This book provides an in-depth analysis of the different ways in which bilingual speakers switch from one language to another in the course of conversation. This phenomenon, known as code-mixing or code-switching, takes many forms. Pieter Muysken adopts a comparative approach to distinguish between the different types of code-mixing, drawing on a wealth of data from bilingual settings throughout the world. His study identifies three fundamental and distinct patterns of mixing – ‘insertion’, ‘alternation’ and ‘congruent lexicalization’ – and sets out to discover whether the choice of a particular mixing strategy depends on the contrasting grammatical properties of the languages involved, the degree of bilingual competence of the speaker or various social factors. The book synthesises a vast array of recent research in a rapidly growing field of study which has much to reveal about the structure and function of language.

Pieter Muysken is Professor of Linguistics and Latin American Studies at Leiden University. He is co-editor, with Lesley Milroy, of One Speaker, Two Languages (1995) and has published widely on Andean linguistics, creole studies and language contact.
BILINGUAL SPEECH

A TYPOLOGY OF CODE-MIXING

PIETER MUYSKEN
This book is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without
the written permission of Cambridge University Press.

First published 2000

Typeset in Times 9.5/13pt    [c1]

A catalogue record for this book is available from the British Library

isbn 0521 77168 4 hardback

Transferred to digital printing 2002
CONTENTS

List of figures vi
List of tables vii
Preface xi
List of abbreviations xiv

1 The study of code-mixing 1
2 Differences and similarities between languages 35
3 Insertion 60
4 Alternation 96
5 Congruent lexicalization 122
6 Function words 154
7 Bilingual verbs 184
8 Variation in mixing patterns 221
9 Code-mixing, bilingual speech and language change 250

References 279
Author index 298
Subject and language index 304
FIGURES

1.1 Schematic representation of the three main styles of code-mixing and transitions between them \hspace{1cm} page 9
1.2 Schematic representation of two ways of conceptualizing the relation between various mixing patterns 31
1.3 The present classification of language interaction phenomena as compared with that of Myers-Scotton and Poplack 32
8.1 Localization of a number of contact settings involving Dutch in the triangle alternation, insertion, congruent lexicalization 245
8.2 Localization of a number of contact settings in the triangle alternation, insertion, congruent lexicalization 246
8.3 Factors governing the choice of a particular code-mixing strategy in different communities 248
9.1 Typology of language contact phenomena 265
# TABLES

1.1 Classification of code-mixing theories in terms of the notions of asymmetry, word order, categorical equivalence, peripherality, and functional elements  

| 3.1 | Features of code-mixing and borrowing according to Sankoff and Poplack (1984) |
| 3.2 | Language choice for animal names, differentiated for grammatical context and morphological shape (based on Muysken, Kook, and Vedder 1996) |
| 4.1 | Number, percentage of total number of borrowings, and type–token ratio (TTR) of French words in Brussels Dutch (Trefers-Daller 1994) |
| 4.2 | The use of French discourse markers in Shaba Swahili/French bilingual speech |
| 5.1 | Examples of two paradigms in sociolinguistic research: the variation paradigm and the code-mixing paradigm |
| 5.2 | Multi-word switches in the Ottersum corpus (based on Giesbers’ tables 4.3.2 and 4.3.5) |
| 5.3 | Single-word switches in the Ottersum corpus (based on Giesbers’ tables 4.3.2 and 4.3.5) |
| 5.4 | Single Dutch elements in the Sranan corpus (based on Bolle’s tables 5.2 and 5.11) |
| 6.1 | The classification of Dutch elements in terms of a number of the criteria listed |
| 6.2 | Myers-Scotton’s system for distinguishing between system (s) and content (c) morphemes (1993b: 101) |
| 6.3 | Functional elements particularly affected in different linguistic domains |
| 6.4 | English verbs in a Spanish grammatical context (adapted from Pfaff 1979: 299) |
List of tables

7.2 Detailed analysis of occurrences of doubling with Spanish verbs that occur in four *waynos* or more 190
7.3 Non-loans contrasted with low and high frequency loans; independent occurrences taken together with occurrences in first position of a doublet 190
7.4 The constituency of mixed verbal compounds in Sarnami (based on Kishna 1979) 199
7.5 The form of the verb in mixed verbal compounds in Sarnami (based on Kishna 1979) 201
7.6 The form of mixed verbal compounds involving a verb + particle combination in Sarnami (based on Kishna 1979) 202
7.7 The constituency of mixed verbal compounds in American Greek (based on Seaman 1972) 212
7.8 The use of the Turkish verb *yap-* ‘do’ in three generations of Turkish migrants in the Netherlands (based on data in Backus 1996: 192, table 5.1) 213
7.9 Summary of the arguments given so far, and their implication for the nominalization, inflection-carrier, and adjunction analyses 218
7.10 Classification of the bilingual verbs in the different settings in terms of the notions insertion (finite, infinitive, compound), alternation, and congruent lexicalization 220
8.1 Diagnostic features of the three patterns of code-mixing 230
8.2 Relative proportion of different categories involved in mixes from Spanish to English in Pfaff (1979) and Poplack (1980) (discounting 747 single noun mixes in Pfaff’s table 5) 232
8.3 A quantitative survey of the sample sentences cited in the major sources for Spanish/English code-mixing in the United States 233
8.4 Code-mixes in Spanish/English in New York (based on Poplack 1980: 602), but with the percentages calculated across the switch directions rather than for the categories switched in one direction, as in the original table 235
8.5 Directionality of multi-word switches in Moroccan Arabic (MA)/Dutch (Du) code-switching (based on Nortier’s table 7.4, 1990: 126) 236
8.6 Directionality of single word switches in Moroccan Arabic
List of tables

(MA)/Dutch (Du) code-switching (based on Nortier’s table 7.15, 1990: 141) 237

8.7 Features of the corpus of Moroccan Arabic–Dutch switching (based on Nortier 1990) 238

8.8 Single-word mixes in the Nairobi corpus (based on Myers-Scotton 1993b) 239

8.9 EL islands (multi-word mixes) in the Nairobi corpus (based on Myers-Scotton 1993b) 239

9.1 The relation of the three mixing strategies to the processes of contact-induced language change 274
This book is an attempt to analyse the recent work on code-mixing and code-switching from a single grammatical perspective and to relate it to the general study of grammar contact. Since code-mixing research is being carried out in many different places and by many different researchers, by necessity my book cites many studies by others. Rather than mostly reporting on my own data, I discuss a great many data sets (many of which were gathered in recent master’s and doctoral dissertations in the Netherlands) and analyses. The reason for this is that I feel this is the only way to advance at the present stage, which is characterized by a proliferation of case-studies in many locations. Thus I want to use this occasion to gratefully acknowledge my indebtedness to all researchers in the field of language contact studies, and apologize beforehand for remaining errors in the interpretation of their data and views.

In early 1981 I was working with two colleagues in Montreal: Anna-Maria DiSciullo, who had participated in a study of the Italo-Quebecois speech community and is also a member of this community, and Rajendra Singh, expert on Hindi/English code-mixing and an accomplished switcher himself. We applied the theory of grammatical government to bilingual speech data and proposed that language mixing was structurally constrained by the government relation, just then being introduced into generative linguistics. In later research this constraint turned out to make the wrong predictions, but the insight behind it still has some value, independent of the theoretical framework adopted. Here I will try to use notions that are as little theory specific as possible, although I cannot eradicate my background in theoretical generative linguistics. This book could not have been written without my participation in the Network on Code-Switching and Language Contact of the European Science Foundation (1990–1992). This network involved a number of meetings which did much to stimulate interest in this area and to bring together interested researchers. For me it was an intense learning experience. I want to thank all participants with whom I had the opportunity
Preface

to discuss the issues at hand, Penelope Gardner-Chloros and Georges Lüdi for setting up the network, and Wouter Hugenholtz and Pat Cosgrove for keeping it intact.

I want to dedicate this work to my students. A number of students in Amsterdam wrote master’s and doctoral theses that have been important for me both because of the insights gained and for the rich data sets made available. In chronological order these include:

- Sita Kishna worked on Sarnami (Hindustani) and argued in detail that there is no difference between insertional code-mixing and borrowing;
- Rob Crama and Heleen van Gelderen studied Dutch/English code-mixing and underlined the importance of homophonous diamborps in facilitating mixing;
- Jacomine Nortier, in her careful study of Moroccan Arabic/Dutch code-mixing tested many constraints in the literature against her data and demonstrated the complete empirical inadequacy of the government constraint;
- Jeanine Treffers-Daller showed the importance of peripherality in Brussels Dutch/French code-mixing: many instances turned out to involve the margins of the clause rather than core elements;
- Liesbeth Adelmeijer focussed on Dutch/English code-mixing and argued for the importance of looking at surface rather than underlyng word order in accounting for mixing possibilities;
- Peter Bakker showed in his study of Michif, the language of Cree-French métis that emerged in the early nineteenth century in Canada, that out of mixing a whole new language can emerge;
- Silvia Kouwenberg, in her description and analysis of Berbice Dutch Creole, showed the importance of surface convergence patterns in language change;
- Jette Bolle demonstrated how diverse the patterns of mixing become when two languages in contact – Surinamese Dutch and the version of Sranan creole spoken in Amsterdam – are very similar;
- Rosita Huwač, in her work on Moluccan Malay/Dutch language contact underlines the historical dimension: different layers of contact varieties coexist in this very complex bilingual community;
- Vincent de Rooij studied Swahili/French mixing patterns and carefully analysed the French discourse markers in Swahili
Preface

bilingual speech, arguing that saliency contributes to the switching of French conjunctions through paratactic adjacency.

Their work is more explicitly cited in the text and referenced in the bibliography. I have tried to give detailed citations for examples of code-mixing cited from various studies, except where I was fortunate enough to be able to use some of the raw data collected by various researchers. In these cases I have cited the main publication coming out of that research. Here I want to thank particularly Rosita Huwæ, Sita Kishna, Vincent de Rooij, and Jeanine Treffers-Daller for making their data available to me.

I should also mention the work of several fellow researchers from other Dutch universities: Herman Giesbers (Nijmegen) carefully collected and analysed materials on Dutch dialect/standard code-switching; Rik Boeschoten (Tilburg) pointed to the importance of asymmetry in mixing patterns, particularly of first generation migrants; Ad Backus (Tilburg) was the first to explore the variety of mixing patterns found in a number of different networks among several generations of Turkish migrants; Henk de Wolf (Utrecht) initiated the modern study of Frisian/Dutch code mixing; finally Louis Boumans (Nijmegen) studied the dynamics in the evolving Moroccan Arabic/Dutch bilingual community.

Other names of colleagues that immediately come to mind for special thanks include René Appel, Hugo Baetens-Beardsmore, Abdeláli Bentahila, Lynne Drapeau, Penelope Gardner-Chloros, François Grosjean, Helena Halmai, Roeland van Hout, Annick De Houwer, Georges Lüdi, Jürgen Meisel, Lesley Milroy, Carol Myers-Scotton, Carol Pfaff, Petr Pitha, Shana Poplack, Henriette Schatz, André Tabouret-Keller. In addition I want to thank a number of generations of undergraduate students at the Linguistics Department of the University of Amsterdam for suffering through a series of confused lectures and seminars full of half-baked ideas. Vincent de Rooij, Jeanine Treffers-Daller, Carol Myers-Scotton, Rik Boeschoten, Suzanne Romaine, Martin Haase, and various anonymous readers for the publisher commented on earlier versions of this manuscript.

Finally, I want to thank the Netherlands Institute for Advanced Studies (NIAS) for making it possible to complete a first draft of this book during the academic year 1995/96 in an atmosphere both relaxing and inspiring, and Pilar van Breda-Burgueño for helping with corrections of the manuscript. Citi Potts of Cambridge University Press caught a great number of the remaining errors.

xiii
ABBREVIATIONS

A     adjective
ABL    ablative case
AC    accusative case
ADV    adverb
AF    affirmative
AG    agentive
agr    agreement
ALL    allative
APPL    applicative
ASP    aspect
AUX    auxiliary
BEN    benefactive
C    clause
CAU    causative
CIS    cislocative (near or toward speaker)
CONJ    coordinating conjunction
COMP    complementizer, subordinating conjunction
CONN    connective
CONSEC    consecutive
COP    copula
CPR    co-preterite
DA    dative
DEF    definite
DEM    demonstrative
DET    determiner
DIM    diminutive
DO    direct object
DUB    dubitative
DUR    durative
EL    embedded language
xiv
**List of abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>EMPH</td>
<td>emphatic</td>
</tr>
<tr>
<td>ESS</td>
<td>essive case</td>
</tr>
<tr>
<td>Excl</td>
<td>exclusive</td>
</tr>
<tr>
<td>F</td>
<td>feminine</td>
</tr>
<tr>
<td>Fr</td>
<td>formative</td>
</tr>
<tr>
<td>FOC</td>
<td>focus</td>
</tr>
<tr>
<td>FU</td>
<td>future tense</td>
</tr>
<tr>
<td>GE</td>
<td>genitive</td>
</tr>
<tr>
<td>HAB</td>
<td>habitual</td>
</tr>
<tr>
<td>I-</td>
<td>internal (language)</td>
</tr>
<tr>
<td>IA</td>
<td>intransitive action</td>
</tr>
<tr>
<td>ILL</td>
<td>illative</td>
</tr>
<tr>
<td>IM</td>
<td>imperative</td>
</tr>
<tr>
<td>IMPF</td>
<td>imperfective</td>
</tr>
<tr>
<td>IND</td>
<td>indicative</td>
</tr>
<tr>
<td>INDEF</td>
<td>indefinite</td>
</tr>
<tr>
<td>INF</td>
<td>infinitive</td>
</tr>
<tr>
<td>INFL</td>
<td>inflection (node)</td>
</tr>
<tr>
<td>INS</td>
<td>instrumental</td>
</tr>
<tr>
<td>INT</td>
<td>intensive</td>
</tr>
<tr>
<td>IO</td>
<td>indirect object</td>
</tr>
<tr>
<td>IP</td>
<td>inflection phrase</td>
</tr>
<tr>
<td>LO</td>
<td>locative</td>
</tr>
<tr>
<td>M</td>
<td>masculine</td>
</tr>
<tr>
<td>ML</td>
<td>matrix language</td>
</tr>
<tr>
<td>N</td>
<td>noun</td>
</tr>
<tr>
<td>NC1, 2, …</td>
<td>noun class 1, 2, …</td>
</tr>
<tr>
<td>NEG</td>
<td>negation element</td>
</tr>
<tr>
<td>NOM</td>
<td>nominalizer</td>
</tr>
<tr>
<td>NON-PST</td>
<td>non-past (often present) tense</td>
</tr>
<tr>
<td>O</td>
<td>object</td>
</tr>
<tr>
<td>P</td>
<td>adposition</td>
</tr>
<tr>
<td>PAR</td>
<td>partitive</td>
</tr>
<tr>
<td>PART1, 2, …</td>
<td>participle 1, 2, …</td>
</tr>
<tr>
<td>PERF</td>
<td>perfective</td>
</tr>
<tr>
<td>PL</td>
<td>plural marker</td>
</tr>
<tr>
<td>POSS</td>
<td>possessive</td>
</tr>
<tr>
<td>pp</td>
<td>past participle</td>
</tr>
<tr>
<td>PP</td>
<td>adpositional phrase</td>
</tr>
<tr>
<td>PR</td>
<td>progressive aspect</td>
</tr>
</tbody>
</table>

© Cambridge University Press  www.cambridge.org
List of abbreviations

PRD  predicate
PREP  proposition
PRES  present tense
PRET  preterite
PST  past
Pron  pronoun
Q  quantifier
RC  root change
RE  reflexive
REL  relative clause marker
REP  repetition
RES  resultative
S  subject
SD  sudden discovery tense
SUB  adverbial subordinator
TA  transitive action
TO  topic
V  verb
VP  verb phrase
W  word
Y/N  yes/no question
lex  first person plural exclusive
Ipl  I plural
Is  I singular