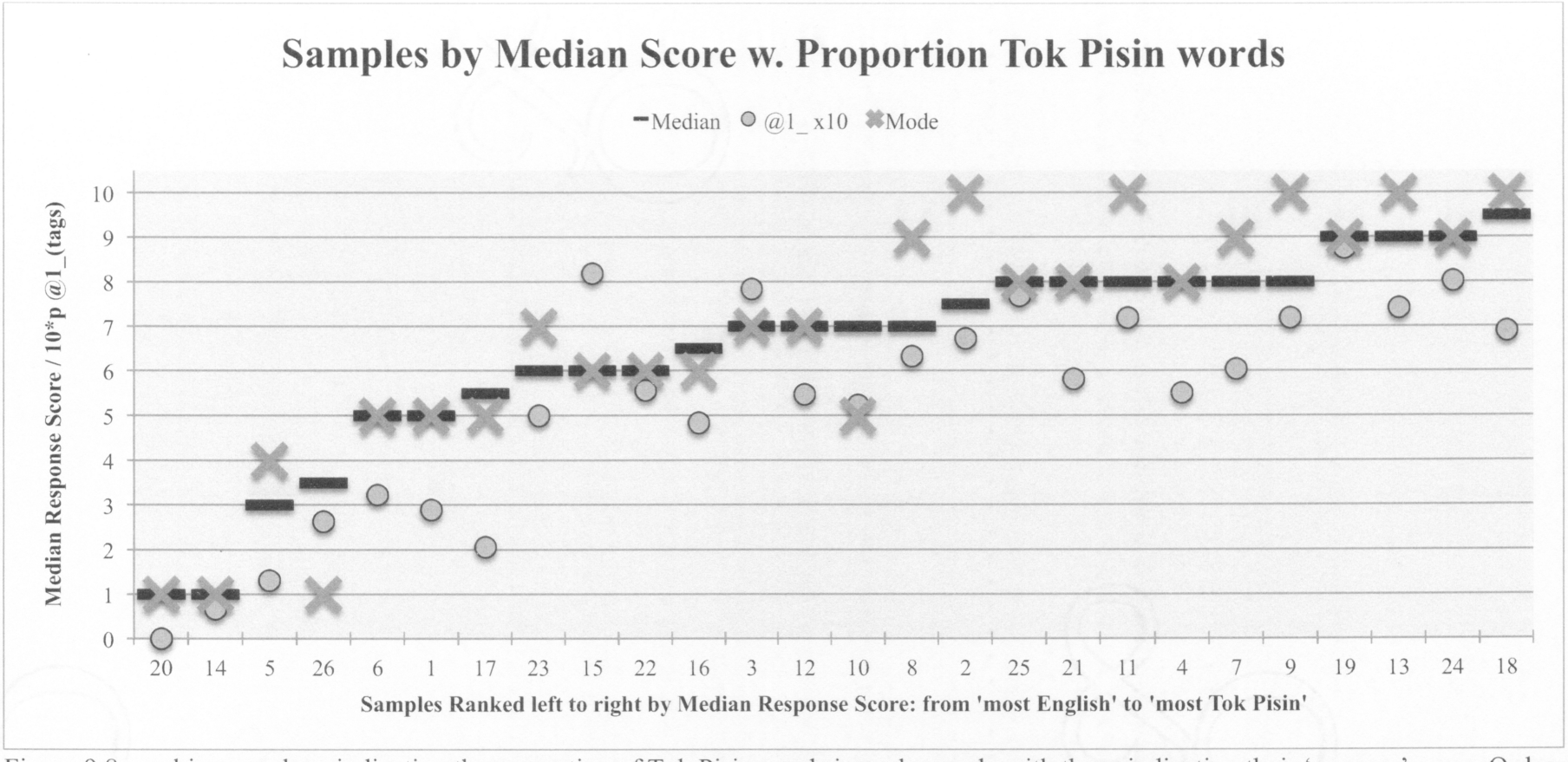


Figure 9.8 combines markers indicating the proportion of Tok Pisin words in each sample with those indicating their 'average' score. Ordered from 'most English' at the left to 'most Tok Pisin' at the right, by 'average' score, it can be seen that the proportion of Tok Pisin words (@1_x10) follows a similar tendency. The exceptions can be explained in terms of the types and prominence of CS in the samples.

Towards the 'most Tok Pisin' of the scale, the many novel English words in sample 18 (321) (indicated by a lower proportion of Tok Pisin words relative to samples with 'more English' median scores) were clearly translated by the speaker (a radio announcer). This, again, explains the 'more Tok Pisin' median score this sample received in terms of the salience of mixed language.

In each of the cases where the proportion of Tok Pisin words seemed out of step with the median score either the smaller size of the sample or the prominence, or otherwise, of the type of language mixing appeared to affect the responses represented in the median score.

Furthermore, there is a correlation between different types of CS and the proportions of material involved from each language, which also explains the placement of samples on the median scale: Alternations may involve equal proportions of material, thus leading to mid-scores; the Insertion of EL islands and mixed constituents create slightly unequal proportions, which tend towards the matrix language when smaller constituents, such as Adj+N or single words are involved; significantly, Congruent Lexicalization can involve varying proportions of words from either language, and this may be reflected in these samples' spread across the scale.

It is clear that the type of language-mixing is still significant: though it is possible that these informants were rapidly evaluating mathematical counts of words just as CLAN programs do (our first question, above). What is even more interesting is that even where the small amounts of English words in some samples were grammatically active and significant a relatively 'more Tok Pisin' median score was being assigned. Language mixing that may be considered linguistically significant, such as that which violates proposed constraints on CS, suggests a change in the morphosyntactic structure of a language, or indicates structural borrowing or convergence, is not necessarily salient for these informants.

9.3 Perceptions of CS: Interpretations from the Listening Exercise

The pattern of responses in the data collected from informants was assessed through frequency-analysis. As well as indicating strong agreement between responses to the 'most English' and 'most Tok Pisin' samples, this analysis revealed an association between samples selected as indicating Congruent Lexicalization and responses towards the middle of the scale. This was taken to indicate a perception by the informants of these samples as representing a 'mix' midway between Tok Pisin and English.

A comparison of the scores given to the samples by informants in the two groups revealed some apparent differences, but these were found to have no statistical significance. A larger study may establish factors that are significant in affecting responses to mixed language.

Using the mode-average of all response scores, some patterns emerged from the data. Taking the median-average of all response-scores for each sample as an indicator, the samples could be broadly characterised as different varieties of language. In terms of English, structure appeared to be more important than pronunciation in deciding whether a sample was 'more Tok Pisin'. In general, samples featuring Alternation were scored below 6. Samples of Tok Pisin indicating Congruent Lexicalization received scores with mid-scale averages and a large range of responses. Samples with Islands of inserted English, including Adj+N combinations (s7.3.2), and mixed constituents were scored lower than those with cultural loans and idiomatic phrases. This suggests that EL Islands, mixed V-TR constituents, and Adj+N constituents from English were more salient than cultural loans, but less so than clear Alternations.

In several cases, the presence of English system morphemes, discourse markers, and conjunctions alone did not seem to result in a lower score. For these informants, high scored 'Tok Pisin' samples could include discourse, structure, phrases, and, for some, even pronouns,

from English so long as they were well integrated into the sentence. This validates the interpretations of the 4M model (s7.3.4.3), by suggesting that these elements are not perceived as diluting the Tok Pisin of the samples. As indicated by Mühlhäusler's comments on 'post-pidgin' Tok Pisin (Mühlhäusler 1985d: 147-148) (see s3.4.3 (22), above), the presence of these features is salient for linguists as representing grammatical change. For these UPNG informants, however, quantity of English in mixed language was more salient (s9.2.2.4) than the type of material, and consequently the UPNG informants had a high tolerance for English system morphemes in the 'most Tok Pisin' samples.

This indicates that language mixing behaviour (CS) types that for linguists indicate a change in the Tok Pisin grammar may not be significant for bilingual speakers. Such a finding suggests the facility with which structural change and borrowing can take place through unmarked CS.

9.3.1 Extending the Listening Exercise

Devised and conducted while in the field, the UPNG pilot of the '*skelim tok Inglis na tok Pisin*' listening exercise was intended primarily to gather measurable feedback from speakers on the mixed language data analysed in terms of language contact and typologies of CS in Chapters 5, 6, 7, and 8. It successfully showed that speakers differentiated 'English' from 'Tok Pisin' in clear cases, and identified Alternation at a mid-point between them.

As conducted, this experiment collected data that suggests a hierarchy of salience for the Insertion of different constituents, and also that Congruent Lexicalization is not particularly salient for the UPNG students who participated in the study. Following from such valuable first findings, there are a number of suggestions for further research in addition to simple repetition of the initial study.

Firstly, a controlled selection of standard English, PNG English, 'Urban' Tok Pisin, and 'Rural' Tok Pisin data in the range of samples.

Secondly, a controlled selection of CS types for each type of language. The pilot study contained a bulk of Insertional CS, often occurring alongside Congruent Lexicalization or Alternation, as this is characteristic of mixed language in PNG, and the core corpus in particular.

Ideally the number of samples should be small, providing a tighter set of types of mixed language for evaluation. Also, the samples would all be of a similar length: the samples used in the pilot had a mean length of 22 seconds but with a high sd of 8.72 seconds. Shorter, clearer samples of particular language/CS type combinations would provide a more comparable set.

Thirdly, an extended survey, including a significant number of participants from cities, towns, and villages in different provinces throughout PNG, and collecting responses from people of different age, region of origin, gender, and level of education would allow statistical comparison for different demographic variables.

Finally, an adjusted scale. The ten-point scale used in the pilot exercise discussed here had no clear mid-point. Though a five-point scale might not allow for enough different responses, a seven-point could be sufficient. The number 'scores' could be associated with native-speaker categorisations, such as *tru tru* 'pure/real', *bigpela* 'correct', *liklik* 'casual', and *haphap* 'mixed', for example.

In planning further research, the quantitative data collected from the Listening Exercise could be used to identify individuals and groups from futher qualitative analysis.

The framework for analysis and the value of an exercise of this type in collecting feedback from speakers has been established. These modifications to the research design would

improve on the initial pilot. The question of the sociolinguistic significance of different types of mixed language in PNG today remains to be thoroughly investigated.

9.3.2 Conclusions from the Pilot Exercise

While these results are purely indicative, being based on the responses of a highly restricted group of educated bilinguals, the observed correlations between linguistic observations of types of CS in the samples and the responses of these informants are interesting. They suggest that some grammatical borrowing from English, as interpreted under the 4M classification, is less salient for these bilingual informants in diluting the 'purity' of Tok Pisin than behaviour suggesting Congruent Lexicalization. Alternation is highly salient, but there is a high tolerance for Insertional CS by these informants in terms of the scores they gave the exercise samples.

Taken together, the responses to the samples in this listening exercise suggest that mixtures of Tok Pisin and English that appear to involve grammatical and structural changes may not be as salient to all speakers as those that involve clear switches, frequent exchanges, or larger chunks of material. Further investigation may reveal this perceptual gap as a significant factor in the development of new, mixed varieties.

This exercise did not assess the social evaluations of the speech samples or their speakers. It does suggest, however, that a sociolinguistic matched guise experimental assessment of such values in PNG would benefit from the careful selection and control of the types of CS in the selected samples, in addition to any considerations such as accent or pidgin/creole features.

Most importantly, however, the listening exercise suggests that behaviour that linguists consider remarkable may go unnoticed by speakers. As such behaviour includes changes to the constituents permitted by the morphosyntactic 'matrix language frame' (MLF) this may presage a path to widespread grammatical change through language-mixing.

10 Conclusions

This chapter draws together the themes and findings discussed in this dissertation (s10.1) and frames answers to the research questions proposed in Chapter 4 (s10.2). It summarises the value and contributions of the research project (s10.3). It also suggests some conclusions that can be drawn and indicates future lines of research (s10.4).

10.1 Findings on Previous Research

For this study of Tok Pisin, two significant themes of enquiry were identified from the literature (see Chapter 3). In addition, three specific questions were proposed in response to previous research. These two themes and three questions are addressed in this section.

The first research theme related to the role of bilingual behaviour in varieties of Tok Pisin: specifically, whether Urban Tok Pisin could be characterised entirely in terms of CS.

As discussed earlier, Mühlhäusler characterised 'Urban' Tok Pisin in terms of the influence of English as borrowing and interference, and he also highlighted CS as indicating that a continuum of varieties between Tok Pisin and English had not developed. Subsequent research has distinguished different types of CS and their roles in borrowing, and has allowed a categorisation of 'interference'.

While strict Alternation between languages allows "code-switching to be pinpointed" (Mühlhäusler 1985d: 148), some types of CS within the sentence are harder to characterise. Phonologically and morphosyntactically integrated L2 forms raise significant questions about the distinction between CS and borrowing, and make 'pinpointing' CS more problematic. Furthermore, the presence of L2 system morphemes violates various proposed constraints, including the MLF model.

Interpreting such CS forms as produced by 'Congruent Lexicalization' based on a shared grammatical system actually means that specific CS cannot be "pinpointed": it is distributed

across the surface forms of the speech act and licensed by congruent structures at the deeper level of grammar.

Distinguishing 'Rural' from 'Urban' varieties of Tok Pisin on the basis of the locality of residence has been called into question by regional studies (Romaine 1992; Smith, G. 2002). My Tok Pisin data show more than simple lexical borrowing or minor functional and restricted phonological borrowing. In terms of Thomason and Kaufman's (1988: 74-76) 'Borrowing Scale' the Tok Pisin of these speakers falls somewhere between type 3 (slightly more structural borrowing as a result of more intense contact) and type 4 (moderate structural borrowing as a result of strong cultural pressure). However, the mechanisms of this 'borrowing' were not investigated by Thomason and Kaufman.

Our analysis of phonology together with the appearance of L2 morphological affixes and other system (function) words in both Tok Pisin and PNG English appears to indicate that these grammatical changes can all be considered as developing through CS. Informed by studies of bilingual behaviour outside of pidgin/creole communities, we can appreciate 'Urban' Tok Pisin as a variety that has developed through specific types of CS: unmarked Alternation, Insertion, and Congruent Lexicalization.

The second theme of the research related to the concepts of decreolization and the development of a pidgin/creole continuum: specifically, whether CS could have a role in creating a continuum in Papua New Guinea.

This is indicated by the literature. Both Muysken (2000) and Auer (1999) indicate how continua of CS behaviour can pattern Thomason and Kaufman's 'Borrowing Scale': as community bilingualism moves from Alternation to Insertion, from marked to unmarked CS, and then from congruent grammatical systems to regularised grammatical roles for different elements of each language in a new 'fused lect', so the appearance of L2 elements 'borrowed' into a language develops along Thomason and Kaufman's scale. As Auer observes: "a

bilingual community may stabilize on a certain point on the continuum" (Auer 1999: 329). Alternatively, as Myers-Scotton (1993a, 1996) suggests, speakers may "turnover" the matrix language they are using and approach CS from the direction of the L2. With these findings of bilingualism leading to restructuring and changes in the L1 of many communities over time, the second theme of this research, whether CS could lead to the development of a pidgin/creole continuum in PNG, gains traction.

Different types of CS between Tok Pisin and English have been identified in my data. Appreciating the relationship between types of CS and degrees of language mixing calls into question Mühlhäusler's (1985d: 148) position that CS indicates a separation of languages. Convergence (Chapter 5) facilitates and is encouraged by CS, and may produce stable intermediate varieties that create a Tok Pisin/PNG English continuum. Alternation (Chapter 6) keeps languages separate, while Insertion (Chapter 7) fits elements of one language into the structure of another. Congruent Lexicalization (Chapter 8) can include active English system morphemes in Tok Pisin, or vice versa, and is based on an intermediate grammar. Whether this intermediate grammar is stable and regular enough to be considered truly mesolectal remains, following on from this research, an open question. Tok Pisin/English bilinguals do not consistently associate all types of CS with mixed language (Chapter 9), and mixing remains salient. There is, however, a possibility that continuing data collection and CS analysis might actually lead to evidence in support of Bickerton's continuum hypothesis (Bickerton 1975a: 24).

In addition to the general themes illuminated by my research, findings on three questions emerging from the literature were made.

From Pidgin and Creole studies came two questions. First, the role of creolization and expansion as processes involved in the regularisation of a pidgin language was raised. Tok Pisin, we found, was most accurately characterised as an expanded pidgin. Grammatically, it

can be termed a 'creole' language. Despite remaining predominantly a second language, its role for many Papua New Guineans is as a fully functioning language of wider communication, and it is often their primary language in day-to-day life.

My research found that children's developments recorded by Romaine (1990, 1992, 1999) and G.Smith (1990, 2002) in the 1980s are common features of adult speech in the early 21st century. In this way this dissertation records the continuation of use for new forms of language.

Second, the question of decreolization was at the heart of this study: with extensive bilingualism with English becoming the norm for many Tok Pisin speakers, would they adopt English forms and so change their language?

It was found that English used in Tok Pisin could either be characterised in terms of deliberate Alternations (Chapter 6), indicating a clear differentiation of the two languages, or the Insertion of English into Tok Pisin sentence structures (Chapter 7): indicating that the structure of Tok Pisin was not changing specifically to become more like English. Where Congruent Lexicalization was indicated (Chapter 8), it could not be argued that this represented decreolization: this type of CS had a role in PNG English as much as it did in Urban Tok Pisin.

Thirdly, the long-standing question about the development of a pidgin/creole continuum (also discussed below s10.2.1) was addressed in terms of this typology of CS.

It was found that congruence between PNG English and Tok Pisin (Chapter 5) and the presence of Congruent Lexicalization (Chapter 8) provide preliminary indications that structured, intermediate varieties may be emerging from CS in the community. Feedback from bilingual Tok Pisin/English speakers (Chapter 9) indicated that some language mixing with structural impacts was not highly salient, suggesting that some types of CS were not perceived as affecting the language as strongly as others.

From the CS literature came an additional question: about the types and consequences of CS between a creole and its superstrate. Individual switches may be locally meaningful or may be unmarked: the presence of unmarked switching between a creole and its superstrate suggests the erosion of diglossia and the distinction between languages. Cases of CS between sentences do not require structural explanations. CS within sentences, however, may be based on a single 'matrix' language, or on a 'composite' or 'congruent' shared matrix: intermediate grammatical structures are 'mesolectal' in many discussions of creole languages.

Marked and unmarked, inter- and intra-sentential: all of these types of CS were found in the Tok Pisin/English data of the core corpus. The Matrix Language Frame (MLF) model was shown to have explanatory power in interpreting much intra-sentential use of English in Tok Pisin, and vice versa. However, not all 'anglicised' Tok Pisin could be understood in terms of a single matrix language. We saw evidence of Insertional CS in PNG English as well as in Tok Pisin, which indicated the possibility for some speakers of Myers-Scotton's matrix language "turnover" (Myers-Scotton 1993a, 1996). This is now associated with "composite CS" in an explicit engagement with Muysekn's typology (Myers-Scotton 2002). Significantly, Muysken's (2000) Congruent Lexicalization of a congruent Tok Pisin/English grammar is a viable interpretation for some data. Overall, we can characterise 'Urban Pidgin' as LM under Auer's (1999) typology of mixed languages.

The following section summarises the specific findings of this research.

10.2 Project Findings : Conclusions

Investigating 'the state of Tok Pisin' twenty-five years after Independence, this research found evidence of many different English elements being used in Tok Pisin. The extent to which this represented borrowing, was the result of decreolization, or could be explained as instances of bilingual behaviour was the subject of the analyses in the previous chapters of this dissertation.

We have seen evidence for all three structural types of CS behaviour: Alternation (Chapter 6), Insertion (Chapter 7), and Congruent Lexicalization (Chapter 8). We have also found that strict diglossia no longer holds in Papua New Guinea, and that different individuals and different communities have different and variable norms influencing their use of language. Not least among these are the eloquent rhetorical flourishes of bilingual speakers.

In addition to strict Alternation, we have seen examples of both locally meaningful, marked, combinations of Tok Pisin and English ('Codeswitching' (Auer 1999)), and also frequent unmarked switching ('Language Mixing' (ibid.)). The latter, for educated Tok Pisin/English bilinguals, seems to represent a default choice: it is the normal mode of communication and characterises the 'Urban' sociolect.

The findings have also shown examples of the Insertion of English content into Tok Pisin: not only individual words, but also compounds and whole EL constituents. In addition, we have seen the prominent presence of English conjunctions and discourse markers in the Tok Pisin of the core corpus, and also of English pronouns. Morphological affixes from English have been shown to be inactive in speech, being supported by their equivalents from the Tok Pisin morphosyntactic 'frame'. Other L2 system morphemes such as determiners, the copula, and the predicate marker have also been identified in mixed language: indicating a violation of the 4M classification at the heart of the MLF model, and suggesting Congruent Lexicalization underlies at least some speakers' language.

Though this analysis has identified extensive and intricate CS in the data, there is insufficient data to identify regularised forms in the community as a whole. The consistent use of English conjunctions and discourse markers in specific, fixed, grammatical roles in Tok Pisin, and, indeed, of Tok Pisin conjunctions and discourse markers in PNG English, is not established by this data. Consequently, we cannot describe a Tok Pisin/English 'fused lect' (Auer 1999).

The analysis of responses to the Listening Exercise (Chapter 9), however, does indicate that different types of language mixing are salient for educated Papua New Guineans.. There is evidence that Alternation is clearly distinguished from Insertion as language mixing. Fascinatingly, while Tok Pisin system morphemes in English seemed salient, the UPNG student groups classified Tok Pisin with well-integrated English discourse markers and conjunctions, and some system morphemes, as comparably less mixed. In this 'elite' community there appears a high tolerance for well-integrated English elements in speech they characterised as 'Tok Pisin'.

In greater detail than the general findings discussed above (s10.1), the findings and answers to our 3 specific research questions are summarised in s10.2.1-s10.2.3. Additional findings from this research are summarised in s10.2.4.

10.2.1 CS and Creole Continuum

As discussed in Chapter 3, the question of whether Tok Pisin and English in Papua New Guinea remain separate or form a continuum was explicitly addressed by Bickerton (1975a), and has been reviewed by Mühlhäusler (1985d), Siegel (1997, 2008), and G.Smith (2000, 2002).

My Tok Pisin/English CS data shows marked Alternation, and this is an indicator of separate systems (Auer 1998b: 13) (see also s3.2.4.2 above). The data also shows unmarked switching and Insertions, including core 'borrowing', which indicate that this separation is under pressure. Indications of Congruent Lexicalization make it clear that the languages are not separate for all speakers at all times. Romaine's "intermediate varieties" (Romaine 1992: 321) (see s3.4.3) are best understood as examples of Congruent Lexicalization.

As Mühlhäusler (1985d: 80) observed, 'anglicised' Tok Pisin need not resemble English any more closely than Rural Tok Pisin. As the "unmarked code" (Myers-Scotton 1993a) for

educated bilinguals, however, it does represent an intermediate variety of behaviour, one structured by the congruence between changes in expanded Tok Pisin and in PNG English.

The core corpus data has shown that congruence between Tok Pisin and PNG English is high, and throughout my dataset the separation of the two languages is in some cases unclear (see especially Chapter 9).

This project proceeded from the hypothesis that Tok Pisin and English were separate languages in PNG, that a CS analysis would explain variation in Tok Pisin, and that a smooth continuum was not present. However, while no conclusive evidence for a community-wide continuum is presented here, it is clear in some cases that intermediate constructions exist. There is strong support from this analysis that 'anglicised' Tok Pisin is characterised by CS, though some cases of CS are specific to congruent systems. These congruent grammars evolve not through the development (expansion or reduction) of one language alone, but through the dynamic adaptation of both Tok Pisin and English, and it is this bilingual process that may produce the mesolect.

10.2.2 CS and Borrowing

The question of the relationship between CS and borrowing in the processes of change of Tok Pisin is one that this research has not been able to answer conclusively. A position was taken (see esp. s7.2) to include all cultural English material in Tok Pisin as CS forms, except where documented in Mihalic's pre-Independence dictionary (Mihalic 1971), since investigating monolingual usage of particular words was beyond the scope of this study (see also s3.2.3).

However, the analysis of mixed constituents (s7.3.4) has shown, in contrast to earlier research, that English morphology present in Tok Pisin is not necessarily grammatically active. The indications of Congruent Lexicalization (Chapter 8) indicate that 'borrowed' system morphemes of other types may be dependent on bilingual grammars. In addition, the

presence of English words replacing Tok Pisin words ('core' borrowing) was seen as significant.

The presence of various different elements from English in Tok Pisin indicates that cultural contact between English and Tok Pisin is increasing, as per Thomason and Kaufman's (1988: 74-76) 'Borrowing Scale' (see also s10.1 above). Applying this scale to data from a future, more wide-ranging study of mixed language in different communities may enable a characterisation of those communities in terms of degree of contact.

10.2.3 Types of CS in Tok Pisin

Evidence of all three of Muysken's (2000) types of bilingual behaviour has been found in the data of the core corpus. It is clear that Tok Pisin/English CS is varied and complex. 'Urban Pidgin' can be characterised as an LM variety (Auer 1999). The presence of different types of CS behaviour in the speech community creates the preconditions for intermediate varieties: either through the regularisation of forms produced in Congruent Lexicalization, or the roles of L2 grammatical morphemes in the formation of a fused lect. Additional data and analysis is required before we can assess the extent of such mixed varieties of Tok Pisin.

10.2.4 Additional Findings

Beyond the specific questions relating to the structural analyses and pidgin and creole and code-switching theory, this research has produced two other significant findings.

10.2.4.1 Hierarchy of Salience

The Listening Exercise experiment described in Chapter 9 showed that different types of CS must be attended to in any characterisation of Tok Pisin. Different types of CS are perceived as affecting the 'mix' of Tok Pisin and English in different ways. It is possible that different communities may respond to Alternation and to different types of constituent Insertion and

Congruent Lexicalization in varied ways. There are indications that low-salience CS can open the way for structural borrowing and linguistic change.

Extending this exercise to other informant groups in the future could provide another measure of community standards.

10.2.4.2 Mixed Language and Identity

The remarkable presence of unmarked CS in the data of the core corpus, and the role of Alternation, Insertion, and Congruent Lexicalization in conversational 'Urban' Tok Pisin and political speeches, reinforces Mühlhäusler's assertion that the use of English in Tok Pisin is a marker of urban identity (Mühlhäusler 1985d: 140). Following G.Smith's (2002) observations we modify this to suggest that using English with Tok Pisin is a marker of a 'modern, educated' identity. It is also clear, however, that PNG English has a flavour and character that distinctively expresses a 'National' identity, partly through its congruence with Tok Pisin.

One consequence of the use of English as the language of instruction in schools has been an expansion of the community of Papua New Guineans familiar with that language. The 2000 National Census data (see s3.1.4 esp. Figure 3.1) indicates that the proportion of rural people under 30 years of age who report literacy in both Tok Pisin and English is at the same level as for the urban population under 60. This reflects rural bi-literacy in 2000 at a level equivalent to urban bi-literacy at the time G.Sankoff was working in Buang (Sankoff, G. 1977, 1979, 1980d). These figures show the growth of a community with new educational and linguistic resources across both urban and rural populations.

This research shows that CS is the significant linguistic articulation of that identity. It may be the case that individual negotiations of status can be read through the types of CS speakers employ. This, however, was beyond the scope of the present study.

10.3 Summary of the Research

This research has established the value of an analysis of 'Urban Pidgin' using a typology of CS. Such analysis can inform future interpretations of variation in Tok Pisin, as well as the role of English 'borrowings' in the development of the language.

By focusing on adult, urban, bilingual speech this dissertation has characterised the expressive and dynamic language of a growing section of the Papua New Guinean people, a group whose influence and authority are central in the development of this nation.

The dataset collected by this researcher is unique and accessible. It has been coded in accordance with international standards for mixed language corpora, and digital copies of the original recordings are of archival quality. It reflects a time and contexts for which limited recordings of primary spoken language are available. In order to give due weight to this collection of recorded bilingual speech a coding and storage scheme was devised with the potential to be transferred to additional datasets. Most importantly, these recordings from naturalistic settings included higher levels of mixed and intermediate language than in earlier research corpora.

In comparison to earlier work on Tok Pisin, this dataset included high levels of usage of the predicate marker *i* when compared with G.Smith's (2002) findings in regional corpora of children's speech. Considering that the dataset of the core corpus was of mixed language, so that a proportion of the speech was in English, this difference indicates that the Tok Pisin of the core corpus contains significantly more predicate marking than did G.Smith's data. This may be a consequence of the CS involved, or of the aging that has taken place for the post-Independence generation between the time of G.Smith's data collection and my own.

In addition, this research invalidated Laycock's claim that 'anglicised' Tok Pisin was dying out (Laycock 1985a): twenty-five years after Independence, I was able to make extensive recordings of such varieties, which continue to develop today.

The listening exercise discussed in Chapter 9 represents a novel contribution: developing a methodology to collect quantitative feedback on mixed language from native speakers. Its significance lies in that it not only provides information on the salience of different types of mixed language for the UPNG students who participated in the exercise, thus providing insights into subliminal licensing of language change, but also establishes the value and viability of very simple experimental methodologies in collecting information on language intuitions in PNG. It is hoped that an extension of this and similar experiments to communities across the country will improve our understanding of variation in Tok Pisin in terms of how it is perceived by Papua New Guinean speakers.

Overall, this research has cleared the way for a more comprehensive study of variation in Tok Pisin by providing a framework for the analysis of interactions between Tok Pisin and English on the basis of a typology of bilingual behaviours. It has highlighted specific behaviours that, should they become widespread, would indicate the formation of established intermediate lects.

Ultimately the notion of CS itself is a construct of analysis (Gardener-Chloros 1995: 86): it is an intellectual technique, a construct used to divide and comprehend the continual variety of human speech, just as 'languages' themselves are constructs. As this research has shown, separating CS from 'borrowing', 'interference' or the processes of development of pidgins/creoles is largely a question of definitions. Nevertheless, demonstrating that 'anglicised' Tok Pisin can be understood in terms of CS establishes the agency of Papua New Guinean Tok Pisin/English bilinguals in the transformation, development, or maintenance of their national language.

10.4 Suggestions for Further Research

It is hoped that the analytic framework established by this research will be used as a further stepping stone for research into the rich and dynamic linguistic space of Papua New Guinea.

A theoretical, methodological, and analytical basis has been laid here for research on CS between a creole and superstrate and on the overlap between CS typologies and creole continua. This typology of CS has been shown to distinguish varieties of Tok Pisin based on individual bilingual behaviour from those that may develop into regularised widespread norms.

In this century, the practical obstacles of recording and storing large volumes of natural language data have been swept aside by advances in digital technology. The transcription, coding, and analysis of such data, however, remains the main time-consuming task of any researcher. It is hoped that the framework described here can be followed by future fieldworkers in the development and expansion of more substantial corpora.

Further research into urban varieties of Tok Pisin in different communities, and the collection of an even larger dataset of natural language would open up this research on the national scale.

This dataset could be used to assess whether the comparably high levels of Tok Pisin grammatical morphology found in the core corpus of mixed Tok Pisin/English speech was characteristic of mixed language, or whether it was a feature of adult, conversational, or bilinguals' language, as opposed to the language of children discussed in, for example, G.Smith's dataset (Smith, G. 2002).

In addition, such further research may draw sociolinguistic connections between types of mixing, social and educational factors, and the articulation of identity. Aspects of the social and educational development of different communities can be assessed on this basis. In turn, assessments of different types of mixed language by different groups of informants (as per Chapter 9) can be used to understand differences in the transparency and efficacy of different varieties of Tok Pisin, which in turn could be used as an evidence base for educational and language policy in Papua New Guinea.

At the international level, this method of linguistic analysis can be extended to other pidgin/creole communities to explore the dynamics between these languages and their lexifiers, and so contribute to on-going debates on changes in these languages and creole continua in general.

In summary, this work provides an important new investigation into the dynamics of pidgin/creole/superstrate language communities in general, and a new perspective on Tok Pisin and English in Papua New Guinea in particular. In addition, it contributes to contemporary studies of bilingualism, and to the growing literature and data on code-switching behaviour.

REFERENCES:

Aceto, M. (1999). Looking beyond decreolisation as an explanatory model of language change in creole speakings communities. *Journal of Pidgin and Creole Languages* **14**(1): 93-121.

Adler, R. (1975). Olsem wanem Tok Pisin i ken helpim ol manmeri long Papua Niugini? In K. A. McElhanon (ed.), *Tok Pisin i go we?* Port Moresby: Linguistic Society of Papua New Guinea. **Kivung Special Publication Number One:** 121-126.

Ahai, N. (1984). Social Networks and Urban New Guinea Pidgin (Tok Pisin). *York Papers in Linguistics* **11**: 9-18.

Aitchison, J. (1981). *Language Change: Progress or Decay*. Bungay, Suffolk: Fontana Paperbacks.

Alfonzetti, G. (1998). The conversational dimension in code-switching between Italian and a dialect in Sicily. In P. Auer (ed.), *Code-switching in conversation*. London, New York: Routledge: 180-214.

Alleyne, M. (1971). Acculturation and the cultural matrix of creolization. In D. H. Hymes (ed.), *The pidginization and creolization of languages*. 1968: 169-186.

Anderson, B. R. O. G. (1983). *Imagined communities : reflections on the origin and spread of nationalism*. London: Verso.

Auer, P. (1984). *Bilingual Conversation*. Amsterdam: John Benjamins.

Auer, P. (1995). The pragmatics of code-switching: a sequential approach. In L. Milroy and P. Muysken (eds.), *One Speaker, Two Languages*: 115-135.

Auer, P., ed. (1998a). *Code-switching in conversation : language, interaction and identity*. London ; New York: Routledge.

Auer, P. (1998b). Introduction: Bilingual Conversation Revisited. In P. Auer (ed.), *Code-switching in Conversation*. Amsterdam: John Benjamins: 1-24.

Auer, P. (1999). From codeswitching via language mixing to fused lects: Toward a dynamic typology of bilingual speech. *International Journal of Bilingualism* **3**(4): 309-332.

Azuma, S. and C. Myers-Scotton (1990). *A Frame-Based Model of Codeswitching*. Proceedings from the 26th Regional Conference, held at Chicago: CLS.

Bailey, B. L. (1971). Jamaican Creole: Can Dialect Boundaries be Defined? In D. H. Hymes (ed.), *Pidginization and Creolization of Languages*. Cambridge: CUP: 341-348.

Bailey, G. and N. Maynor (1998). Decreolization? In P. Trudgill and J. Cheshire (eds.), *The Sociolinguistics Reader*. London: Arnold: 240-262.

Bailey, R. W. and M. Görlach, eds. (1984 [1982]). *English as a world language*. Cambridge: Cambridge University Press.

Baing, S., B. Deutrom, et al. (2008). *Papua New Guinea Tok Pisin English Dictionary*. South Melbourne: Oxford University Press.

Baker, P. (1993). Australian influence on Melanesian Pidgin English. *Te Reo* **36**: 3-67.

Bakker, P. (1997). *A Language of our own. The genesis of Michif, the mixed Cree-French language*. New York: Oxford University Press.

Bakker, P. (2006). Pidgins versus creoles and pidgincreoles. In S. Kouwenberg and J. V. Singler (eds.), *Handbook of pidgin and creole studies*.

Bakker, P. and M. Mous, eds. (1994). *Mixed languages. Fifteen case studies in language intertwining*. Amsterdam: IFOTT.

Barron, C. (1986). Lexical Nativisation in Papua New Guinean English. *Research Report*. Lae: Department of Language and Communication Studies.

Bender, E. M. (2004). *On the Boundaries of Linguistic Competence: Matched-Guise Experiments as Evidence of Knowledge of Grammar*. Stanford: Stanford University.

Benítez-Torres, C. (2009). *Inflectional vs. Derivational Morphology in Tagdal: A Mixed Language*. 38th Annual Conference on African Linguistics, held: Cascadilla Proceedings Project.

Benítez-Torres, C. (2011). The pragmatics of Berber and Songhay vocabulary in Tagdal: A Northern Songhay language. *Spciety of Pidgin and Creole Linguistics Summer Conference*. University of Accra, Ghana: unpublished.

Bentahila, A. and E. E. Davies (1995). Patterns of codeswitching and patterns of language contact. *Lingua* **96**: 75-93.

Bickerton, D. (1973). On the Nature of a Creole Continuum. *Language* **49**: 640-669.

Bickerton, D. (1975a). Can English and Pidgin Be Kept Apart? In K. A. McElhanon (ed.), *Tok Pisin i go we?* Port Moresby: Linguistic Society of Papua New Guinea: 21-27.

Bickerton, D. (1975b). *Dynamics of a creole system*. London ; New York: Cambridge University Press.

Bickerton, D. (1981). *Roots of Language*. Ann Arbor: Karoma.

Bickerton, D. (1984). The language bioprogram hypothesis. *The Behavioural and Brain Sciences* **7**: 173-188.

Bickerton, D. (1991). On the supposed "Gradualness" of creole development. *Journal of Pidgin and Creole Languages* **6**(1): 25-58.

Bickerton, D. (1992). The sociohistorical matrix of creolization. *Journal of Pidgin and Creole Languages* **7**(2): 307-318.

Bickerton, D. (2010). Creoles, Capitalism and Colonialism. In R. Cohen and P. Toninato (eds.), *The Creolization Reader: studies in mixed identities and cultures*. Abingdon, Oxon: Routledge: 39-48.

Blom, J.-P. and J. J. Gumperz (1972). Social meaning in linguistic structure: Code-switching in Norway In J. J. Gumperz and D. H. Hymes (eds.), *Directions in Sociolinguistics*. New York: Holt, Rinehart and Winston: 407-434.

Blommaert, J. (2010). *The sociolinguistics of globalization*. Cambridge: CUP.

Bloomfield, L. (1935). *Language*. London: Allen and Unwin.

Bock, K. and W. Levelt (1994). Language production: grammatical encoding. In M. A. Gernsbacher (ed.), *Handbook of Pyscholinguistics*. New York: Academic Press: 945-984.

Boeschoten, H. (1998). Codeswitching, codemixing, and code alternation: What difference. In R. Jacobsen (ed.), *Codeswitching worldwide*. Berlin: Mouton de Gruyter: 15-24.

Bokamba, E. (1988). Code-Mixing, Language Variation, and Linguistic Theory: Evidence from Bantu Languages. *Lingua* **76**: 21-62.

Bolton, K. and B. B. Kachru, eds. (2003). *World Englishes: Critical Concepts in Linguistics*. Abingdon: Routledge.

Brash, E. (1975). Tok Pisin! *Meanjin Quarterly* **34**(3): 320-327.

Britto, F. (1986). *Diglossia: A study of the theory with application to Tamil*. Washington D.C.: Georgetown University Press.

Chaudenson, R. (1992). *Des îles, des hommes, des langues: essais sur la créolisation linguistique et culturelle*. Paris: L'Harmattan.

Chaudenson, R. (2001). *Creolization of language and culture*. London: Routledge.

Cheshire, J. (1982). *Variation in an English dialect: A sociolinguistic study*. Cambridge: Cambridge University Press.

Cheshire, J., ed. (1991). *English around the world: sociolinguistic perspectives*. Cambridge: Cambridge University Pres.

Churchill, W. (1911). *Beach-La-Mar : the jargon or trade speech of the western pacific*. Washington: Carnegie Institution of Washington publication.

Clyne, M. (1967). *Transferrence and triggering: Observations on the language assimilation of postwar German speaking migrants in Australia*. The Hague: Nijhoff.

Clyne, M. (1997). Some of the things trilinguals do. *The International Journal of Bilingualism* **1**: 95-116.

Cohen, R. and P. Toninato, eds. (2010). *The Creolization Reader: studies in mixed identities and cultures*. Routledge Student Readers. Abingdon, Oxon: Routledge.

Crowley, T. (1990). *Beach-La-Mar to Bislama : The Emergence of a National Language in Vanuatu*. Oxford: Clarendon Press.

Crowley, T. (1990a). Serial verbs and prepositions in Bislama. In J. W. M. Verhaar (ed.), *Melanesian Pidgin and Tok Pisin*. Amsterdam: John Benjamins. **20**.

Dahl, Ö. (1985). *Tense and Aspect Systems*. Oxford: Basil Blackwell.

Davani, J. A. B. (1984). *Current trends and patterns in the use of English amongst students at UPNG*. unpublished B.A. honours thesis dissertation, University of Papua New Guinea.

De Groots, N. (2008). Foreward to Papua New Guinea Tok Pisin English Dictionary. In S. Baing, B. Deutrom and C. A. Volker (eds.), *Papua New Guinea Tok Pisin English Dictionary*. South Melbourne: Oxford University Press.

DeCamp, D. (1971). Towards a generative analysis of a post-creole speech continuum. In D. H. Hymes (ed.), *Pidginization and Creolization of Languages*. Cambridge: CUP: 349-370.

DeGraff, M. (1999a). Creolization, Language Change, and Language Acquisition: An Epilogue. In M. DeGraff (ed.), *Language Creation and Language Change: Creolization, Diachrony, and Development*. Cambridge, Mass: MIT Press.

DeGraff, M., ed. (1999b). *Language Creation and Language Change: creolization, diachrony, and development*. Cambridge, Mass: MIT.

DeGraff, M. (2003). Against Creole Exceptionalism. *Language* **79**(2): 391-410.

Devette-Chee, K. (2011). Decreolization of Tok Pisin: Is there a Tok Pisin - to - English Continuum? *Language & Linguistics in Melanesia* **29**: 95-103.

DiSciullo, A., P. Muysken, et al. (1986). Government and Code-Mixing. *Journal of Linguistics* **22**: 1-24.

Dorney, S. (1990). *Papua New Guinea : people, politics and history since 1975*. Milsons Point, N.S.W.: Random House.

Dueber, D. (2009). 'The English we speaking': Morphological and syntactic variation in educated Jamaican speech. *Journal of Pidgin and Creole Languages* **24**(1): 1-52.

Dutton, T. (1973). *Conversational New Guinea Pidgin*. Canberra: Australian National University.

Dutton, T. (1978). Tracing the pidgin origin of Hiri (or Police) Motu: issues and problems. In S. A. Wurm and L. Carrington (eds.), *Second International Conference on Austronesian Linguistics: proceedings*. Canberra: ANU: 1351-1375.

Dutton, T. (1982). The Melanesian response to linguistic diversity: the Papuan example. In R. J. May and H. Nelson (eds.), *Melanesia beyond diversity*. Canberra: Research School of Pacific Studies, ANU. **I:** 251-261.

Dutton, T. (1983). Birds of a feather: a pair of rare indigenous pidgins from the Gulf of Papua. In Woolford and Washabaugh (eds.), *The Social Context of Creolization*. US: Karoma: 77-105.

Dutton, T. and D. R. Thomas (1985). *A New Course in Tok Pisin (New Guinea Pidgin)*. Canberra: Australian National University.

Eastman, C. M. (1992). Codeswitching as an urban language-contact phenomenon. *Journal of Multilingual and Multicultural Development* **13**(1&2): 1-18.

Edwards, W. (1980). Varieties of English in Guyana: Some comparisons with BEV. *Linguistics* **18**: 289- 309.

Edwards, W. (1983). Code selection and shifting in Guyana. *Language in Society* **12**: 295-311.

Edwards, W. (1984). A community-based approach to the provenance of urban Guyanese Creole. *York Papers in Linguisitcs* **65**: 99-113.

Elias-Olivares, L. (1976). *Ways of speaking in a Chicano community: a sociolinguistic approach*. PhD thesis, University of Texas.

Faraclas, N. (2000). Telephone interview. A. Paliwala. unpublished personal communication: Field Notes.

Ferguson, C. A. (1962). Problems of teaching languages with diglossia. In E. D. Woodworth and R. Di Pietro (eds.), *Georgetown University Round Table of Languages and Linguistics*. Washington D.C.: Georgetown University Press.

Ferguson, C. A. (1971). Absence of copula and the notion of simplicity: a study of normal speech, baby talk, foreigner talk and pidgins. In D. H. Hymes (ed.), *Pidginization and creolization of languages*. Cambridge: CUP: 141-150.

Ferguson, C. A. (1972 (1959)). Diglossia. In P. P. Giglioli (ed.), *Language and Social Context*. London: Penguin Books: 232-251.

Fishman, J. A. (1969). Bilingualism with and without diglossia; diglossia with and without bilingualism. *Journal of Social Issues* **23**(2): 29-38.

Fishman, J. A. (1972a). *Language in sociocultural change*. Stanford: Stanford University Press.

Fishman, J. A. (1972b). *The sociology of language*. Rowley, Mass: Newbury House Publishers.

Fishman, J. A. (1972c). *Sociolinguistics: A brief introduction.* Rowley, Mass: Newbury House Publishers.

Fishman, J. A. (1972d (1962)). Who speaks what language to whom and when? In J. B. Pride and J. Holmes (eds.), *Sociolinguistics: Selected readings*. Baltimore: Penguin Books: 15-32.

Fishman, J. A. (1976). The sociologiy of language: An interdisciplinary social science approach to language in society. In J. A. Fishman (ed.), *Advances in the sociology of language*. The Hague: Mouton. **Col. 1 (2nd ed.; 1st ed. In 1971):** 217-404 :286-299.

Foley, W. A. (1986). *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press.

Foley, W. A. (1988). Language Birth: The Processes of Pidginization and Creolization. In F. Newmeyer (ed.), *The Cambridge Survey: Language*. Cambridge: Cambridge University Press. **iv**.

Foley, W. A. (1997). *Anthropological linguistics : an introduction*. Malden, MA: Blackwell Publishers.

Franklin, K. J. (1990). On the translation of official notices into Tok Pisin. In W. M. Verhaar (ed.), *Melanesian Pidgin and Tok Pisin*. Amsterdam: John Benjamins: 323-344.

Gal, S. (1979). *Language shift : social determinants of linguistic change in bilingual Austria*. New York: Academic Press.

Gardener-Chloros, P. (1991). *Language selection and switching in Strasbourg*. Oxford: Clarendon Press.

Gardener-Chloros, P. (1995). Code-switching in community, regional and national repertoire: the myth of discreteness in lingustic systems. In L. Milroy and P. Muysken (eds.), *One speaker, two languages: inter-disciplinary perspective on code-switching*: 68-89.

Gewertz, D. B. and F. K. Errington (1999). *Emerging class in Papua New Guinea : the telling of difference*. Cambridge, U.K. ; New York: Cambridge University Press.

Giacalone Ramat, A. (1995). Code-switching in the context of dialect/standard language relations. In L. Milroy and P. Muysken (eds.), *One Speaker, Two Languages*: 45-67.

Gillespie, R. (2002). Dating the first Australians. *Radiocarbon* **44**(2): 455-472.

Goodman, M. (1985). Review of Bickerton 1981. *International Journal of American Linguistics* **51**: 109-137.

Goulden, R. J. (1987). *The Melanesian content in Tok Pisin* PhD 2nd draft, University of Toronto.

Gumperz, J. J. (1968). The speech community. *International encyclopedia of social sciences* **9**: 381-386.

Gumperz, J. J. (1971a). Hindi-Punjabi Code-Switching in Delhi (1964). In A. S. Dil (ed.), *Language in Social Groups*. Stanford, California: Stanford University Press: 205-219.

Gumperz, J. J. (1971b). *Language in social groups*. Stanford, Calif.: Stanford University Press.

Gumperz, J. J. (1982). *Discourse Strategies*. Cambridge: CUP.

Gumperz, J. J. and E. Hernandez-Chavez (1971a). Bilingualism, bidialectalism and classroom interaction. In C. Cadzen, J. V. and D. H. Hymes (eds.), *The Functions of Language in the Class-Room*. New York: Teacher's College Press.

Gumperz, J. J. and E. Hernandez-Chavez (1971b). Cognitive Aspects of Bilingual Commmunication. In W. H. Whiteley (ed.), *Language use and social change*. Oxford: Oxfford University Press: 111-125.

Gupta, A. F. (1994). *The Step-Tongue: Children's English in Singapore*. Clevedon, Philaedlphia, Adelaide: Multilingual Matters.

Hall, R. A. (1943). *Melanesian Pidgin English, Grammar, Texts, Vocabulary*. Baltimore: Linguistic Society of America.

Hall, R. A. (1955). *Hands off Pidgin English*. Sydney: Pacific Publications.

Hall, R. A. (1962). The life cycle of pidgin languages. *Lingua* **11**: 151-156.

Harvey, J. (2007). *Tok Pisin on the Internet*. PhD Dissertation, Ball State University.

Haugen, E. (1950). The analysis of linguistic borrowing. *Language* **26**: 210-231.

Haugen, E. (1956). *Bilingualism in the Americas: A bibliography and research guide*: American Dialect Society.

Haugen, E. (1969). *The Norwegian language in America*. Philadephia: The University of Pennsylvania Press.

Healy, L. R. (1975). When is a word not a Pidgin word? In K. A. McElhanon (ed.), *Tok pisin i go we*. Port Moresby: Linguistic Society of Papua New Guinea: 36-42.

Heath, J. (1989). *From Code-Switching to Borrowing: foreign and diglossic mixing in Moroccan Arabic*. London and New York: Kegan Paul International.

Hinrichs, L. (2006). *Codeswitching on the web: English and Jamaican Creole in e-mail communication*. Amsterdam: Benjamins.

Ho, M.-L. and J. T. Platt (1993). *Dynamics of a Contact Continuum: Singaporean English*. Oxford: Clarendon Press.

Holm, J. (1990). Review: Melanesian Pidgin and the Oceanic substrate. By Roger M. Keesing. *Journal of Pidgin and Creole Languages* **5**(1): 147-152.

Holzknecht, S. (1989). Sociolinguistic anaysis of a register: birthday notices in Papua New Guinea *Post Courier*. *World Englishes* **8**(2): 179-192.

Householder, F. W. J. (1962). Greek diglossia. In E. D. Woodworth and R. Di Pietro (eds.), *Georgetown University Round Table on Languages and Linguistics*. Washington D.C.: Georgetown University Press.

Hudson, A. (1994). Diglossia as a Special Case of Register Variation. In D. Biber and E. Finegan (eds.), *Sociolinguistic Perspectives on Register*. Oxford, New York: Oxford University Press: 294-313.

Hudson, C. L. and I.-M. Eigsti (2003). The lexical competence hypothesis: A cognitive account of the relationship between vernacularization and grammatical expansion in creolization. *Journal of Pidgin and Creole Languages* **18**(1): 1-79.

Hymes, D. H. (1967). Models of the interaction of language and social setting. *Journal of Social Issues* **23**(2): 8-28.

Hymes, D. H. (1968). *Pidginization and creolization of languages*. Pidginization and creolization of languages, held at University of the West Indies, Mona, Jamaica: Cambridge University Press.

Hymes, D. H., ed. (1971). *Introduction to Part III, Pidginization and creolization of languages.* Cambridge: Cambridge University Press.

Irvine, A. (2004). A good command of the English language: Phonological variation in the Jamaican acrolect. *Journal of Pidgin and Creole Languages* **19**(1): 41-76.

Jake, J. L. (1994). Intrasentential code switching and pronouns: on the categorical status of functional elements. *Linguistics* **32**: 271-298.

Jake, J. L. and C. Myers-Scotton (1997). Codeswitching and compromise strategies: implications for lexical structure. *The International Journal of Bilingualism* **1**: 25-39.

Jourdan, C. (1989). Nativization and anglicization in Solomon Islands Pijin. *World Englishes* **8**(1): 25-35.

Jourdan, C. and R. M. Keesing (1997). From Fisin to Pijin: Creolization in process in the Solomon Islands. *Language in Society* **26**: 401-420.

Kachru, B. B. (1978). Code-Mixing as a Communicative Stratehy in India. In J. E. Alatis (ed.), *International Dimensions of Bilingual Education.* Washington DC: Georgetown Univ. Press**:** 107-124.

Kale, J. (1990). Controllers or victims: Language and education in the Torres Strait. In R. B. Baudaulf and A. Luke (eds.), *Language Planning and Education in Australiasia and the South Pacific.* Clevedon: Multilingual Matters**:** 107-120.

Kay, P. (1978). Variable Rules, Community Grammar, and Linguistic Change. In D. Sankoff (ed.), *Linguistic Variation.* New York, London: Academic Press**:** 71-82.

Keck, V. (1993). Talk about a changing world. Young Yupno Men in Papua New Guinea Debate their Future. *Canberra Anthropology* **16**(2): 67-96.

Keesing, R. M. (1988). *Melanesian Pidgin and the oceanic substrate.* Stanford, Calif.: Stanford University Press.

Keesing, R. M. (1991). The expansion of Melanesian Pidgin: Further early evidence from the Solomons. *Journal of Pidgin and Creole Languages* **6**(2): 215-229.

Kulick, D. (1992). *Language shift and cultural reproduction : socialization, self, and syncretism in a Papua New Guinean village.* Cambridge [England] ; New York: Cambridge University Press.

Kulick, D. and C. Stroud (1990). Code-switching in Gapun: Social and linguistic aspects of language use in a language shifting community. In W. M. Verhaar (ed.), *Melanesian Pidgin and Tok Pisin.* Amsterdam: John Benjamins**:** 205-234.

Kwamwangamalu, N. M. and C. L. Lee (1991). "Mixers" and "Mixing": English across Cultures. *World Englishes* **10**: 247-261.

Labov, W. (1970). The study of language in its social context. *Studium Generale* **23**: 30-87.

Labov, W. (1971a). The notion of "system" in Creole languages. In D. H. Hymes (ed.), *Pidginization and Creolization of Languages.* Cambridge: CUP.

Labov, W. (1971b). On the adequacy of natural languages. In J. V. Singler (ed.), *Pidgin and creole tense-mood-aspect systems.* Amsterdam/ Philadelphia: John Benjamins**:** 1-58.

Labov, W. (1972). *Sociolinguistic Patterns.* Philadelphia: University of Philadelphia Press.

Labov, W. (1973). The boundaries of words and their meanings. In C.-J. N. Bailey and R. Shuy (eds.), *New ways of analyzing variation in English*. Washington, D.C.: Georgetown University Press: 340-373.

Lalla, B. and J. D'Costa (1990). *Language in Exile: Three Hundred Years of Jamaican Creole*. Tuscaloosa, Alabama: The University of Alabama Press.

Lambert, W. E., H. Giles, et al. (1975). Language attitudes in a French-American community. *Linguistics* **158**(21): 127–152.

Lattey, E. (1979). Beyond Variable Rules. In S. A. Wurm (ed.), *Papers in Pidgin and Creole Linguistics 2*. Canberra: Linguistic Circle of Canberra. **57:** 21-36.

Lawton, D. (1980). Language attitude, discreteness and code shifting in Jamaican Creole. *English World-Wide* **1**: 221-226.

Laycock, D. C. (1966). Papuans and Pidgin: Aspects of Bilingualism in New Guinea. *Te Reo: Proceedings of the Linguistic Society of New Zealand* **9**: 44-51.

Laycock, D. C. (1970). *Material in New Guinea Pidgin (Coastal and Lowlands)*. Canberra: ANU.

Laycock, D. C. (1985a). The future of Tok Pisin. In S. A. Wurm and P. Mühlhäusler (eds.), *The Handbook of Tok Pisin*. Canberra, A.C.T., Australia: Dept. of Linguistics, Research School of Pacific Studies, Australian National University: 665-668.

Laycock, D. C. (1985b). Phonology: Substratum elements in Tok Pisin phonology. In S. A. Wurm and P. Mühlhäusler (eds.), *The Handbook of Tok Pisin*. Canberra, A.C.T., Australia: Dept. of Linguistics, Research School of Pacific Studies, Australian National University: 295-307.

Laycock, D. C. (1985c). Tok Pisin and the census. In S. A. Wurm and P. Mühlhäusler (eds.), *The Handbook of Tok Pisin*. Canberra, A.C.T., Australia: Dept. of Linguistics, Research School of Pacific Studies, Australian National University: 223-231.

Le Page, R. B. (1960). An historical introduction to Jamaican Creole. In R. B. Le Page and D. DeCamp (eds.), *Jamaican Creole: Creole Language Studies 1*. London: MacMillan: 1-124.

Le Page, R. B. (1977). Processes of pidginization and creolization. In A. Valdman (ed.), *Pidgin and Creole Linguistics*. Bloomington IN: University of Indiana Press: 222-255.

Le Page, R. B. and A. Tabouret-Keller (1985). *Acts of identity : Creole-based approaches to language and ethnicity*. Cambridge ; New York: Cambridge University Press.

Lefebvre, C. (1986). Relexification in creole genesis revisted: The case of Haitian Creole. In P. Muysken and N. Smith (eds.), *Substrata versus universals in creole genesis*. Amsterdam: Benjamins: 297-300.

Lefebvre, C. and P. Muysken (1988). *Mixed categories : nominalizations in Quechua*. Dordrecht, Holland ; Boston: Kluwer Academic Publishers.

Leopold, F. (1939-49). *Speech development of a bilingual child; a linguist's record*. Evanston, Ill.: Northwestern University Press.

Levelt, W. (1989). *Speaking from intention to articulartion*. Cambridge MA: MIT Press.

Lewis, M. P., ed. (2009). *Ethnologue: Languages of the World*. Dallas, Tex: SIL International.

LIPPS Group, T. (2000). The LIDES coding Manual: A document for preparing and analysing language interaction data. *The International Journal of Bilingualism* **4**(2): 278.

Lynch, J. (1979). Changes in Tok Pisin Morphology. *13th PNG Linguistic Society Congress*. Port Moresby: Paper.

Maschler, Y. (1998). On the transition from code-switching to a mixed code. In P. Auer (ed.), *Code-switching in conversation*. London, New York: Routledge: 125-150.

McElhanon, K. A., ed. (1975). *Tok pisin, i go we? : proceedings of a conference held at the University of Papua New Guinea, Port Moresby, P.N.G., 18-21 September, 1973*. University of Papua New Guinea: Linguistic Society of Papua New Guinea.

McWhorter, J. H. (1998). Identifying the Creole Prototype: vindicating a typological class. *Language* **74**(4): 788-818.

Melchers, G. and P. Shaw (2011). *World Englishes*. Abingdon: Hodder Education.

Migge, B. (2007). Code-switching a social identities in the Eastern Maroon community of Suriname and French Guiana. *Journal of Sociolinguistics* **11**(1): 53-73.

Mihalic, F. (1971). *The Jacaranda Dictionary and Grammar of Melanesian Pidgin*. Hong Kong: Jacaranda Press Pty Ltd.

Mihalic, F. (1975). Standardisation in Pidgin. In K. A. McElhanon (ed.), *Tok Pisin i go we?* Port Moresby: Linguistic Society of Papua New Guinea: 54-58.

Mihalic, F. (1990). Obsolescence in the Tok Pisin vocabulary. In W. M. Verhaar (ed.), *Melanesian Pidgin and Tok Pisin*. Amsterdam: John Benjamins: 263-274.

Milroy, J. (1992). *Linguistic variation and change: on the historical sociolinguistics of English*. Oxford: Blackwell.

Milroy, L. (1980). *Language and social networks*. Oxford: Blackwell.

Milroy, L. and P. Muysken, eds. (1995). *One speaker, two languages : cross-disciplinary perspectives on code-switching*. Cambridge [England] ; New York: Cambridge University Press.

Mitchell, A. G. (1946). *The Pronunciation of English in Australia*. Sydney: Angus and Robertson.

Moffat, S. (1990). *Becoming bilingual: a sociolinguistic study of the communication of young mother tongue Panjabi-speaking children*. PhD Dissertation, University of Newcastle Upon Tyne.

Mufwene, S. S. (1988). Column: Starting on the wrong foot. *Journal of Pidgin and Creole Languages* **3**(1): 109-117.

Mufwene, S. S. (1994). New Englishes and Criteria for Naming Them. *World Englishes* **13**: 21-31.

Mufwene, S. S. (2000). Creolization is a social, not a structural, process. In I. Neumann-Holzschuh and E. W. Schneider (eds.), *Degrees of restructuring in creole languages*. Amsterdam: John Benjamins: 65-84.

Mufwene, S. S. (2001). *The ecology of language evolution*. Cambridge: Cambridge University Press.

Mufwene, S. S. (2008). What do creoles and pidgins tell us about the evolution of language. In B. Laks, S. Cleuziou, J.-P. Demoule and P. Encrevé (eds.), *In the proceedings of*

the conference on The Origin and Evolution of Languages: Approaches, Models, Paradigms. London: Equinox.

Mufwene, S. S. (2009). Restructuring, hybridization, and complexity in language evolution. In E. O. Aboh and N. Smith (eds.), *Complex Processes in New Languages*. **CLL Volume 35:** 367-400.

Mühlhäusler, P. (1974). *Pidginization and simplification of language*. Canberra: ANU.

Mühlhäusler, P. (1975). Sociolects in New Guinea Pidgin. In K. A. McElhanon (ed.), *Tok Pisin i go we?* Port Moresby: Linguistic Society of Papua New Guinea: 59-75.

Mühlhäusler, P. (1978). *Samoan Plantation Pidgin and the origins of New Guinea Pidgin*. Canberra: Australian National University.

Mühlhäusler, P. (1979a). *Growth and Stucture of the Lexicon of New Guinea Pidgin*. Canberra: Australian National University.

Mühlhäusler, P. (1979b). Sociolects in New Guinea Pidgin. In S. Wurm (ed.), *New Guinea and Neighbouring Areas: A Sociolinguistic Laboratory*. The Hague: Mouton: 225-242.

Mühlhäusler, P. (1980). Structural expansion and the process of creolization. In A. Valdman and A. Highfield (eds.), *Theoretical Orientations in Creole Studies*. London: Academic Press: 19-56.

Mühlhäusler, P. (1984 [1982]). Tok Pisin in Papua New Guinea. In R. W. Bailey and M. Görlach (eds.), *English as a World Language*. Cambridge: Cambridge Uiversity Press: 439-466.

Mühlhäusler, P. (1985a). External History of Tok Pisin. In S. A. Wurm and P. Mühlhäusler (eds.), *The Handbook of Tok Pisin*. Canberra, A.C.T., Australia: Dept. of Linguistics, Research School of Pacific Studies, Australian National University: 35-64.

Mühlhäusler, P. (1985b). Good and Bad Pidgin: nogut yu toktok kranki. In S. A. Wurm and P. Mühlhäusler (eds.), *The Handbook of Tok Pisin*. Canberra, A.C.T., Australia: Dept. of Linguistics, Research School of Pacific Studies, Australian National University: 275-291.

Mühlhäusler, P. (1985c). History of the study of Tok Pisin. In S. A. Wurm and P. Mühlhäusler (eds.), *The Handbook of Tok Pisin*. Canberra, A.C.T., Australia: Dept. of Linguistics, Research School of Pacific Studies, Australian National University: 15-33.

Mühlhäusler, P. (1985d). Internal Development of Tok Pisin. In S. A. Wurm and P. Mühlhäusler (eds.), *The Handbook of Tok Pisin*. Canberra, A.C.T., Australia: Dept. of Linguistics, Research School of Pacific Studies, Australian National University: 75-166.

Mühlhäusler, P. (1985e). Language planning and the Tok Pisin lexicon. In S. A. Wurm and P. Mühlhäusler (eds.), *The Handbook of Tok Pisin*. Canberra, A.C.T., Australia: Dept. of Linguistics, Research School of Pacific Studies, Australian National University: 595-664.

Mühlhäusler, P. (1985f). Syntax of Tok Pisin. In S. A. Wurm and P. Mühlhäusler (eds.), *The Handbook of Tok Pisin*. Canberra, A.C.T., Australia: Dept. of Linguistics, Research School of Pacific Studies, Australian National University: 341-421.

Mühlhäusler, P. (1985g). Tok Pisin Theoretical Issues. In S. Wurm and P. Mühlhäusler (eds.), *The Handbook of Tok Pisin*. Canberra, A.C.T., Australia: Dept. of Linguistics, Research School of Pacific Studies, Australian National University: 443-483.

Mühlhäusler, P. (1985h). Variation in Tok Pisin. In S. A. Wurm and P. Muhlhausler (eds.), *The Handbook of Tok Pisin*. Canberra, A.C.T., Australia: Dept. of Linguistics, Research School of Pacific Studies, Australian National University: 233-273.

Mühlhäusler, P. (1986). *Pidgin and creole linguistics*. Oxford: Basil Blackwell.

Mühlhäusler, P. (1987). Tracing the predicate marker in the Pacific. *English Worldwide* **8**: 97-121.

Mühlhäusler, P. (1990). On the origins of the predicate marker in Tok Pisin. In J. W. M. Verhaar (ed.), *Melanesian Pidgin and Tok Pisin*. Amsterdam: John Benjamins. **20:** 235-250.

Mühlhäusler, P. (1996). *Linguistic ecology : language change and linguistic imperialism in the Pacific region*. London ; New York: Routledge.

Mühlhäusler, P. (1997). *Pidgin and Creole Linguistics*. London: University of Westminster Press.

Mühlhäusler, P., J. A. Bennett, et al. (1979). Some English-based pidins in the southwestern Pacific. In S. Wurm (ed.), *New Guinea and Neighbouring Areas: A Sociolinguistic Laboratory*. The Hague: Mouton: 53-80.

Mühlhäusler, P., T. E. Dutton, et al. (2003). *Tok Pisin Texts: from the beginning to the present*. Amsterdam: John Benjamins.

Mundhenk, N. (1990). Linguistic decisions in the Tok Pisin Bible. In W. M. Verhaar (ed.), *Melanesian Pidgin and Tok Pisin*. Amsterdam: John Benjamins: 345-374.

Muysken, P. (1995). Code-Switching and Grammatical Theory. In L. Milroy and P. Muysken (eds.), *One Speaker, Two Languages*. Cambridge: CUP: 177-198.

Muysken, P. (1997). Alternation, insertion, congruent lexicalization. In M. Pütz (ed.), *Language Choices: Conditions, Constraints and Consequences*. Amsterdam: Benjamins: 361-380.

Muysken, P. (2000). *Bilingual Speech: a typology of code-mixing*. Cambridge: Cambridge University Press.

Myers-Scotton, C. (1986). Diglossia and code-switching. In J. A. Fishman, A. Tabouret-Keller, M. Clyne, B. Krishnamurti and M. Abdulaziz (eds.), *The Fergusonian Impact: In honor of Charles A. Ferguson on the occasion of his 65th birthday*. Berlin, New York: Mouton de Gruyter. **2:** 403-415.

Myers-Scotton, C. (1992). Comparing Codeswitching and Borrowing. *Journal of Multilingual and Multicultural Development* **13**(1&2): 19-40.

Myers-Scotton, C. (1993a). *Duelling languages : grammatical structure in codeswitching*. Oxford : New York: Clarendon Press ; Oxford University Press.

Myers-Scotton, C. (1993b). *Social motivations for codeswitching : evidence from Africa*. Oxford: Clarendon Press.

Myers-Scotton, C. (1995). A Lexically-Based Production Model of Codeswitching. In L. Milroy and P. Muysken (eds.), *One Speaker, Two Languages*. Cambridge: CUP: 233-256.

Myers-Scotton, C. (1996). A way to dusty death: the matrix language turnover hypothesis. In L. Grenoble and L. Whaley (eds.), *Endangered Languages*. Cambridge: Cambridge University Press: 289-316.

Myers-Scotton, C. (1997a). Code-switching. In F. Coulmas (ed.), *The handbook of sociolinguistics*. Oxford: Blackwell: 217-237.

Myers-Scotton, C. (1997b). "Matrix Language Recognition" and "Morpheme Sorting" as Possible Structural Strategies in Pidgin/Creole Formation. In A. K. Spears and D. Winford (eds.), *The Structure and Status of Pidgins and Creoles*. Amsterdam, Philadelphia: John Benjamins. **19:** 152-174.

Myers-Scotton, C. (1997c). Rational actor models and social discourse analysis. In E. R. Pedro (ed.), *Discourse Analysis, proceedings of the 1st international conference on discourse analysis*. Lisbon: Edicoes Colibri and Associacao Portuguesa de Linguistica: 177-199.

Myers-Scotton, C., ed. (1998a). *Codes and consequences : choosing linguistic varieties*. New York: Oxford University Press.

Myers-Scotton, C. (1998b). A theoretical interoduction to the markedness model. In C. Myers-Scotton (ed.), *Codes and consequences: choosing linguistic varieties*. New York: Oxfod University Press: 18-38.

Myers-Scotton, C. (2002). *Contact Linguistics: Bilingual Encounter and Grammatical Outcomes*. Oxford: Oxford University Press.

Myers-Scotton, C. (2006 [1989]). Codeswitching with English: Types of switchin, types of communities. In K. Bolton and B. B. Kachru (eds.), *World Englishes: Critical Concepts in Linguistics*. Abingdon: Routledge. **IV:** 214-233.

Myers-Scotton, C. and J. L. Jake (1995). Matching lemmas in a bilingual language competance and production model: evidence from intrasentential code switching. *Linguistics* **33**: 981-1024.

Myers-Scotton, C. and J. L. Jake (2000). Four types of morpheme: evidence from aphasia, code switching, and second language acquisition. *Linguistics* **38**(6): 1053-1100.

Myers-Scotton, C. and J. L. Jake (2009). A universal model of code-switching and bilingual language processing and production. In B. E. Bullock and A. J. Toribio (eds.), *The Cambridge Handbook of Linguistic Code-switching*. Cambridge: CUP: 336-358.

Myers-Scotton, C., J. L. Jake, et al. (1996). Arabic and constraints on codeswitching. In M. Eid and D. Parkinson (eds.), *Perspectives on Arabic linguistics IX*. Amsterdam: Bejamins: 9-43.

Nagy, N. and M. Meyerhoff (2008). Introduction Social Lives in Language. In N. Nagy and M. Meyerhoff (eds.), *Social lives in language*. Amsterdam: John Benjamins.

Narakobi, B. (1983). *The Melanesian Way*. NZ: University of the South Pacific.

National Statistical Office, C. (1994). Report on the 1990 National Population and Housing Census in Papua New Guinea. Port Moresby: National Statistical Office: 397.

National Statistical Office, C. (2002a). *2000 Census Basic Tables - National Level*. Port Moresby: National Statistical Office.

National Statistical Office, C. (2002b). Papua New Guinea 2000 Census National Report. Port Moresby: National Statistical Office: 99.

Nekitel, O. (1994). Review of *Atlantic Meets Pacifc: A global view of pidginization and creolization*, Francis Byrne and John Holm. *Language and Linguistics in Melanesia* **25**(2): 194-196.

Nekitel, O. (1998). *Voices of Yesterday, Today and Tomorrow, Language, Culture and Identity*. New Delhi: UBS Publishers' Distributors Ltd.

Nekitel, O. (2000). Personal interview. A. Paliwala. unpublished personal communication: Field Notes.

Nettle, D. and S. Romaine (2000). *Vanishing Voices: the extinction of the world's languages*. Oxford: Oxford University Press.

Nishamura, M. (1985). *Intrasentential code-switching in Japanese-English*. PhD, University of Pennsylvania.

Nishamura, M. (1986). Intrasentential code-switching: the case of language assignment. In J. Vaid (ed.), *Language Processing in Bilinguals: Psycholinguistic and Neuropsychological Perspecives*. Hillsdale, NJ: Lawrence Erlbaum Associates: 123-143.

Noel, J. (1975). Legitimacy of Pidgin in the Development of Papua New Guinea toward nationhood. In K. A. McElhanon (ed.), *Tok Pisin i Go We?* Port Moresby: Linguistic Society of Papua New Guinea: 76-84.

O'Connell, J. F. and F. J. Allen (2004). Dating the colonization of Sahul (Pleistocene Australia-New Guinea): A review of recent research. *Journal of Archaeological Science*(31): 835-853.

O'Connell, J. F. and J. Allen (2007). Pre-LGM Sahul (Pleistocene Australia-New Guinea) and the archaeology of early modern humans. In P. Mellars, K. Boyle, O. Bar-Yosef and C. Stringer (eds.), *Rethinking the Human Revolution*. Cambridge: McDonald Institute for Archaeological Research: 395-410.

O'Donnell, W. R. and L. Todd (1980). *Variety in Contemporary English*. London: Allen & Unwin.

Oesch Serra, C. (1998). Discourse connectives in bilingual conversation: The case of an emerging Italian-French mixed code. In P. Auer (ed.), *Code-switching in conversation*. London, New York: Routledge: 101-124.

Paliwala, A. (2003). Three types of creole / superstrate code mixing in Tok Pisin. *SPCL Summer Conference*. University of Hawai'i: unpublished.

Paliwala, A. (2004). 'Re-evaluating the Tri-Lingual Parliament'. *SPCL 2004*. Curacao: unpublished.

Paliwala, A. (2011). Tok Pisin and English in Papua New Guinea – the value of Census data. *SPCL Summer Conference*. University of Ghana, Accra: unpublished.

Paliwala, A. (2012). Language in Papua New Guinea: the Value of Census Data. *Language & Linguistics in Melanesia* **30**: 1-31.

Paliwala, A. (unpublished). Book Review: Geoff Smith 'Growing up with Tok Pisin'. *PACE Newsletter*.

Paolillo, J. C. (2002). *Analyzing Liguistic Variation: statistical models and methods*. Stanford: CSLI.

Patrick, P. (1997). Style and register in Jamaican Patwa. In E. W. Schneider (ed.), *Englishes Around the World 2, Caribbean, Africa, Asia, Australasia*. Amsterdam: The Netherlands: John Benjamins: 41-55.

Patrick, P. (1999). *Urban Jamaican Creole: Variation in the Mesolect*. Amsterdam, Philadelphia: John Benjamins.

Pawley, A. (1974). *On Epenthetic Vowels in New Guinea Pidgin*. Working Papers in Linguistics, held at Hawaii: USP Hawaii.

Pfaff, C. (1976). Functional and structural constraints on syntactic variation in code-switching. In S. B. Steever, C. A. Walker and S. S. Mufwene (eds.), *Papers from the Parasession of Diachronic Syntax*. Chicago: Chicago Linguistics Society: 248-259.

Pfaff, C. (1979). Constraints on language mixing: intrasentential code-switching and borrowing in Spanish/English. *Language* **55**: 291-318.

Piau, J. and S. Holzknecht (1985). Current Attitude to Tok Pisin. In S. A. Wurm and P. Mühlhäusler (eds.), *The Handbook of Tok Pisin*. Canberra, A.C.T., Australia: Dept. of Linguistics, Research School of Pacific Studies, Australian National University: 487-493.

Platt, J. and H. Weber (1980). *English in Singapore and Malaysia - Status: Features : Functions*. Oxford: Oxford University Press.

Platt, J., H. Weber, et al. (1984). *The New Englishes*. London: Routledge & Kegan Paul plc.

Poplack, S. (1980). "Sometimes I'll start a sentence in Spanish y terminol Espanol": Toward a Typology of Code-Switching. *Linguistics* **18**: 581-618.

Poplack, S. (1981). Syntactic Structure and Social Function of Code-Switching. In R. Duran (ed.), *Latino Language and Communicative Behaviour*. Norwood, NJ: Ablex: 169-184.

Poplack, S. (1985). *Contrasting patterns of code-switching in two communities*. Methods V: Proceedings of the V International Conference on Methods in Dialectology, held at Victoria: University of Victoria Press.

Poplack, S. (1987). Contrasting patterns of code-switching in two communities. In E. Wande, J. Anward, B. Nordberg, L. Steensland and M. Thelander (eds.), *Aspects of multilingualism: proceedings from the fourth Nordic symposium on bilingualism, 1984*. Sweden: University of Uppsala: 51-77.

Poplack, S. (1993). Variation theory and language contact. In D. R. Preston (ed.), *American dialect research*. Amsterdam: John Benjamins: 251-286.

Poplack, S. and M. Meechan (1995). Patterns of language mixture: nominal structure in Wolof-French and Fongbe-French bilingual discourse. In L. Milroy and P. Muysken (eds.), *One Speaker, Two Languages*: 199-232.

Poplack, S., S. Wheeler, et al. (1987). Distinguishing language contact phenomena: Evidence from Finnish-English bilingualism. In P. Lilius and M. Saari (eds.), *The Nordic Languages and Modern Linguistics 6*. Helsinki: University of Helsinki Press.

Reed, S. W. (1943). *The Making of Modern New Guinea*. Philadelphia, PA.: American Philosophical Society.

Reyes, R. (1976). Language Mixing in Chicano Bilingual Speech. In J. D. Bowen and J. Ornstein (eds.), *Studies in Southwest Spanish*. Rowley, Mass: Newbury House: 182-188.

Rickford, J. R. (1977). The question of prior creolization in Black English. In A. Valdman (ed.), *Pidgin and creole linguistics*. Bloomington: Indiana University Press: 190-221.

Rickford, J. R. (1979). *Variation in a Creole Continuum: Quantitative and Implicational Approaches*. PhD Dissertation, University of Pennsylvania.

Rickford, J. R. (1986). The need for new approaches to social class analysis in Sociolinguisitcs. *Language and Communication* **6**: 215-221.

Rickford, J. R. (1987). *Dimensions of a Creole continuum : history, texts & linguistic analysis of Guyanese Creole*. Stanford, Calif.: Stanford University Press.

Rickford, J. R. (1999). *African American vernacular English : features, evolution, educational implications*. Malden, Mass.: Blackwell Publishers.

Roberts, S. J. (2004). *The emergence of Hawai'i Creole English in the early 20th century: The sociohistorical context of creole genesis*. PhD PhD Dissertation, Stanford.

Romaine, S. (1988). *Pidgin and Creole languages*. London ; New York: Longman.

Romaine, S. (1989a). *Bilingualism*. Oxford: Basil Blackwell.

Romaine, S. (1989b). English and Tok Pisin (New Guinea Pidgin English) in Papua New Guinea. *World Englishes* **8**(1): 5-23.

Romaine, S. (1990). Change and variation in the use of bai in young children's creolized Tok Pisin in Morobe Province. In W. M. Verhaar (ed.), *Melanesian Pidgin and Tok Pisin*. Amsterdam: John Benjamins: 187-204.

Romaine, S. (1991). The Pacific. In J. Cheshire (ed.), *English around the world*. Cambridge: Cambridge University Press: 619-636.

Romaine, S. (1992). *Language, education, and development : urban and rural Tok Pisin in Papua New Guinea*. Oxford: Clarendon Press.

Romaine, S. (1993). The decline of predicate marking in Tok Pisin. In F. Byrne and J. Holm (eds.), *Atlantic meets Pacific: A global view of pidginization and creolization*. Amsterdam: John Benjamins: 251-260.

Romaine, S. (1994). On the creation and expansion of registers: Sports reporting in Tok Pisin. In D. Biber and E. Finegan (eds.), *Sociolinguistic Perspectives on Register Variation*. Oxford: Oxford University Press: 59-81.

Romaine, S. (1999). Grammaticalization of the Proximative in Tok Pisin. *Language* **75**(2): 322-351.

Romaine, S. (2000). *Language in Society An Introduction to Sociolinguistics*. Oxford, New York: Oxford University Press.

Romaine, S. and F. Wright (1987). Short forms in Tok Pisin. *Journal of Pidgin and Creole Languages* **2**: 63-67.

Salisbury (1967). Pidgin's respectable past. *New Guinea* **2**(2).

Salisbury, R. F. (1962). Notes on Bilingualism and Linguistic Change in New Guinea. *Anthropological Linguistics* **4**(7): 1-13.

Salmons, J. (1990). Bilingual discourse marking: Code switching, borrowing, and convergence in some German-American dialects. *Linguistics* **28**: 453-480.

Sankoff, D., ed. (1978). *Linguistic Variation: models and methods*. New York, London: Academic Press.

Sankoff, D. and S. Laberge (1978). Statistical Dependence Among Successive Occurrences of a Variable in Discourse. In D. Sankoff (ed.), *Linguistic Variation*: 119-126.

Sankoff, D. and S. Poplack (1981). A Formal Grammar for Code-Switching. *Papers in Linguistics: International Journal of Human Communication* **14**(1): 3-46.

Sankoff, D., S. Poplack, et al. (1990a). The case of the nonce loan in Tamil. *Language variation and change* **2**: 71-101.

Sankoff, D., S. Poplack, et al. (1990b). A congruence approach to the syntax of code-switching. *International Journal of Bilingualism* **2**: 1-20.

Sankoff, G. (1977). Creolization and Syntactic Change in New Guinea Tok Pisin. In B. G. Blount and M. Sanches (eds.), *Sociocultural dimensions of language change*. New York: Academic Press Inc: 119.

Sankoff, G. (1979). The genesis of a language. In K. C. Hill (ed.), *The genesis of language*. Ann Arbor: Karoma: 23-47.

Sankoff, G. (1980a). Language Use in Multilingual Societies: Some Alternate Approaches. In G. Sankoff (ed.), *Social Life of Language*. Philadelphia: University of Philadelphia Press: 29-46.

Sankoff, G. (1980b). Multilingualism in Papua New Guinea. In G. Sankoff (ed.), *Social Life of Language*. Philadelphia: University of Philadelphia Press: 95-132.

Sankoff, G. (1980c). Political Power and Linguistic Inequality in Papua New Guinea. In G. Sankoff (ed.), *Social Life of Language*. Philadelphia: University of Philadelphia Press: 5-28.

Sankoff, G., ed. (1980d). *The Social Life of Language*. Philadelphia: University of Pennsylvania Press.

Sankoff, G. (1980e). Variability and Explanation in Language and Culture: Cliticization in New Guinea Tok Pisin. In G. Sankoff (ed.), *Social Life of Language*. Philadelphia: University of Philadelphia Press: 257-270.

Sankoff, G. (1991). Using the future to explain the past. In F. Byrne and T. Huebner (eds.), *Development and structures of Creole languages: essays in honor of Derek Bickerton*. Amsterdam; Philadelphia: John Benjamins: 61-74.

Sankoff, G. and Brown (1980). The Origins of Syntax in Discourse: A Case Study of Tok Pisin Relatives. In G. Sankoff (ed.), *Social Life of Language*. Philadelphia: University of Philadelphia Press: 211-256.

Sankoff, G. and P. Brown (1976). On the origins of syntax in discourse: a case study of Tok Pisin relatives. *Language* **52**(3): 631-666.

Sankoff, G. and P. Kay (1974). A language-universals approach to pidgins and creoles. In D. DeCamp and I. F. Hancock (eds.), *Pidgins and Creoles: Current Trends and Prospects*. Washington: Washington Georgetown University.

Sankoff, G. and S. Laberge (1980 (1973)). On the Acquisition of Native Speakers by a Languages. In G. Sankoff (ed.), *Social Life of Language*. Philadelphia: University of Philadelphia Press: 195-210.

Sankoff, G. and C. Mazzie (1991). Determining noun phrases in Tok Pisin. *Journal of Pidgin and Creole Languages* **6**(1): 1-24.

Savic, J. M. (1995). Structural convergence and language change: Evidence from Serbian/English code-switching. *Language in Society* **24**: 475-492.

Schneider, E. W. (2011). *English around the world: An introduction*. Cambridge: Cambridge University Press.

Schuchardt, H. (1909). Die Lingua Franca. In G. Gilbert (ed.), *Pidgin and Creole Languages: Selected Essays by Hugo Schuchardt*. Cambridge: Cambridge University Press: 65-88.

Sebba, M. (1993). *London Jamaican: A case study in language contact*. London: Longman.

Sebba, M. and T. Wooton (1998). We, they and identity: Sequential versus identity-related explanation in code-switching. In P. Auer (ed.), *Code-Switching in Conversation: Language, interactions, and identity*. London, New York: Routledge: 262-289.

Secretariat, L. a. A. (1995). Literacy and Awareness in Papua New Guinea, 1994. *Annual report of the Literacy and Awareness Secretariat*. Port Moresby: Department of Education.

Siegel, J. (1997). Pidgin and English in Melanesia: Is there a continuum? *World Englishes* **16**(2): 185-204.

Siegel, J. (1998). Review of Verhaar, JWM, Towards a reference grammar of Tok Pisin: An experiment of corpus linguistics. *Journal of Pidgin and Creole Languages* **13**: 172-180.

Siegel, J. (1999). Transfer constraints and substrate influence in Melanesian Pidgin. *Journal of Pidgin and Creole Languages* **14**(1): 1-44.

Siegel, J. (2000). Substrate influence in Hawai'i Creole English. *Language in Society* **29**: 197-236.

Siegel, J. (2004). Morphological simplicity in pidgins and creoles. *Journal of Pidgin and Creole Languages* **19**(1): 139-162.

Siegel, J. (2008). *The Emergence of Pidgin and Creole Languges*. Oxford: Oxford University Press.

Siegel, J. (2010). Bilingual literacy in creole contexts. *Journal of Multilingual and Multicultural Development* **31**(4): 383-402.

SIL, S. I. o. L. Ethnologue. from http://www.ethnologue.com/show_country.asp?name=PG.

Singler, J. V. (1992). Rejoinder: Nativization and pidgin/creole genesis: a reply to Bickerton. *Journal of Pidgin and Creole Languages* **7**(2): 319-333.

Singler, J. V. (1996). Theories of creole genesis, socihistorical considerations, and the evaluation of evidence: The case of Haitian Creole and the Relexification Hypothesis. *Journal of Pidgin and Creole Languages* **11**(2): 185-230.

Singler, J. V. (2006a). Children and creole genesis. *Journal of Pidgin and Creole Languages* **21**(1): 157-173.

Singler, J. V. (2006b). Yes, but not in the Carribean. *Journal of Pidgin and Creole Languages* **21**(2): 337-358.

Slobin, D. (1977). Language change in childhood and history. In J. MacNamara (ed.), *Language Learning and Thought*. New York: Academic Press: 185-214.

Smeall, C. (1975). A quantitative analysis of variation: *i* in Tok Pisin. In C. Cogen and e. al (eds.), *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society*. Berkley: Linguistics Society: 403-409.

Smith, A.-M. (1978). The Papua New Guinea dialect of English. Port Moresby: Educational Research Unit, University of Papua New Guinea: 43.

Smith, A.-M. (1986). *Papua New Guinea English.* PhD Dissertation, University of Papua New Guinea.

Smith, A.-M. (1988a). English in Papua New Guinea. *World Englishes* **7**(3): 299-308.

Smith, A.-M. (1988b). The use of aspect in Papua New Guinea. In B. McCarthy (ed.), *Asian-Pacific Papers occasional papers 10*: Applied Linguistics Association of Australia**:** 109-134.

Smith, G. (1990). Creolized Tok Pisin: Uniformity and variation. In R. Baldauf and A. Luke (eds.), *Language Planning and Education in Australasia and the South Pacific.* Clevedon: Multilingual Matters**:** 197-209.

Smith, G. (1994). Husat i bin othoraizim? New verbs in Manus Tok Pisin,. *Seventh International Conference of the Society for Pidgin and Creole Linguistics.* Georgetown, Guyana.

Smith, G. (1998). English and Melanesian Pidgin in the Admiralty Islands. *Links and Letters* **5**: 107-123.

Smith, G. (2000). Tok Pisin and English: The Current Relationship. In J. Siegel (ed.), *Processes of Language Contact.* Quebec, Canada: Fides**:** 271-293.

Smith, G. (2002). *Growing up with Tok Pisin - Contact, creolization, and change in Papua New Guinea's national language.* London: Battlebridge.

Smith, G. (2004a). Tok Pisin: morphology and syntax. In B. Kortmann, K. Burridge, R. Mesthrie, E. W. Schneider and C. Upton (eds.), *A Handbook of Varieties of English: Volume 2: Morphology and Syntax.* Berlin, New York: Mouton de Gruyter. **2:** 720-741.

Smith, G. (2004b). Tok Pisin: phonology. In B. Kortmann, K. Burridge, R. Mesthrie, E. W. Schneider and C. Upton (eds.), *A Handbook of Varieties of English: Volume 1: Phonology.* Berlin, New York: Mouton de Gruyter. **1:** 710-728.

Smith, N. (1994). An annotated list of creoles, pidgins, and mixed languages. In J. Arends, P. Muysken and N. Smith (eds.), *Pidgins and Creoles.* Amsterdam: John Benjamins.

Smithies, M. and S. Holzknecht (1981). Errors in Papua New Guinea written English at the tertiary level. *RELC Journal* **12**: 10-34.

Soares, P., T. Rito, et al. (2011). Ancient Voyaging and Polynesian Origins. *American Journal of Human Genetics* **Feb 3, 2011 DOI: 10.1016/j.ajhg.2011.01.009.**

Sobin, N. (1976). Texas Spanish and lexical borrowing. *Papers in Linguistics* **9**: 15-47.

Stroud, C. (1998). Perspectives on cultural variability of discourse and some implications for code-switching. In P. Auer (ed.), *Code-Switching in Conversation: Language, interactions, and identity.* London, New York: Routledge**:** 321-348.

Sumbuk, K. (2002). Personal interview. A. Paliwala. unpublished personal communication: Field Notes.

Swigart, L. (1992). Two codes or one? The insiders' view and the description of codeswitching in Dakar. *Journal of Multilingual and Multicultural Development* **13**(1 & 2): 83-102.

Tavurvur (2008). The Absence of "National Identity" Within PNG. *tubuans & dukduks* http://garamut.wordpress.com/2008/10/29/the-absence-of-national-identity-within-png/ 2011.

Thomas, D. R. (2000). Personal interview. A. Paliwala. unpublished: personal communication: Field Notes.

Thomason, S. G. and T. Kaufman (1988). *Language contact, creolization, and genetic linguistics*. Berkeley: University of California Press.

Timm, L. A. (1981). Diglossia old and new – A critique. *Anthropological Linguistics* **23**(8): 356-367.

Todd, L. (1974). *Pidgins and creoles*. London: Keegan Paul.

Traugott, E. C. (1977). Pidginization, creolization, and language change. In A. Valdman (ed.), *Pidgin and creole linguistics*. Bloomington: Indiana University Press: 70-98.

Valdman, A. (1991). Decreolization of dialect contact in Haiti? In F. Bryne and T. Huebner (eds.), *Development and Structures of Creole Languages: Essays in Honor of Derek Bickerton*. Amsterdam: Benjamins: 75-88.

Verhaar, J. W. M. (1991). The function of *i* in Tok Pisin. *The Journal of Pidgin and Creole Languages* **6**: 231-266.

Verhaar, J. W. M. (1995). *Toward a reference grammar of Tok Pisin: an experiment in corpus linguistics*. Honolulu: University of Hawai'i Press.

Wakizaka, M. (2008). *On the issues of language contact and languages shift in Tok Pisin - Focusing on two "non-standard" varieties: Highlands Pidgin and Anglicised Pidgin.* Master of Art by research Dissertation, The University of Adelaide, Australia.

Walczyński, M. (2012). *A Living Language : Selected aspects of Tok Pisin in the press (on the basis of Wantok newspaper)*. Nysa: Publishing Office PWSZ w Nysie.

Walsh, D. S. (1978). Tok Pisin syntax - the East Austronesian factor. In S. A. Wurm (ed.), *Papers in Pidgin and Creole Linguistics No 1*. Canberra: Linguistic Circles of Canberra. **1:** 185-196.

Wei, L. (1998). The 'why' and 'how' questions in the analysis of conversational code-switching. In P. Auer (ed.), *Code-Switching in Conversation: Language, interactions, and identity*. London, New York: Routledge: 156-176.

Weinreich, U. (1953). *Languages in Contact*. The Hague: Mouton.

Williams, J. P. (1993). Documenting the Papuan-based pidgins of insular New Guinea. In F. Bryne and J. Holme (eds.), *Atlantic meets Pacific: A global view of pidginization and creolization*. Amsterdam: Benjamins: 355-367.

Winer, L. (1993). *Trinidad and Tobago*. Amsterdam, The Netherlands: John Benjamins.

Winer, L. (1995). Penny cuts: Differentiation of Creole varieties in Trinidad, 1904-1906. *Journal of Pidgin and Creole Languages* **10**: 125-155.

Winford, D. (1992). Back to the past: The BEV/Creole connection revisited. *Language Variation and Change* **4**: 311-357.

Winford, D. (1993). Variability in the use of perfect have in Trinidadian English. *Language Variation and Change* **5**: 141-187.

Winford, D. (1997). Re-examining Caribbean English creole continua. *World Englishes* **16**(2): 233-279.

Winford, D. (2003). *An introduction to contact linguistics*. Oxford: Blackwell.

Wogan, P. (2004). Deep Hanging Out: Reflections on Fieldwork and Multisited Andean Ethnography. *Identities: Global Studies in Culture and Power*(11): 129-139.

Woolford, E. (1979a). *Aspects of Tok Pisin Grammar*. Canberra: Linguistic Circle of Canberra.

Woolford, E. (1979b). Variation and change in the i 'Predicate Marker' of New Guinea Tok Pisin. In S. A. Wurm (ed.), *Papers in Pigin and Creole Linguistics No. 2*. Canberra: Linguistic Circle of Canberra. **57:** 37-49.

Wurm, S. A. (1977). The Nature of New Guinea Pidgin. In S. Wurm (ed.), *New Guinea Area Languages and Study vol 3*. Canberra: ANU.

Wurm, S. A. (1980). Standardization and Instrumentalisation in Tok Pisin. In A. Valdman and A. Highfield (eds.), *Theoretical Orientations in Creole Studies*. New York: Academic Press.

Wurm, S. A. (1985). The Status of Tok Pisin and attitudes towards it. In S. A. Wurm and P. Mühlhäusler (eds.), *The Handbook of Tok Pisin*. Canberra, A.C.T., Australia Dept. of Linguistics, Research School of Pacific Studies, Australian National University: 65-74.

Wurm, S. A. and P. Mühlhäusler, eds. (1985a). *Handbook of Tok Pisin (New Guinea Pidgin)*. Canberra, A.C.T., Australia: Dept. of Linguistics, Research School of Pacific Studies, Australian National University.

Wurm, S. A. and P. Mühlhäusler (1985b). Introduction. In S. A. Wurm and P. Mühlhäusler (eds.), *The Handbook of Tok Pisin*. Canberra, A.C.T., Australia: Dept. of Linguistics, Research School of Pacific Studies, Australian National University: 3-11.

Yarupawa, S. (1986). Milne Bay Informal Variety of English. *Research Report*. Lae: Department of Language and Communications Studies.

APPENDIX I : Audio Recordings

The following table provides general information on the primary data: audio recordings made during fieldwork in Papua New Guinea. These recordings were made on a range of media: audio cassette tape; MiniDisc; and direct recording to a laptop hard drive. Each separate piece of media has a unique identifier. Texts drawn from those media are identified with similar filenames, with added extensions to indicate time and/or subdivision of the primary recording. Extracts from those texts, such as the examples cited in this dissertation, are located either by the line reference from the transcribed file, or by the time index of the original recording.

Of the total length of recorded media of c8,750 minutes, just under 20% was empty, leaving the 7,000 minutes of recorded audio notes in Chapter 4, Table 4.1.

--

File Name	Date	Contents	Recording Medium	Mins
TiT#01	2000.03.30	A N Y & K play tic tac toe + recordings from my room - Waigani, Port Moresbi	Cassette Tape	90
TiT#02	2000.03.31	small talk outside	Cassette Tape	90
TiT#03a	2000.03.31	small talk on Fridy; translating with K. for CWA booking; P and A talking - Waigani, Port Moresbi	Cassette Tape	90
TiT#03b	2000.03.31	playing GT2	Cassette Tape	90
TiT#04a	2000.03.31	At BN's :erratic sounds, recording cuts out	Cassette Tape	90
TiT#04b	2000.03.31	cont. BN TP/Eng alternation : potential value for Peng discourse; possible useful side between BN and 'Arox'	Cassette Tape	90
TiT#05a	2000.03.31	at BN's cont. Darts At Naing's - plenty of TP/Eng mix, due to presence of Kiwi interlocutor?	Cassette Tape	90
TiT#05b	2000.03.31	cont BN's : loud music interferes with recording A & J spak	Cassette Tape	90
TiT#06a	2000.04.01	Pikinin sport & underneath the house : J & G using TP despite some tok ples; G & A with CS necessary to complete some sentences	Cassette Tape	90

File Name	Date	Contents	Recording Medium	Mins
TiT#06b	2000.04.01	cont underneath the house; GT2; P & Q in background with some CS?	Cassette Tape	90
TiT#07	2000.05.11	HoP debate recorded off the TV	Cassette Tape	90
TiT#08a	2000.04.01	From Room : too quiet	Cassette Tape	90
TiT#08b	2000.04.01	Clear sound of youngester and CS : MN Stori	Cassette Tape	90
TiT#09	2000.04.02	Men and Women talking - quiet. Adam Snores	Cassette Tape	90
TiT#10a	2000.04.06	Ladies calling for Sago, some Jack, then P & J - quiet	Cassette Tape	90
TiT#10b	2000.04.07	YUMI fm : 08:40-09:25 - contrast morning talkback (w Acrolectal features?) with news broadcast (cultural loans only) and Very Quick Woman (poss NS friend)	Cassette Tape	90
TiT#11	2000.04.07	YUMI fm : 13:15-> : - Adverts, news, and talkback	Cassette Tape	90
TiT#12a	2000.02.08	PD & T talking on porch of CWA - Madang Town	Cassette Tape	90
TiT#12b	2000.04.09	Lutheran Church service	Cassette Tape	90
TiT#13a	2000.04.10	M & PD stori likli : Various after visit to DWU : buai men near DWU ; boat business ; & Kranket Island w Diego	Cassette Tape	90
TiT#13b	2000.04.10	A & W talking inside CWA	Cassette Tape	90
TiT#14	2000.04.11	YUMI fm : 06:15 & 07:30 o samting	Cassette Tape	90
TiT#15a	2000.04.11	some fellas (highlands) in the rain on the porch of the CWA	Cassette Tape	90
TiT#15b	2000.04.13	long Kranket Island	Cassette Tape	90
TiT#16a	2000.04.13	A & J (mostly)	Cassette Tape	90
TiT#16b	2000.04.13	TMA with PD & also end of W	Cassette Tape	90
TiT#17a	2000.04.14	TMA with M	Cassette Tape	90
TiT#17b	2000.04.13	TMA with W (ref#16b also)	Cassette Tape	90
TiT#18a	2000.04.15	Mornings @ CWA	Cassette Tape	90
TiT#18b	2000.04.15	Friday Night @ Coasties - meeting the boss	Cassette Tape	90
TiT#19	2000.04.15	Friday Night @ Coasties - the Hospital Guys	Cassette Tape	90
TiT#20	2000.04.16	At Jai Aben : Lapun n Waitskin	Cassette Tape	90
TiT#21a	2000.04.17	In Madang Town	Cassette Tape	90
TiT#21b	2000.04.19	NauFM interview with Moresbi Metropolitan Police Chief @ c.09:30	Cassette Tape	90
TiT#22a	2000.04.19	NauFM talkback CS following interview	Cassette Tape	90

File Name	Date	Contents	Recording Medium	Mins
TiT#22b	2000.04.20	Women waiting for boat from Kranket + A G & C	Cassette Tape	90
TiT#23a	2000.04.20	- @ Marco's Place	Cassette Tape	90
TiT#23b	2000.04.24	Back @ CWA (largely blank)	Cassette Tape	90
TiT#24a	2000.04.24	Tea & Scones @ CWA + YUMI fm 12:00->	Cassette Tape	90
TiT#24b	2000.04.24	Canned Heat taped of YUMI + MSM TSM & A - MSM Stori	Cassette Tape	90
TiT#25	2000.04.27	CWA & A & N and then T R & A (b)	Cassette Tape	90
TiT#26	2000.04.28	A @ Smuggler's & short reel of N & A + Wewak airport	Cassette Tape	90
TiT#27	2000.04.28	Adam & the Mayor : long kikai buai & roun toun & meeting CK and jacket man - Vanimo Town	Cassette Tape	90
TiT#28	2000.04.30	Walk to top town and Sarade Nait Wantaim CK - Vanimo High School	Cassette Tape	90
TiT#29	2000.04.30	To Wutung	Cassette Tape	90
TiT#30a	2000.05.02	Rasta Stori - Lido Village	Cassette Tape	90
TiT#30b	2000.05.02	Kevin & Kaikai	Cassette Tape	90
TiT#31a	2000.05.02	Sounds of the Ples	Cassette Tape	90
TiT#31b	2000.05.03	Sounds of Nambis, na toun	Cassette Tape	90
TiT#32a	2000.05.03	*Bad Tape*	Cassette Tape	90
TiT#32b	2000.05.04	Dinner Time in Lido	Cassette Tape	90
TiT#33a	2000.05.04	After rausim skin bilon buai	Cassette Tape	90
TiT#33b	2000.05.05	Breakfast	Cassette Tape	90
TiT#34a	2000.05.06	K & Naing Younger - Waigani, Port Moresbi	Cassette Tape	90
TiT#34b	2000.05.06	Various Radio & (probably blank)	Cassette Tape	90
TiT#35	2000.05.08	HoP greivance debate	Cassette Tape	90
TiT#36	2000.05.08	Some HoP & TMA with Dino & Q	Cassette Tape	90
TiiT#D1	2000.08.07	Brisbane Equipment Test	MiniDisc	70
TiiT#D2	2000.08.10	Big Rooster (mostly) A & W + some home tok - Waigani, Port Moresbi	MiniDisc	70
TiiT#D3	2000.08.11-13	Goroka Market - Goroka Market again, Bird of Paradise Hotel - Mount Kis	MiniDisc	70
TiiT#D4	2000.08.13	McCarthy Museum & dinner stori @ NSI - Goroka	MiniDisc	70
TiiT#D5	2000.08.15	outside room, olympians, air niugini - then blank!	MiniDisc	70
TiiT#T1	2000.08.15	YUMI fm - Goroka : (a) 22:20-23:07	Cassette Tape	90

File Name	Date	Contents	Recording Medium	Mins
		(b) 23:07-23:56		
TiiT#T2	2000.08.16	YUMI fm - Goroka : (a) 08:18 (b) 10:10	Cassette Tape	90
TiiT#T3	2000.08.16	YUMI fm - Goroka : (a) 11:30 (b) 12:33	Cassette Tape	90
TiiT#T4	2000.08.16	YUMI fm - Goroka : (a) 14:00 (b) 15:23	Cassette Tape	90
TiiT#T5	2000.08.16	YUMI fm - Goroka : (a) 18:50 (b) …	Cassette Tape	90
TiiT#D5	2000.08.15	equals #T5 edited	MiniDisc	70
TiiT#T6	2000.08.27	YUMI fm - 21:00	Cassette Tape	90
TiiT#D6-	-	Various Hagen Show Audio Recordings	MiniDisc	70
TiiT#D8		Bundi - actually Hagen Town and at home	MiniDisc	70
TiiiTBlack Disk	Kokoda Memorial Game, 2002	Miscellaneous recordings from the Kokoda Memorial Game	MiniDisc	70
T3/Md01	2002.12.08	L's Morning (outside - secret)	MiniDisc	70
T3/Md02	2002.12.08	L's Children & Adults (outside - shirt)	MiniDisc	70
T3/Md03	2002.12.09	L's - Boyz lo apinun	MiniDisc	70
T3/Md04	2002.12.10	L's boyz & folks	MiniDisc	70
T3/Md05	2002.12.10 / 11	L's renig11? / 14> Sumbuk	MiniDisc	70
T3/Md06	2002.12.11 / 12	L's - roun lo Kumu	MiniDisc	70
T3/Md07	2002.12.13	L's - secy	MiniDisc	70
T3/Md08	2002.12.14 / 26	N / PS Kidz / Cards	MiniDisc	70
T3/Md09	2002.12.18 / 21	Boroko / Work & J Stori	MiniDisc	70
T3/Md10	2002.12.21 / 22	N Stori / Bubu	MiniDisc	70
T3/Md11	2002.12.23	ELICITATION D + Stori	MiniDisc	70
T3/Md12	2002.12.24	L's - Boyz & Beers (Christmas Eve)	MiniDisc	70
T3/Md13	2003.01.01	The Hill	MiniDisc	70
T3/Md14	2003.01.04	Mumu & Bubu	MiniDisc	70
T3/Md15	2003.01.08	Q @ work	MiniDisc	70
T3/Md16	2003.01.08	Last night l'haus	MiniDisc	70
T3/Hd01	2002.12.28	ELICITATION W	Hard Drive	46
T3/Hd02	2003.01.02	ELICITATION N	Hard Drive	10

File Name	Date	Contents	Recording Medium	Mins
T3/Hd03	2003.01.03	ELICITATION J / also visitor	Hard Drive	6
T3/Hd04	2003.01.06	ELICITATION Q	Hard Drive	27
T3/#T1	2002.12.30	YUMIfm - A@ 11:50	Cassette Tape	90
T3/#T2	2002.12.31	YUMIfm top100 - A@ 12:00	Cassette Tape	90
T3/#T3	2002.12.31	YUMIfm top100 - A@ 13:55, B@ 15:15	Cassette Tape	90
T3/#T4	2002.12.31	YUMIfm top100 - A@ 16:00	Cassette Tape	90
T3/#T5	2003.01.06	KraiFM (NBC) 90.7fm A@ 17:20, B@ 18:30	Cassette Tape	90
T3/#T6	37629	YUMIfm - A@ 10:40, B@ 11:45	Cassette Tape	90
T4Md01	2003.03.10-15	?; PNGDF; quiet PR	MiniDisc	70
T4Md02	2003.03.17-18	L's fridge, rhythm, w. DW	MiniDisc	70
T4Md03	2003.03.18	UPNG dept. corridor meeting, + SK	MiniDisc	70
T4Md04	2003.03.19	Moresbi Misc	MiniDisc	70
T4Md05	2003.03.20	Moresbi Misc	MiniDisc	70
T4Md06	2003.03.30	Moresbi Misc	MiniDisc	70
T4Md07	2003.04.02	Moresbi Misc	MiniDisc	70
T4Md08	2003.04.03	Moresbi Misc	MiniDisc	70
T4Md09	2003.04.04	Moresbi Misc	MiniDisc	70
T4Md10	2003.04.06-08	>6 Q&M; >19 Proma	MiniDisc	70

APPENDIX II : 'Core Corpus' Files and Length

Recording Filename	CLAN transcript extension	Text Type	Length (mins)	Length (words)
TiT#07a-Hd01	HoP airports 3spk.cha	Parliament	6:36	794
TiT#07a-Hd02	HoP local government 3spk.cha	Parliament	2:50	451
TiT#07a-Hd03	HoP Death Penalty & Vagrancy Act 4spk.cha	Parliament	8:52	1171
TiT#07a-Hd04	HoP ports 2spk.cha	Parliament	3:04	381
TiT#08_208	MNmuddy.cha	Conversation	0:20	47
TiT#08_502	MNmamaluvana.cha	Conversation	0:10	21
TiT#08_506	MNslip.cha	Conversation	2:14	368
TiT#08_507	MNcsNotahighlander.cha	Conversation	4:01	653
TiT#08_508	MNwettumas.cha	Conversation	0:24	52
TiT#10b_03	YUMIfm globe ad.cha	Radio Advert	0:32	78
TiT#10b_05a	YUMIfm talkback.cha	Radio Talkback	1:36	329
TiT#10b_05b	YUMIfm talkback.cha	Radio Talkback	1:56	436
TiT#10b_07	YUMIfm talkback.cha	Radio Talkback	3:24	786
TiT#24b_04	smokbrus.cha	Conversation	1:16	199
TiT#24b_15	haus tambaran.cha	Conversation	11:49	1983
TiT#35b_02	HoP ombudsman.cha	Parliament	4:28	670
TiT#35b_04	HoP point of order.cha	Parliament	2:59	231
TiT#35b_05	HoP M*CS.cha	Parliament	1:09	141
TiT#35b_202	HoP M* idea.cha	Parliament	10:21	1658
TiT#35b_203	HoP N* army&police.cha	Parliament	9:49	1662
TiT#35b_204	HoP Police M*.cha	Parliament	8:55	1397
TiT#35b_205	HoP MP for W*.cha	Parliament	7:02	1080
	TOTAL:		93:47	14,588

APPENDIX III : CLAN Coding

The LIPPS Group recommended in their Special Edition of the International Journal of Bilingualism (LIPPS 2000) that code-switching researchers use the CLAN programs for encoding and analysis of multilingual data.

The Computerized Language Analysis (CLAN) program was designed to analyse data transcribed in the Child Language Data Exchange System (CHILDES). The LIPPS Group proposed modifications to the CHILDES transcription protocols, implementation of which should be recorded along with other metadata in supplementary Readme documentation accompanying texts in a bilingual dataset.

This appendix provides a reformatted version of the CLAN readme file template associated with the texts in the Tok Pisin-English 'core corpus' database. The readme file template is a complete record of the transcription and coding conventions discussed in Chapter 4. For my data, this template provides details of those conventions in a standardised and transferrable format.

Appendix IV provides a full text from the 'core corpus' coded following these conventions. Appendix VII, Sample 1 provides an additional example.

The following text duplicates the material in the readme file:

Readme.txt

Acknowledgements:

> The LIPPS Group ref. The International Journal of Bilingualism, vol4:2, June2000
>
> Brian MacWhinney & the CLAN project

Researcher:

> Tok Pisin-English database by Adam Blaxter Paliwala, University of Sydney

List of Files:

.CHA version:

Recordings: Tape#.wav

Date:

Transcription format:

Transcriber v.1.4.2

FILE: TiT#.trs

Number of hours transcribed:

0h:00m:00s

Characteristics of the community to which informants belong:

Informants:

Interaction type:

General Comments:

Data Review:

CURRENT:

2003-03-31 : this Readme file supports all current revisions of transcription and coding conventions.

2003-04-15 : this Readme file is the Master for transcription and coding practices. It is used to update or as supplement to the CHILDES files in the TP/Eng corpus database.

2006-02-15 : this Readme file is current and includes conventions for coding on all tiers, including Language Interaction %lic and Variation %var tiers critical to the two possible analyses proposed for highly variable data like this (my thesis presupposes that the bilingual model of mixed language provides for a more satisfactory account, and thus language type variation is principally coded in the main tier).

Special transcription and coding practices:

Changes to 00deplides.cha @ 00depadd.cha:

Strings / tiers added:

%ort: * #*

- this new tier has been added for the inclusion of standard orthographic transcription of the main tier

%glo: [unclear]

- this allows the standard entry '[unclear]' to fill for quiet or confused sections of the audio file.

%glo: *{*}

- the string '*{*}' is added to allow simple morphological comments to be added to glosses on this tier

%tra: [unclear]

- this allows the standard entry '[unclear]' to fill for quiet or confused section of the audio file.

%tr?: * #*

- this new tier indicates uncertain translations. It is an open entry field.

%lic: * #*

- this new tier indicates language interaction coding allowing for programs to search for identified code-mixing features (see below)

%var: * #*

- this new tier indicates language variation data coding allowing for programs to search for identified TokPisin (1) and English (2) variants of c.40 linguistic features (see below)

%sec: * #*

- this new tier is the Section header automatically transferred from the Transcriber file.

%tur: * #*

- this new tier is the Turn header automatically transferred from the Transcriber file.

%syn: * #*

- this new tier is the synchronous, or 'overlapping', header automatically transferred from the Transcriber file.

Transcription conventions:

Main Tier:

Main Tier transcriptions are produced by automatic transfer from the Transcriber v.1.4.2 software File/Export/Childes format. The .cha file produced has time index marking with the %syn tier, and turn marking with the %tur tier. These are automatic features of the transfer process and are inserted ABOVE the speaker tiers they delimit.

Main Tier transcriptions are therefore phonetic transcripts, reflecting the actual pronunciation of words. While this has the advantage of keeping the main tier representation close to the original audio and maintaining a significant level of phonological awareness significant to the present research, it may make the transcripts difficult to interpret for users lacking familiarity with the varieties involved.

Speaker Tiers are renamed using a basic find/replace operation. The automatic trasnfer process assigns speaker tier identities (*pk1,2,3,etc.) in sequence. These generic, text-specific identity markers are replaced with an informant-specific three character string.

The %ort: tier has been added to supplement the main tier text, giving a representation of the data using the orthographic standards of Mihalic's Dictonary and Grammar of Neo-Melenesian (1971) for Tok Pisin, and the Oxford English Dictionary for English. Contracted forms common to spoken registers of both languages - such as 'blo/ng' for Tok Pisin (+animate) preposition 'bilong' and -'s for English cliticised

copula 'is', for example - are retained on the %ort tier as they are in common use in written forms of spoken register and help retain some of the flavour of the original.

The following conventional representations are used in the main tier to represent phonemes as found in English (British) pronunciations of the sounds indicated by the letters capitalised in the orthographic standard written examples given. (IPA characters are impossible in this readme file, and so the 'phoneme(s)' column is largely included as a redundant feature.)

Vowels:		
on main tier	phoneme(s)	example
uh	ə	brothER, emphAsis
ei	eɪ	hAY, debAte, nAme
ai	aɪ	tIme, fIght
au	aʊ	hOUse, OUtside
A	æ	hAd, Acting
e	ɛ	hEAd, bEd
E	e	hEAd, bEd (Australian)
a	a	hEArt, pArt
I	ɪ	hId, fIsh
i	i	hEEd, wE
o	ɒ	hOd, bOd, dOg
u^	ʌ	bUd, hUt
u	u	yOU, trUth
oo	ʊ	gOOd, cOUld
aw	ɔ	lAW, cAUght

NB. An 'r' following a vowel most frequently indicates an r-colouring on the vowel- unusual in British or Australian English

NB. Other combinations of vowels (e.g. 'au'), where present, should be taken to indicate diphthongal movement.

NB. Capital 'I', where used as part of a sequence (aI), indicates the English 1st person singular pronoun as a dipthong.

NB. As can be deduced from the above listing, the use of character doubling (e.g. 'oo') DOES NOT indicate character length (indicated by [:]), but rather is used to denote a discrete vowel sound.

Consonants:		
representation	phoneme(s)	Example (standard English)
th	θ	THink, baTH
dh	ð	broTHer, THen
ks	ks	siX
sh	ʃ	SHirt, fiSH
ch	ʧ	CHain, churCH
dj	ʤ	juDGe, villaGe
ng	ŋ	raNG, runniNG
r	ɹ	*indicates an alveolar or approximant /r/
rr	r	*indicates a trilled or tapped /r/

Dummy Tier:

*xxx: a dummy (first) speaker tier is added before the default Transcriber %sec %tur tiers which disrupt parsing by appearing before speaker tiers. It is commonly attributed to the first speaker, and left blank.

CODING scheme: [cf. LIDES s4.2]

@1	Standard Tok Pisin pronunciation ref. Mihalic / Dutton & Thomas
@10	Abbreviated Tok Pisin : TP=0
@11	Variant pronunciation of Tok Pisin : TP=1
* loan pron? * @31	Shared / Loan / stock phrase with Tok Pisin pronunciation : Borrowed=TP1
@33	New Borrowing / coining : English with

	Tok Pisin morphology : BorrowedMix
@12	Tok Pisin with English phonology or semantics : 'Anglicised' Tok Pisin : TP=Eng2
@3	Identical in both languages
@30	Abbreviated shared item
@21	English with Tok Pisin phonology or semantics : 'Papua New Guinean English' : Eng=TP1
@33	New Borrowing / coining : Tok Pisin with English morpholgy [!?] : BorrowedMix
* loan pron? * @32	Shared / Loan / stock phrase with English pronunciation : Borrowed=Eng2
@22	Variant pronunciation of English
@20	Abbreviated English
@2	Standard English pronunication ref. AM Smith (1986)

nb. the regular pattern here is that 'standard' production receives the single numeral [x]/[y] language definition. A second digit is added to the language designator [x]/[y] to reflect pronunciation:

x0=abbreviations, xx=language-internal variant, xy=transferred phonology.

Shared items and borrowed items have the first digit 3, with 33 the designation for a morphological mix.

OTHER conventions:

- Overlapping speech is indicated throughout by two methods. The presence of two different speaker tiers below one %tur tier indicates both speakers talk within the interval defined in in the %tur. The # character is used in early .trs transcripts to indicate syllables poken by the other speaker while one speaker is silent. Here the LIPPS recommendation to indicate overlapping sections of speech with [>] to introduce the overlap and [<] to close is incorporated also. Sections of speech on

concurrent speaker tiers within these boxed arrows should be taken as overlapping sections.

- wavey brackets are used to enclose comments in the main (speaker) tier

%glo and %tra tiers give English language equivalents.

Neither are included where %ort reveals a standard (written) English translation.

%glo tier morphological abbreviations:

Pronoun gloss:

1,2,3	first, second, third person
sg/pl/dl/tl	singular/plural/dual/trial number
(ex)	exclusive pronoun (for first person)
(in)	inclusive pronoun (for first person)
e.g. 1pl(ex)	first person plural exclusive pronoun
3dl	third person dual pronoun

Other abbreviations:

NEG	negative/negator
-	(in %glo) affix indicator
TR	transitive (suffix)
PROG	progressive
COP	copular
Qtag	question tag or particle
PREP	preposition
LOC	locative (preposition)
{-/+a}	inanimate/animate object
DET	determiner
PM	predicate marker i

%lic is the language interaction tier, where interpretations of interaction patterns in the data are included.

In general, the format of this tier follows that of the LIPPS recommendations in the LIDES coding manual, though certain alternations to the codes involved and the syntax of the coding scheme have been made.

%lic Codes:

Category codes

c	code switch
b	borrowing
y	syntactic calque
e	semantic calque
f	function
d|x	direction: T for Tok Pisin; E for English (lower case marking may be used)
|or|	alternative interpretation follows

eg. format: c|xx:d|x:f|xx |or| y|xx:x,x

Word Type codes

N	noun
V	verb
Adj	adjective
Adv	adverb
Prep	preposition
Conj	conjunction
Comp	complementiser
Pron	pronoun
Q	question word
Tag	tag
Mod	Modal
PM	predicate marker
Cop	Copula
Num	numeral
Temp	temporal marker
Det	determiner
Loc	locative (preposition
Pl	plural marker
Aux	auxilliary
Neg	negative particle

Grammatical role codes

Pred	predicate
Subj	subject
Obj	object
Rc	relative clause
Sc	subordinate clause
Mc	main clause

Specifics within Categories:

c | code mix, specifics

turn

te	turn external
ti	turn internal

interS

se =sentence external

t	tag
qt	quote
cs	coordinate sentence #nb. Sentence external switching is more directly functional
ss	subordinate sentence

Eg format: c|ti:se:f|xx

f | function, specifics

rept	Repetition
intr	Interrogation / Interrogative
addr	Addressee / Topic
qual	Qualification / Qualitative
demo	Demonstrative
expl	Explicative
dirc	Directive
decl	Declarative
quot	Quotation
idio	Idiom
disc	Discourse pattern / markers / identity

focu	Focus

intraS

si =sentence internal

1	constituency	
a	nonconstituent	
	S	single word
	M	multiple word
	G	grammatical

Eg. format:

1aS/M/R/G:(S/M:N/V/Adj/Adv/Prep/Conj/Comp/Pron/Q/Tag/Mod/PM/Cop;

G:Pred/Subj/Obj/Rc/Mc)

R =ragged

b	constituent	
	CP	complementiser phrase
	VP	verb phrase
	NP	noun phrase
	AdjP	adjectival phrase
	AdvP	adverbial phrase
	PP	prepositional phrase

Eg. format: 1b:CP/NP/VP/AdjP/AdvP/PP/Pron/Q/Tag

mode

2	mix mode	
Ins	insertion	
	F	function word (MLF model)
	C	content word (MLF model)
	I	island (MLF model)
	M	mixed constituent (MLf model)
Alt	alternation	
CLx	congruent lexicalisation	

fluency

3	fluency	
	f	fluent
	n	not fluent (pauses

		and interruptions)
	c	self-correction features

direction

d|	direction of switch	
	P	to Tok Pisin
	E	to English

Eg. format: %lic: c|tx:sx:1x(xx.xx)-2x:xx-3x:d|P

borrowing

b |	borrowing, specifics	
type		
	core	core
	cult	cultural
	est	established
	non	nonce
	coin	mixed word /coin
	shar	shared word
integration		
	pi	phonological integration
	mi	morphological integration
	sy	syntactic integration
	f	full integration
	p	partial integration
	n	no integration
	d	double marking

Eg. format: %lic: b|core/cult/non:pi~f/p/n-mi~f/p/n-si~f/p/n

%lic syntax:

The format of these codes is quite variable in the corpus. The syntax of the coding scheme essentially isolates individual phenomena in each tier by assigning a new string to each observation. Within the string, specific characteristics are separated by the : (colon) character, but a loose formal hierarcy is created through the use of other characters (see General notes below).

%lic special characters:

-	(Category code)|(specific1):(specific2a,specific2b):(specific3)-(Specific3observation):(xxx):(Category code)|(specific1): etc.
|	(vertical bar) character indicating that what follows relates to a certain category of interaction,
:	(colon) character speparates specific observations
~	(tilda) character links multiple observations in the same area
-	(hyphen horizonal bar) character links related observations on specifics in sequence
|or|	indicates the strings which follow represent an alternative interpretation from a different theoretical perspective
.	(period) character delimits strings relating to single language interaction phenomena.

Due to constraints in the formatting of the tier structure, it is left up to the reader to correctly recognise the parts of speech each string identifies.

To assist in identification, the sequence in which specific language interaction phenomena appear on the main tier is rigorously duplicated in the order in which the strings are presented here.

General:

If more than one type of language interaction occurs on a single tier, then additional discrete strings are included, separated by the period mark.

Alternative interpretations of particular interactions are indicated by the |or| symbol between alternatives.

Where |or| is between strings delimited by period marks, it indicates that what follows is an alternative to all prior observations on that tier.

A %lic: string such as:

c|ti:se:1aG:Subj.Pron-2Ins-3f:pi:d|e

can be interpreted as indicating the presence of a phonologically integrated (pi), English (d|e), Pronominal Subject (Subj.Pron), as a Grammatical nonconstituent (1aG) insertion outside of a sentence.

Of course, this is not a logical string, as insertion of subjects is a sentence internal phenomenon, but the syntax is illustrated.

Additional comments (asides), can be included where absolutely necessary in {xx} wavey brackets.

%var is the variation coding tier, where data is recorded on the forms of identified TokPisin/English variables identified in the corpus.

In general this tier is formatted so that the order of variables presented reflects the order in which they appear in the main tier coded transcript. This data allows inter-language variation in the corpus as a whole, and for sections of the corpus as selected for analysis, to be evaluated in terms of both the presence/absence of certain variant forms in the transcribed data, and additionally with some reference to the language context in which variants occur. In this way the coding on the %var tier post-theoretically incorporates data which reflects the hypothesis that the TiT TPEng corpus recordings are not of a single language with a single axis of variation: "+/- standard features", but of two languages, thereby necessitating the inclusion of codes for data on a sencondary axis of variation: "langauge of utterance".

So a varible such as 01 (dh) has a total of four possible values, reflecting both its form and its context:

(dh):		
	1p	"phoneme [d] in TokPisin context";
	1e	"phoneme [d] in English context";
	2p	"phoneme [ð] in Tokpisin context";

	2e	"phoneme [ð] in English context".

Language Context Type:

p	Tokpisin
e	English

Special Characters:

-!-	Instance of special note
-x	Non-standard application of variant : suggests incomplete acquisition of variable rule)
v-	Similar-to Variable : Variable follows the same rule as the cited example [may be useful] (esp. Phonological Variables 03. and 04. short/long and high/low vowels)

%var codes:

Variable Category / Name	Variable-values	Variables
Phonological		
(dh)	1=[d]; 2=[ð]	01. d vs d~dh - dis vs dis~this
(th)	1=[t]; 2=[ɵ]	02. t vs t~th - teet vs teet~teeth
(I)	1=[ɪ]; 2=[i:]	03. i vs ee - rɪfreshment vs ri:freshment (Verhaar 1995: 9 - no short/long vowel)
(a)	1=[e]; 2=[æ]	04. e vs a - ken vs kan (Verhaar 1995: 9 - no high/low vowel distinction)
(uh)	1=[ɪ/u/ɑ/æ]; 2=[ə]	05. (non-central vowel sounds) vs ə
(dj)	1=[s]; 2=[dʒ]	06. s vs s~j~dj - sas vs charj, sas vs djudj (#check corpus for "j"→ "y")
(ch)	1=[s]; 2=[tʃ]	07. s vs s~ch -
(sh)	1=[s]; 2=[ʃ]	08. s vs s~sh -
(z)	1=[s]; 2=[z]	09. s vs s~z - is vs is~iz [stress?]
(pf)	1=[p/f]; 2=[p] [f]	10. p/f vs p or f
(ng)	1=[-n]; 2=[-ng]	27. -ng vs -n (# revise data for partial merge with 40.)
(-t/d)	1=ø; 2=[T/D]	28. -T/D terminal (# revise data for partial merge with 40.)
(kg)	1=[k/g]; 2= [k] [g]	32. k/g vs k or g
(aw)	1=[o]; 2=[ɔ]	35. ɒ/o vs ɔ - kot vs kawt
(o)	1 =[æ] ; 2=[o]	36. æ vs ɒ/o - gat vs got

(CC)	1= Laycock's permitted clusters; 2= others	37. Consonant Clusters (see table)
(DIP)	1= Limited set (Laycock + Smith); 2= others	40. Dipthongs (see table)
Syntactic		
Poss	1=N + bilong + Agent; 2=Agent + -'s + N	11. possession word order - N + bilong + Agent vs Agent +'S + N
Pass	?	23. Passives
Rela	?	24. relative constructions
Pred	1= *i* as predicate marker; 2=ø	15. consistency of predicate marking with i
Prep	1=long; 2=to/on/by etc.	26. prepositions
DO	1 = ø; 2= DO+V	43. dummy auxiliary DO
Cop	1=ø; 2=is/be etc.	29. copula
TMA		
ConA	1a= wok long + V; 1b= V + i stap; 2= be + V-ing	12. continuous aspect - wok long V, V i stap vs be + V-ing
PasT	1= invariant forms: bin + V; 2= variant forms / V-ed	13. past tense verb - invariant forms, bin + V vs Eng variants (nb. Ref. also 28.)
ComA	1= V + pinis; 2 = V-ed	14. completive - V + pinis vs V-ed
Cond	1=sapos; 2=if	30. conditional - if vs sapos
FutT	1= bai +V; 2=will/may etc.	33. future - bai + V vs Eng variants
HabA	1= save; 2= usually/often etc.	38. habitual aspect - save + V vs English variants
Prox	1= laik; 2= ø	39. proximative aspect - laik + V vs none
Modal	1= ø; 2= can, may, shall, will (modal uses)	41. modal auxiliaries - ø vs English can, may, shall, will
Time	1= taim etc.; 2= when etc.	42. temporal reference - taim vs when
Morphology		
Nega	1= no, nogat; 2= DO + not	16. negation - no, nogat, vs DO + not
ArtD	1= ø, wanpela; 2= the	17a. Articles: definite article - ø, wanpela vs the;
ArtI	1= ø; 2= a	17b. indefinite article - ø vs a
Pron	1= mi/em/en; 2=	18. Pronouns - mi vs I; em vs them; en vs it

	I/them/it	
Conj	1= na; 2= and	19. conjunctions - na vs and; but
Plur	1= ol + N; 2= N-s/z	20. plural marking - ol + N vs N-s
Numb	1= -pela; 2= bare	21. numbers - wan vs wanpela etc.
Trans	1 = -im; 2= ø	31. transitive verbs - -im vs no marking
Adj	1= -pela; 2= bare	34. 'adjectives'
Lexical		
LexV	[various]	22. lexical variation - diwai vs tree; pikinini vs child; people vs manmeri; papa vs father; mama vs mother. [# REM: insertional CS might account for this...] nb. Also using this category for English loan words generally being used instead of circumloqutions...
Disc	1=a; 2=eh etc. ?	25. Discourse
Loc	1=TP adverbials; 2=English Adverbials	44. Adverbial Location

CONSONANT CLUSTERS (NOT PERMITTING EPENTHETIC VOWELS)

from Laycock, Don, 'Phonology: Substratum elements in Tok Pisin phonology' (1985: 296) (in Wurm & Mühlhäusler The Handbook of Tok Pisin (1985))

English	Tok Pisin	English Example	Tok Pisin equivalent
mp	mp	company	kampani
	m	lamp	lam
mb	mb	number	namba
nt	nt	mountain	maunten
	n	cent	sen
nd	nd	window	windua
	n	hand	han
nč	nis	winch	winis
	ns	?	
nǰ	nis	change	senis
	ns	engine	ensin
ŋk	ŋk	monkey	manki
		bank	beng
ŋg	ŋg	finger	pinga

nb. phonological variable value '2' codes which allow for the presence of EITHER value eg. (dh):2 = dis~this are represented at the level of %var coding as indicating the presence of an ABSOLUTE value. Subsequent analysis of eg. the total variation of realisations of this feature by any one speaker must take place before the general "2 possible values" evaluation represented by eg. d~dh, can be conclusively assigned.

CONSONANTS:

nb. Verhaar 1995 [8]

"Tok Pisin has no word-final voiced consonants"

CONSONANTS	
-d	→ -t
-ds	→ -s
-dj	→ -s
-g	→ -k
-v	→ -p

"Tok Pisin has no word-final consonant clusters"

=> eg. "happened" = [hæpənd] → [hæpən]

"Acrolectal Tok Pisin has some exceptions to these laws"

eg. difens

Hailans

generally this rule mandates-T/D deletion and removal of terminal -ng

= variables 27. and 28.

VOWELS AND DIPTHONGS

Verhaar 1995: 6-7

simple vowels

a	[æ]
e	[e]
i	[ɪ]
o	[o]
u	[ʊ]

dipthongs
ai
au
oi
Smith 2002: 44
ia
iu

#evidence of Thomason & Kaufmas's (1988:74) "slight structural borrowing"

%var syntax

The format for coding this tier is to present each occurance of a variable form along with the codes identifying its value as a single string. Strings are seperate by spaces. Strings are generally presented in the order in which variables are to be found on the main tier.

The structure of a code string is typically:

[Variable]:[Value][Language Context Code]

A %var string such as Numb:1p can be interpreted as indicating the presence of a number (Variable 'Numb') with the, traditional Tokpisin, -pela suffix (Value '1') included, occurring within the context of language could be identified as generally Tokpisin (Language Context Code 'p').

The %var tier may include any number of strings, each relating to a separate occurrence of a given variable.

However, please note that not all repeat occurrences of a particular variable may be included at this time.

Analysis of the %var tier

Significant for analysis of the coding data may be the occurrence of contrasting values for a single variable the occurrence of contrasting variable values in a single language context [this reflects the general tendency for Variable Value codes to present typically Tokpisin variants as '1' codes, and typically English variants as '2' codes, so that '1' values may be expected to occur with the 'p' tag = in a Tokpisin context, and '2' values with an 'e' tag => in an English context.) any patterns in the occurrance of variables in different language contexts this may be particularly significant for variables where there is a null value for one language eg. Trans]

REM: ALL CODING IS ACCUMULATIVE AND NOT EXCLUSIVE.

THE ABSENCE OF CODING FOR A PARTICULAR FEATURE DOES NOT GUARANTEE IT IS NOT PRESENT.

YOUR CONTRIBUTIONS ARE ALWAYS APPRECIATED.

APPENDIX IV : Example 'Core Corpus' Text – CLAN format

The following is a direct copy of the CLAN format file for a short conversational text from the core corpus.

It includes the full file details, including the file headers identifying the participant codes (@Participants) and the time indexes (%sec, %tur, %syn for the whole file and the particular lines of trasncription).

This appendix illustrates the transcription and coding conventions in use. In addition to the main transcription tier (*[Participant code])with the phonetic transcription and Turbo Language Tags, the orthographic transcription tier (%ort), the variation coding tier (%var), and the language interaction coding tier (%lic) are included. Another example including the CLAN coding tiers is presented in Appendix VII, Sample 1.

Filename: TiT#08_208 – Mnmuddy.cha

```
@Font:          Win95:Courier:-13:0
@Begin
@Participants:MNM MN Mother, YNT youngN Teenager, KLT
        KL Teenager, PLM PL Mother
@Languages:  Tok Pisin (1), English (2), Shared (3), Unknown (4), (see
        coding scheme in readme document)
*MNM:           .
%sec:  type=report startTime=0.000 endTime=19.925
%tur:  mode= startTime=0.000 endTime=4.277
*MNM:           &unclear .
%syn:  1.050
*MNM:           &eh yu:pla@10 go@1 mekim@1 wanem@1 &ay &ay # !
%ort:  eh yupela go mekim wanem hey hey !
%glo:  eh 2pl go make-TR what hey hey !
%tra:  eh, where are you(pl) going hey hey !
%var:  Pron:1p Trans:1p .
%syn:  2.740
*MNM:           no@2 pleiin@21 a@1 its@20 maddy@21 .
%ort:  no playing a it's muddy .
%glo:  NEG play-PROG Qtag 2sg{-a}-COP muddy .
%tra:  no playing okay (question) it is muddy .
%lic:  c|ti:si:1aG:Q-2Ins-3a:d|P .
%var:  (-ng):1e Nega:2e Disc:1e Pron:2e Cop:2e (uh):1e .
%tur:  mode= startTime=4.277 endTime=5.481
*YNT: &hey &ay ## .
%ort:  hey hey .
%tra:  hey hey .
%tur:  mode= startTime=5.481 endTime=6.044
*KLT: ol@1 lain@1 lo@10 +...
```

%ort: ol lain lo ...
%glo: 3pl group PREP{-a} ...
%tra: those people (trails off) ...
%lic: c|te:se:f|expl:d|P .
%var: Plur:1p LexV:1p Prep:1p .
%tur: mode= startTime=6.044 endTime=15.315
*MNM: yu:pela@1 lai@10 go@1 lo@10 we@12 ### ?
%ort: yupela laik go lo we ?
%glo: 2pl like go PREP{-a} where ?
%tra: where do you want to go ?
%lic: c|te:se:2Alt:f|addr .
e|ti:si:1aS:Q-2Alt:d|E .
%var: Pron:1p ?Prox:1p Prep:1p .
%syn: 7.786
*MNM: no@2 ken@1 go@1 pilai@11 plis@1 .
%ort: no ken go pilai plis .
%var: Nega:1p (a):1p LexV:1p (I):1p
%tur: mode= startTime=8.978 endTime=13.671
*MNM: its@20 maddy@21 and@2 drrain@21 an@20 ## wi:@2 don@20 wan@21 duhti@2 fit@21 runing@2 insaid@2 .
%ort: it's muddy and (the) rain and ## we don't want dirty feet running inside .
%lic: c|ti:se:cs:f|expl:d|E .
%var: Pron:2e Cop:2e (uh):1e Conj:2e (-t/d):2e ConA:1e Conj:2e (-t/d):1e Pron:2e Nega:2e (-t/d):1e (uh):2e (I):1e ConA:2e (ng):1e .
*PLM: # ## # ## # # xxx .
%tur: mode= startTime=13.671 endTime=15.315
*MNM: yu:pla@10 planti@10 lis@4 reko@4 lainim@1 ia@1 ?
%ort: yupela plantim this? reco? lainim hia ?
%glo: 2pl plant-TR? DET? reco? lined-TR here ?
%tra: did you leave this (something) lying here ?
%var: Pron:1p Trans:1p ArtD:1p Trans:1p .
%tur: mode= startTime=15.315 endTime=16.186
*KLT: yea@22 .
%ort: yes .
%tur: mode= startTime=16.186 endTime=19.925
*MNM: nodher@4 spaisis@4 wo@2 aI@2 teik@2 yu:@2 ##### .
%ort: another? spices? wo? I take you .
%glo: another? spices? ? 1sg take 2sg .
%tra: I will take you (something) .
%var: (dh):2e Pron:2e .
@End

APPENDIX V : Software Used in this research

Microsoft Office : http://office.microsoft.com/en-us/

Trasncriber : http://transag.sourceforge.net/ (originally : http://trans.sourceforge.net/en/presentation.php)

Soundforge : http://www.sonycreativesoftware.com/soundforgesoftware

Audacity : http://audacity.sourceforge.net/

CD Wave : http://www.milosoftware.com/cdwave/

Praat : http://www.fon.hum.uva.nl/praat/

CLAN : http://childes.psy.cmu.edu/

APPENDIX VI : Listening Exercise Introductory text

Text of pre-recoded introduction to the 'Skelim Mix lo tok Inglis na Tok Pisin' Listening Exercise cassette:

"Hello. Adam hia. Thank you for agreeing to participate in this language research. This tape contains the listening material for one exercise: 'skelim mix lo tok Inglis na tok Pisin'. You will need a pen and a piece of paper to write your responses. ... Following this introduction there are a number of samples of speech recorded here in PNG. There are 26 samples in total, and each plays for about 20 second. Each sample is introduced by number before it is played. You should write the numbers from 1 to 26 down the left side of your paper, and your age and place of origin at the top. ... For this exercise we listen to each sample and give it a 'score' from 1 to 10. We use the number 1 to indicate tru tru English, and 10 for tru tru Tok Pisin. The other numbers from 2 to 9 are used to record a 'mix' between English and Tok Pisin. The lower the number, the more we think the sample is English. For example: 5 and 6 can be used to indicate a half half mix. Please write the score you give for each sample next to it's number on your paper. ... You can listen to each sample as many times as you want. To help you make up your mind about the socre, you may like to think about the parts of speech involved in the mix. ... You can write any comments you like on your paper. You maybe able to guess the gender, age, place of origin, and probable occupation of the speaker. You can also include these observations next to each sample number. ... With this exercise we are not testing you. There are no right or wrong answers. We are interested in your opinions and intuitions, so please think for yourself before making an answer. ... Thank you again for joining in this part of the Tok Pisin / English survey. The Melanesian And Pacific Studies (MAPS) centre is hosting a seminar on some Linguisitic aspects of this research at 4pm Thursday 1st May, at the Arts Lecture Theatre, UPNG campus. You are welcome to join us there. ... Now: Listening Excercise begins:"

APPENDIX VII : Listening Exercise Samples

The samples are reproduced here as coded in CLAN as for the Core Corpus.

As well as providing details of the sample texts used in the Listening Exercise described in Chapter 9, the first text also provides an additional example of the CLAN coding conventions described in Appendix III, illustrated in Appendix IV, and used throughout the core corpus. Here, in order to make the main tier easier to read, the Turbo Language Tags are ommitted.

Sample 1 includes the full file details, including the file headers identifying the participant codes (@Participants) and the time indexes (%sec, %tur, %syn for the whole file and the particular lines of trasncription). In addition to the main transcription tier (*[Participant code]) and the orthographic transcription (%ort), the variation coding tier (%var) and the language interaction coding tier (%lic) are included.

For the other samples here only the main tier and orthographic transcriptions have been reproduced, and formatting applied to aid the reader.

In all cases the main tier has been reformatted to follow the standards in the dissertation text: in Lucidia San Unicode font, with English in ***bold italics***.

Sample 1 : Clan format example : single Alternation

@Begin
@Participants: RRR Madang_R Informant, AAA ABP Researcher
%sec: type=report startTime=0.000 endTime=17.554
%tur: mode= startTime=0.000 endTime=6.494
*RRR: yu yu yu olsem yu stap ### long disela # y yu stap insaid steret lo disela kain .
%ort : yu yu yu olsem yu stap , long dispela , y- yu stap insaid stret long dispela kain ...
%var : ConA:1p Prep:1p ArtD:1p LexV:1p Prep:1p ArtD:1p ConA:1p Prep:1p
%tur: mode= startTime=6.494 endTime=7.088
*AAA: yes .
%tur: mode= startTime=7.088 endTime=7.741
*RRR: *stadi* a ?
%ort : study a ?
%var : LexV:2p Disc:1p
%lic : c|ti:si:1aS:N-2InsC-3n:d|e:b|cult:pi~f:sy~f "dispela kain stadi a?"
%tur: mode= startTime=7.741 endTime=8.460
*AAA: yes .
%tur: mode= startTime=8.460 endTime=17.554
*RRR: au yu thing ols anem # *ts gud tu go bek an len ## tokpisin in klasrum o tokpisin s oredi dhe # ts mater of yusing it ?*
%ort : how you think olsem wanem: it's good to go back and learn tokpisin in classroom or tokpisin is already there, it's matter of using it ?
%var : LexV:2pe DAux:1e (th):2pe Cop:2pe Adj:2e Prep:2e (a):1e (CC):1e Dip:1e Prep:2e ArtD:1e Cop:2e (aw):1e (dh):2e Cop:2e Artl:1e Prep:2e (ng):2e Pron:2e
%lic : c|ti:si:1bAdjP-2Insl:Q-3f:d|p:pi~n:sy~n:f|intr "au yu thing ols anem"
@End

Sample 2 (11) : Insertion : single words, Adj+N, V - integrated

Participants: NIU Niusman_Announcer

Length : 20.062 seconds

*NIU: ol toktok blon ***otonomi*** blon Bowganvill em ol bigpla toktok we bai kamap lon namba foa ***politikol rroun*** # blon toktok namel lon ***nasnol gavman*** # no lidaman blon Bowganvill .

%ort : ol toktok bilong autonomy bilong Bouganvile em ol bigpela toktok we bai kamap long namba foa political round bilong toktok namel long national government no leaderman bilong Bouganville .

*NIU: gavman bai stap staat wawk lon ***tenderim*** o putim aut Finans Pasuhfic lon ol lain # husat i laik baim lon wik i kam aninit lon displa ***praivitaizashin*** # ***eksuhsaiz*** blon ol .

%ort : government bai stap stat wok long tender-im o putim aut Finance Pacific long ol lain husat i laik baim long wik i kam aninit long dispela privatisation excercise bilong ol .

Sample 3 : Insertions : single words and Adj+N

Participants: MPM Member of Parliament M

Length : 19.505 seconds

*MPM: ol ***sevis*** nau mi luat long en .

%ort : ol service nau mi lukaut long en .

*MPM: ***helt edyukeishin*** ## ***rowd infuhsruksiuh*** ## disela olget i ***rundaun*** .

%ort : health, education , road infrastructure, dispela olgeta i run down .

*MPM: ##### disl ol ***suhvis*** nau mipla ks lng en em blong bifo trru ### bifo***indepen*** bla ksm en dluh ***suhvis*** em stap .

%ort : dispela ol service nau mipela kisim long en em bilong bifo tru, bifo independence mipela kisim en; dispela service em stap .

*MPM: #### .

*MPM: nau disela o ***sevis*** em i rrundaun .

%ort : nau dispela ol service em i rundaun .

*MPM: ### .

*MPM: e mi mi gat ***konsuhn*** tru lo displa .

%ort : and mi mi gat concern tru long dispela .

Sample 4 : Insertion : N, Adj+N, Discourse markers, numbers

Participants: NPF NPF_meri, UET Uncle_ET_Presenter

Length: 22.891 seconds

*NPF: usat asol wok lo rausim ***asistid pesen*** blo eNPieF nau ol w& sapos ol ***remembuh*** lo eNPieF suhpos ol laik # osem ## rause nem bl ol lo eNPieF em b osem wanem n(a)o ?

%ort : husat tasol wok long rausim assisted person bilong NPF nau ol w- sapos ol remember long NPF sapos ol laik, olsem, rausim neim bilong ol long NPF em bai olsem wanem nau ?

*UET: ah& mi no klia nau amin ahm& mi tu mi les lon ***kontribyut*** nau .

%ort : ah, mi no klia nau I mean ahm, mi tu mi les long contribute nau .

*NPF: tru yeh .

%ort : tru ia .

*UET: mm& .

*NPF: ose na miai na .

%ort : olsem nau mi ai? na .

*UET: okei .

%ort : okay .

*NPF: putim lili ***koments*** .

%ort : putim liklik comments .

*UET: ***no wuhris*** .

%ort : no worries .

*NPF: thenk yu .

*UET: thenk yu lon ***kol*** wuhrai nau .

%ort : thank you long call goodbye nau .

*NPF: # # # # ## bai .

%ort : bye .

*UET: ***thri tu zero wuhn nain faiv faiv o thri tu zero wuhn nain faiv thri*** suhpos yu laik wokim ***komen*** tetei lo moning ***telefon lains*** .

%ort : three two zero one nine five five o three two zero one nine five three sapos yu laik wokim comment tedei long moning telephone lines .

Sample 5 : Alternation : back and forth

Participants: RRR Madang_R, AAA researcher, TTT Madang_T

Length : 39.288 seconds

*RRR: nogat .

%ort : nogat .

*AAA: yes .

%ort : yes .

*RRR: *al dont al don think papa* a ?

%ort : I don't I don't think papa a ?

*TTT: uhhuh& .

*AAA: mm&

*RRR: *its a sistuhm .*

%ort : it is a system .

*AAA: *sistem .*

%ort : system .

*RRR: *edukeisen sistuhm ### enfows (u)as perents # tu tits au bed* ah& *# kids ## tu adopt tu ## a sistuhm .*

%ort : education sysetem, enforces us parents, to teach our bad ah, kids, to adapt to, a system .

*AAA: sistem yeah .

%ort : system yeah .

*RRR: *yust in skul* a ?

%ort : used in school a ?

*RRR: *al thing all put it dhat wei .*

%ort : I think I'll put it that way .

*RRR: mi mi sa thin *oh wai wai* i pu #### na *edyukesuhn* tu yu tokples long # skul yu *panis* .

%ort : mi mi save think oh why why i putim ... na education too yu tokples long, skul yu punish .

*TTT: hhhh .

*AAA: ah hah .

*RRR: *so yu not alaud tu spik yoa langwids in skul .*

%ort : so you not allowed to speak your language in school .

*AAA: # # # ## # # # tu tokples yeah .

%ort : to tokples yes .

*RRR: *so # its a sistem dhet faws as tu #### thro wan thin aut of tha windo an pu thi adhe wans .*

%ort : so, it is a system that force us to, throw one thing out of the window and put the other ones .

Sample 6 : Alternation : Back and forth; Insertion : Conj, Discourse

Participants: MSM M_Security_Madang, AAA Researcher

Length: 37.895 seconds

*MSM: *yu slip dea an deyl kuk fud fo yu* .

%ort : you sleep there and they will cook food for you .

*MSM: #### .

*MSM: *an* yu kaikai kaikai yu stap tasol ### *an den # dey a pipuhl hu wil gow an ah ## luk fo yo ## fo yo potin* ## .

%ort : and yu kaikai kaikai yu stap tasol, and then, they are people who will go and ah, look for your, for your protein .

*AAA: mmm& ## # # # ### ## em nau .

%ort : em nau .

*MSM: *an potin is ol wayl animuhls ## dhem kil flain foks dheyl kil pig waly pigs* uhm& uh& *bandicoot an watnot* a ?

%ort : and protein is ol wild animals, them kill flying fox they'll kill pig wild pigs uhm uh bandicoot and whatnot a ?

*MSM: *wayl buhds* ##### .

%ort : wild birds .

*MSM: bai ol i kisim kam ol i kukim lo yu .

%ort : bai ol i kisim kam ol i kukim lo yu .

*MSM: *an* ah& ##### nau bai ol i kukim nuh ## bai i ### kukim saksak tuhnim saksak .

%ort : and ah, nau bai ol i kukim nau, bai i, kukim saksak tanim saksak .

*AAA: mm& mmm& .

*MSM: na sup # *deyl mek sup yuno* uh& *get* ah& ### *taro an yam an en dhey mes it* a ?

%ort : na soup, they'll make soup you know uh get ah, taro and yam and then they mash it a ?

Sample 7 : Insertion : Islands, Pronouns, N - unintegrated

Participants: UET UncleET_Announcer

Length : 19.040 seconds

*UET: *wel* suhpos yu bin nau traim ol lo yu we te le moinen no gat thru sori tumas em se i e lain i said bisi bisi stret so ol yu ken uh # yu ken *trai mi gen tumoro moning okei* ah i no *tomoro moning* long mande mon yumi ken traim wokim *lains* blo yumi gen long uh # kisim *mo kols* .

%ort : well sapos yu bin nau traim ol long yu we tede monin no gat through sori tumas em se e? line i sait bisi bisi stret so ol yu ken uh, yu ken try mi gen tomorrow morning okay ah i no tomorrow morning long Mande morn(ing) yumi ken traim wokim lines bilong yumi gen long uh, kisim moa calls .

*UET: em tuhsol nau taim bifo ol maks *wi* got lon taim blo yu lon ste monin theng yu long ah& ### long awl *kawls* stah& thang yu long ah& o *kuhmpuhni* blo yu .

%ort : em tasol nau taim bifo ol remarks we got long taim bilong yu long tede morning thank you long ah, long all calls stah- thank you long ah ol company bilong yu .

Sample 8 : Insertion : multiple : Adj+N, two word V, Islands, N - integrated

Participants: MPT Member of Parliament T

Length : 27.307 seconds

*MPT: na nau mi mi bin r ra wanem rritim wanpla pas i kam long ## *ofis* blo yu minista ### long wail uh& ol wok long uh& *terim ap* isla *trenin ## program* long ol ## mi gat *olmos twelf tausin # pauns of meteriuhls* blong *distrik* blo mi ##### i stap long xxxx *epot* ##### long *eliftim* #### i go antap ### na *ap intil nau* mi no ksi wanpla *ansa* blo yu .

%ort : na nau mi bin r- ra- wanem writim wanpela pas i kam long, ofis bilong yu minister, long while uh& ol wok long uh& tearim up dispela training, program long ol, mi gat almost twelve thousand, pounds of materials bilong district bilong mi, i stap long xxxx epot, long airliftim, i go antap, na up until now mi no kisim wanpela ansa bilong yu .

Sample 9 : Insertion : Adj+N, N, Islands, V - integrated

Participants: SSS caller_S, UET ET_Presenter

Length: 19.180 seconds

*SSS: *las wik* mi bin arim long ## ah& ol i gavman ol i putim isela kain lo blong olim ol *prais* lo ol samting ia .

%ort : last week mi bin harim long, ah& ol i gavman ol i putim dispela kain lo bilong holim ol prais long ol samting hia .

*UET: mmm& ### .

*SSS: &ah beain aselum o samthin .

%ort : &ah behain asylum o samting .

*UET: yes ### .

*SSS: nau lo tinting blo mi gavaman ogata taim a thin sa se yu gavman ### ol i got nau wantaim ol *veri siotem plans* a lo *teigim* samuhla kain ah& .

%ort : nau long tingting bilong mi government olgeta taim ating save se yu gavman, ol i gat nau wantaim ol very shorterm plans a? long takeim sampela kain ah& ...

UET: sotkat .

%ort : shortcut .

*SSS: *priemtiv mesuhs* lo ol *problems* yumi gat .

%ort : preemptive measures long ol problems yumi gat .

Sample 10 : CLx? : Alternation & Insertion : Prep, Det+N, Discourse

Participants: TTT Madang_T, ABP Researcher

Length: 18.390 seconds

*TTT: *laik him # is* em malgebuk a ### ?

%ort : like him, is em Malgebuk(?) a ?

*TTT: mi malgam lo *in* ia tasol lo pi nulpla #### marid blon tu lus lo tok puhles .

%ort : mi Malgam(?) long in hia tasol long pi nulpela, marit bilong (x) tu lus long tok ples .

*TTT: *togs uhn* tokpisin tasol .

%ort : talks in tokpisin tasol .

*ABP: he heh he ##### .

*TTT: *thuh problem* em sap lo mitluh uhm paba ##### *bigos dets uh deni langwids wi speek a* ?

%ort : the problem em stap long mi tupela em papa, because that is the very language we speak a ?

Sample 11 (2) : Insertion : single words, Adj+N, V - integrated

Participants: NIU nius_man_Announcer

Length: 20.062 seconds

*NIU: ol toktok blon *otonomi* blon bowganvill em i bigpela toktok we bai kamap lo namba fo *politikol rroun* # blon toktok namel lon *nasnol gavman* no lidaman blon bowganvill ## .

%ort : ol toktok bilong autonomy bilong Bouganvile em i bigpela toktok we bai kamap long nambe fo political round, bilong toktok namel long national government no lidaman bilong Bouganville .

*NIU: gavman bai staat wawk lon *tenderim* o putim aut Finans Puhsifik lon ol lain # husait i laik baim lon wik i kam aninit lon displa *praivitaiseishin # eksuhsais* blon ol .

%ort : government bai start work long tender-im o putim aut finance Pacific long ol lain, husat i laik baim long wik i kam aninit long dispela privatisation, exercise bilong ol .

Sample 12 : CLx? : Insertions & Alternations : multiple : Adj+N, V, N, Conj, Pron, Adv, Islands

Participants: MpT Minister_for_Transport

Length: 25.449 seconds

*MpT: ah& # .

*MpT: lon ***sivu eiveishin*** # mipela i ***risponsibuhl*** lon ***epawt*** ### ah& em toktok long ***epot*** i bagarap ***an thes # konsuhn*** blong ***sivu eivieishuhn*** ## na ***sefti espe*** blong ***opereishin*** suhpos em ## ***difens eroplein*** ## ***den*** # mi mas go bek lon ahh& # ***thru difens minista*** # na i go bek long ah& ***othoruhti*** blon disela ***eroplein*** em i ***uhstreilian hai komishuhn*** ### a ?

%ort : long civil aviation, mipela i responsible long airport. ah& em toktok long airport i bagarap and there is concern bilong civil aviation, na safety aspect bilong operation sapos em, defence aeroplane, then, mi mas go bek long ahh&, through defense minister, na i go bek long ah& authority bilong dispela aeroplane em i Australian High Commission, a?

Sample 13 : Insertion : N, Conj, Discourse (Pron + Adv)

Participants: FIS tinfis_man, UET Uncle_ET_Presenter

Length: 30.650 seconds

*FIS: ah& mi lai totok long ahm& ### disuhluh ***koment*** ah& Jo vi mekim lo said lo tuna .

%ort : ah& mi laik toktok long ahm&, dispela comment ah& Jo bin mekim long said long tuna .

*UET: yeah ## .

*FIS: aaah& mi lai tok osem em no ken lidim nem blong ah& tuna we ol meikim lon Medeng ***bikos*** ah& ## olgeta tunafis uh& em i no kam lo Medeng tasol em blon awgeta sis blon PapuNuGini ## .

%ort : aaah& mi laik tok olsem em no ken lidim nem bilong ah& tuna we ol mekim long Madang because ah&, olgeta tuna fish uh& em i no kam long Madang tasol em bilong olgeta seas bilong Papua New Guinea .

*FIS: ***dats wai*** mi lai tok osem saos em i mekim ## litimapim neim blo displuh tinfis blon PapuNuGini # em i orait ***bigos*** olgeta tinfis em ogeta tuna i kam blong # olgeta provinsis # ol solwatuh blo olgeta provis blo PapaNuGini ## .

%ort : that is why mi laik tok olsem sapos em i mekim, litimapim nem bilong dispela tinfis bilong Papua New Guinea, em i orait because olgeta tinfis em olgeta tuna i kam bilong, olgeta provinses, ol solwara bilong olgeta provinces bilong Papua New Guinea .

*FIS: em tasol mi laik mekim .

%ort : em tasol mi laik mekim .

Sample 14 : Insertion : V

Participants: ABP Researcher, TTT Madang_T

Length: 17.554

*ABP: *uhl uhran dhuh wuhl # priveni is feimuhs fuh having ## .*

%ort : all around the world, P New Guinea is famous for having .

*TTT: i lavim ## .

%ort : he lavim .

*ABP: uhhh .

*TTT: yap ## .

%ort : yes .

*ABP: *al thing iss wuhn siksth ## ov uhl dhuh langwidjis ## in dhuh wuhl dhei* stuhp *uhp ## in is kuhntri ### .*

%ort : I think it is one sixth, of all the languages , in the world they stap up in this country .

*ABP: *nau wuhn siksth sikstin pesent uh aul dhuh wuhld langwidjis dhuh* stuhp *in PapuNuGini .*

%ort : now one sixth sixteen percent of all the world languages they stap in Papua New Guinea

Sample 15 : Insertion : N, Discourse (Pron+V), V, Adv - Intergrated

Participants: UET Uncle_ET_Presenter, CAL caller

Length: 6.316 seconds

*UET: olsem mi toto lo wuhnpla ol *lisna* lo monin *ah min* ol i ol *diskaradjim* wanpla na *enkaradjin* narapla # *indairektli .*

%ort : olsem mi toktok long wanpela ol listener long morning I mean ol i ol discourage-im wanpela na encourage-im narapela, indirectly .

*CAL: em nau .

*UET: mmm& .

Sample 16 : CLx? : Insertion : V, Adj+N, mixed VP, Conj, V – unintegrated, N

Participants: MPT Member of Parliament T
Length: 17.601

*MPT: mi laik minista cough& osem #### ah& ***rekonaisim*** disuhla ***bawduh epot*** ### inaf ### nesuhl gamuhn ## mi ***stil muhs*** ## ***fanim*** # ***an olso*** ### inuhf lo askim Oseid ##### Oseid long ***fuhndin apgreidin*** ### i go lo # ol disela ***epots*** .

%ort : mi laik minister, olsem, ah& recognise-im dispela border airport, inap, national government, mi still must, fund-im, and also, inap long ask-im Ausaid, Ausaid long fund-im upgrade-im, i go long, ol dispela airports .

Sample 17 : CLx? : Alternation : mixed, Pron, Question; Insertion? Pronouns

Participants: MNM MNM_mother
Length: 28.360 seconds

*MNM: ***firs ting thed in tuhms of yuu pipuhl sed eit tuhti and wi astatinga miting at ten oklok*** yu ken savi olsem wuhnem mipla ***gat ada things tu du*** ### .

%ort : first things third in terms of you people said eight thirty and we are starting a meeting at ten o'clock yu ken savi olsem wanem mipela gat other things to do .

*MNM: ***aam*** no isi long hapim ol nau # gimmi pon ## .

%ort : I'm no isi long hapim ol nau, give me point .

*MNM: ***yuu tel mi dhis lausi ## grr tradisuhn fain mi fo twentieit kina an # dhis chip shuht det al am werin fo twenti kina*** em wanem ?

%ort : you yell me this lousy, grr& tradition fine me for twenty eight kina and, this cheap shirt that I am wearing for twenty kina em wanem ?

*MNM: ## oh& mi tokim ***im*** laughs& .

%ort : oh& mi tokim him .

Sample 18 : Insertion : Island (NP, predicate), Adj+N

Participants: QKM Quikmeri_Announcer

Length: 14.211 seconds

*QKM: yes wan lili isin tokim mi tetei olsem em i ***wuhld hel dei*** # em i taim we ***lai neishins*** i maakim olsem taim bai yumi maakim # ol wanem kain kain ah& ah& sait blong uh& ***wawk helt*** # yumi saw lukim luh isela isa graun so tetei # ah& em i ***wuhl hel dei*** emm ah& ***ilait neishuhns*** i maakim long em .

%ort : yes one liklik pisin tokim mi tedei olsem em i world health day, em i taim we United Nations i makim olsem taim bai yumi makim, ol wanem kain kain ah& ah& sait bilong uh& work health, yumi save lukim long dispela dispela graun so tedei, ah& em i world health day em ah& United Nations i makim long em .

Sample 19 : Insertion : N, Island (idiom)

Participants: MpM Member of Parliament M

Length: 10.588 seconds

*MpM: mi t save kamap ia em em mipla sa totok lo sampla gudpla nau strongpla toktok mila sa wuhkim # .

%ort : mi tu save kamap ia em em mipela save toktok long sampela gupela nau strongpela toktok mipela save wokim .

*MpM: ating mi ***howp*** osem beain sampluh mila muhs # paluhmuhn muhs la- larim na em ***muhs kuhm laif on divi*** lo suhmpluh taim .

%ort : ating mi hope olsem behain sampela mipela must, parliment must la- larim na em must come live on TV long sampela taim .

Sample 20 : PNGE

Participants: uMP Member of Parliament U

Length: 14.396 seconds

*uMP: *dhiis pipuhl hav fo meni taims #### trai tu krusifai dhuh liiduhs uhv dhis neisuhn dhuh paluhmuhn mas asuhet its othoriti an bring dem bifo duh privluhdgis komiti ## weduh it was duh intensin tu unduhmain palamuhnt !*

%ort : these people have for many times, try to crucify the leaders of this nation the parliament must assert its authority and bring them before the privaledges committee, wether it was the intention to undermine parliament !

Sample 21 : CLx? : Alternations : back and forth : flagged

Participants: TTT Madang_T, ABP Researcher

Length: 27.260 seconds

*TTT: *dhei stat* ## ah& ## olsem ### putim ol *buks teksbuks skulbuks* long # tok Bel a ?

%ort : they start, ah&, olsem, putim ol books textbooks schoolbooks long, tok Bel a ?

*ABP: uh& huh& .

*TTT: tok ples # Madang # # *bat Bel agen is* i nuh i n olsem ## langwids olgeta Madang *speeking* .

%ort : tok ples, Madang, but Bel again is i no i no olsem, language olgeta Madang speaking .

*ABP: # # # # yes # Bel a ?

*ABP: em nau .

*TTT: Madang em i ## em i narakain moa long planti hap long PapuaNuGini a ?

%ort : Madang em i, em i narakain mo long planti hap long Papua New Guinea a ?

*ABP: *yeah* .

*TTT: *bikos* Madang alon em # i guht planti *lengwids* .

%ort : because Madang alone em, i got planti language .

Sample 22 : CLx? : Insertions & Alternations : back and forth - idioms

Participants: YJo YumiFM_Jo, UET Uncle_ET_Presenter
Length: 31.300 seconds

*YJo: athin dhuh oul ***liduhsip eksesais*** blo ol lida blo yumi lo Parluhmuhn nau ya .

%ort : atin the whole leadership exercise bilong ol leader bilong yumi long Parliament nau hia .

*UET: mmm& .

*YJo: ating yeah yumi gat pinis a em wok long ol lo Giad Ga kamuhpim gudpla sindaun blo man lo ples ia ***evuhn kondisuhn*** lo ***yuman welbeing*** blo yumi luk # .

%ort : atin yes yumi gat pinis a? em wok long ol long Giad? Ga? kamapim gudpela sindaun bilong man long ples ia even condition long human wellbeing bilong yumi luk .

*UET: tru .

*YJo: ol man na meri blong isuhla kuhnri .

%ort : ol man na meri bilong dispela country .

*UET: mmm& .

*YJo: n a tin ***whateva da plein on da kard*** ol mas lulu lo ***longtem va den*** ## ah& ***an olsouh*** ol maas uh& traim ol ting samla ***proaktiv mezuhs*** # .

%ort : and I think whatever the playing on the card ol man lukluk long longterm rather than, ah& and also ol mas uh& traim ol ting sampela proactive measures .

*UET: rait .

*YJo: la ***wotim*** ol ***problems*** umi gat e umi ## ***aksperimentim*** na umi ***ekspiriensim*** awgeta ***laas twenifaiv yiuhs*** .

%ort : laik votim ol problems yumi gat na yumi experiment-im na yumi experience-im olgeta last twentyfive years .

*UET: rait .

%ort : right .

*YJo: ol *probuhms* yumi gat athing *its uhbuht taim* yumi ## mas lukluk lo *seipin ov* sampla *lontem # kaunsil wuhk den yu wen se* .

%ort : ol problems yumi gat ating it is about taim yumi, mas lukluk long shaping of sampela longterm, council work then you? when? say?

Sample 23 : CLx : back and forth...

Participants: RRR Madang_R, ABP Researcher

Language: 15.790 seconds

*RRR: yeah osem ## nem tula *wuhd* a ?

%ort : yes olsem, nau em tupela word a ?

*ABP: ## UH& # .

*RRR: nowuhn buai ### a ?

%ort :nowuhn? buai, a ?

*RRR: *bikos #### ovisi* uh& ######## *oviyu yo tembe is* # sanis # bipla trru #### *littuhl o topisin* .

%ort : because, overseas uh&, over you your tembe? is, senis, bigpela tru, little of tokpisin .

Sample 24 : Insertion : N, Island (idiom) - integrated

Participants: MPM Member of Parliament M

Length: 35.480 seconds

*MPM: mi ting olosem bai PapuaNuGini # bai fait lo narapla *independen* .

%ort :mi ting olsem bai Papua New Guinea bai paitim long narapela independence .

*MPM: #### blo wanem mogetuh *stouhs* ## olgeta liklik samtin oguhtuh plpuhl blo narapla ausaid i kan kisim ## .

%ort : long wanem olgeta stores, olgeta liklik samting olgeta people bilong narapela autsaid i ken kisim .

*MPM: nau olosem ### mi laik # tok olosem # mi sori tumas lo lukim #### .

%ort : nau olsem, mi laik, tok olsem, mi sori tumas long lukim .

*MPM: a tin plani man ol i toktok pinis long ol *benks* # blong yumi lo insaid lo PapuNuGini .

%ort : ating planti man ol i toktok pinis long ol banks, bilong yumi long insaid long Papua New Guinea .

*MPM: ##### yumi *uhp in guhvman* yu wok lo kirapim awlguhtuh kainkaind *inityushin* long # halefim ol *plpol* ## tasol yumi n puhl inuhp lo halivim #### hau m uhlipl i halivim #### .

%ort : yumi up in government yu wok long kirapim olgeta kainkain institution long, halavim ol people, tasol yumi no full inap long halavim, how mipela i halavim ?

*MPM: eh !

%ort : eh !

*MPM: ## no got *ansa* blong en .

%ort : no gat ansa bilong en .

Sample 25 : CLx? : Insertion : N – unintegrated, N, Conj, Adv, Prep, Conj, Discourse

Participants: YuS Caller_S, UET Announcer_Uncle_ET
Length: 31.579 seconds

*YuS: NCDC na sa luglug lo # ***problems*** insaid lo wanem ya ## ***sidi*** ya ### .

%ort : NCDC nau save lukluk long, problems insaid long wanem hia, city hia .

*YuS: ***so*** em tok lo ***espesli*** lo Gawdens maket a ?

%ort : so em tok long especially long Gordens market a ?

*UET: rait .

*YuS: em wanpla bigpla ***ezampol*** uh ### .

%ort : em wanpela bigpela example hia .

*YuS: taim wanluh # ol wok o wanem # stretim Gawdens makit stap ## ol man o go sali o samti ***autsaid*** stap # .

%ort : taim wanpela, ol wok o wanem, stretim Gordens market stap, ol man ol go salim ol samting autsaid stap .

*YuS: ***at*** taim yu ***lik autsids kondisin*** ia ## .

%ort : at taim yu luk outsides condition hia .

*UET: un& hm& .

*YuS: em no gudpla ## .

%ort : em no gupela .

*YuS: lo said blo ***helt*** i lukim i no gudpla lo man salim keik anta lo pipia na b ol man lo baim no keke ### .

%ort : long said bilong health i lukim i no gudpela long man salim cake antap long pipia na bai ol man long baim na kaikai .

*YuS: ***so*** ### osem mi # ***putin*** suh lili ***bikos*** eNSiDiSi mas luglu lo disuhla stretim klinim ***so det*** ## plani ol man ol salim ## em bai # lo gudpla ***elti*** ol man kan baim lo keke .

%ort : so olsem mi, puting dispela liklik because NCDC man lukluk long dispela stretim klinim so that, planti ol man ol salim, em bai, long gudpela healthy ol man kan baim long kaikai .

Sample 26 : Alternation : repetition and back

Participants: RRR Madang_R

Length: 8.081 seconds

*RRR: ***wen al wuhs*** ah& #### taim mi mangi iet ah& ***pidjin wuhs ah& ### a foruhn tang ### .al had tu luhn hau tu spik it .***

%ort : when I was ah ... taim mi mangki yet ah pidgin was ah ... a foreign tongue . I had to learn to speak it .