

A blurred, artistic photograph of a crowd of people, possibly at a festival or protest, with a bright light source in the upper left. The image is in shades of blue and purple.

OXFORD

LANGUAGE IN SOCIETY

An Introduction
to Sociolinguistics

Suzanne Romaine

second edition

Language in Society

This page intentionally left blank

Language in Society

An Introduction to Sociolinguistics

Second edition

Suzanne Romaine

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford OX2 6DP

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide in

Oxford New York

Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Taipei Tokyo Toronto

Oxford is a registered trade mark of Oxford University Press
in the UK and in certain other countries

Published in the United States
by Oxford University Press Inc., New York

© Suzanne Romaine 1994; 2000

The moral rights of the authors have been asserted

Database right Oxford University Press (maker)

First edition published by Oxford University Press 1994

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
without the prior permission in writing of Oxford University Press,
or as expressly permitted by law, or under terms agreed with the appropriate
reprographics rights organization. Enquiries concerning reproduction
outside the scope of the above should be sent to the Rights Department,
Oxford University Press, at the address above

You must not circulate this book in any other binding or cover
and you must impose this same condition on any acquirer

British Library Cataloguing in Publication Data

Data available

Library of Congress Cataloging in Publication Data

Data available

ISBN 0-19-873192-2

10 9 8 7 6 5

Typeset by RefineCatch Limited, Bungay, Suffolk

Printed in Great Britain by Biddles Ltd

www.biddles.co.uk

Contents

<i>List of Figures</i>	VII
<i>List of Tables</i>	VIII
<i>Preface</i>	IX
<i>Acknowledgements</i>	XII
Chapter 1 Language in Society/Society in Language	1
Language v. dialect	2
Language and dialect in Papua New Guinea	2
Language and dialect in Europe	11
English: language and dialect	17
Accent v. dialect	19
Register and style	21
Speech communities and communicative competence	23
Language, society, and reality: different words, different worlds?	26
Chapter 2 Language Choice	32
Societal multilingualism	35
Domains	44
Diglossia	46
Language shift and death	49
Code-switching	55
Chapter 3 Sociolinguistic Patterns	64
Language and social class	67
Style	74
Language and style	75
Gender	78
Language and age	82
Language and social network	83
Standardization	87
Standardization and its effects in Papua New Guinea	92
Chapter 4 Language and Gender	101
Sex and gender	103
Man-made language?	106
Learning to talk like a lady	122

VI CONTENTS

Gossip v. shop talk	126
Language reform: a misguided attempt to change herstory?	128
Chapter 5 Language Change in Social Perspective	135
Dialectology and language change	135
The relationship between variation and change	142
Using the present to explain the past and the past to explain the present	144
Change in real v. apparent time	150
Language change and social ideology	155
Chapter 6 Pidgin and Creole Languages	167
Distribution	169
Origins and structure	177
The creole continuum	191
Pidgins and creoles in social context	193
Chapter 7 Linguistic Problems as Societal Problems	205
Language and educational failure	205
Difference v. deficit	213
Language differences in the classroom	214
Testing: who decides what is right?	219
Taking appropriate action	220
One flag—one nation?	224
Bilingual education: maintenance or assimilation?	227
Immersion or submersion	230
Semilingualism: a new deficit theory	233
Chapter 8 Conclusions	240
<i>References</i>	249
<i>Index</i>	263

List of Figures

1.1	Map of Papua New Guinea showing New Britain and the distribution of Austronesian and non-Austronesian languages	4
1.2	Languages of north-west New Britain	6
1.3	Dialect continua in Europe	12
1.4	Map of the south Balkan region	16
2.1	Factors affecting language choice for the Buang	47
3.1	Vowel patriotism: the Norwegian language struggle	64
3.2	The Queen's English?	73
3.3	Social and stylistic stratification of postvocalic /r/ in New York City	76
3.4	Woman caught in the act of /h/ dropping	80
4.1	Old maids and old butlers	108
4.2	House persons and cleaning persons	111
4.3	Why women don't want to be called by their first names	114
5.1	The operation of the High German sound shift	137
5.2	S-curve for language change	143
5.3	Map of the eastern United States showing speech areas and incidence of postvocalic /r/	153
5.4	Areas of England where /r/ may still occur in the speech of older, rural, working-class speakers	154
5.5	'Vous as un joli vouvou'	161
6.1	Sign in Tok Pisin in Papua New Guinea warning all cars to stop if they see the red stop sign	168
6.2	Map showing some of the world's pidgins and creoles	170
6.3	Map showing the Caribbean creoles	172
6.4	Three types of creolization	192
6.5	The Guyanese creole continuum	192
6.6	An advertisement for Sun Flower tinned fish in Tok Pisin	199
7.1	Images of standard English and Hawai'i Creole English (HCE)	208
7.2	Papua New Guinean children playing in facilities provided by a mobile pre-school unit 'Skulnabaut' which operates in Tok Pisin	211
8.1	New languages as acts of identity	245
8.2	Jerusalem street signs	246

List of Tables

1.1	Some differences between American and British English	19
2.1	Conflicting census figures for Macedonia 1889–1905	38
2.2	Speakers of Irish according to census returns	42
2.3	Retention of languages by ethnic groups in Australia	50
2.4	Language choice patterns in Oberwart	52
3.1	Percentage of postvocalic /r/s pronounced	68
3.2	Percentage of non-RP forms in Norwich	69
3.3	Incidence of glottal stops in all environments compared to medial position	71
3.4	Verbs without -s	71
3.5	Regional and caste differences in Kannarese	74
3.6	Percentage of forms without final g	75
3.7	Multiple negation in Detroit	79
3.8	Percentage of /t/ forms in the speech of Eskilstuna schoolchildren and adults in relation to social class	82
3.9	Two Belfast women compared	84
4.1	Collocations with <i>man/woman</i> and <i>boy/girl</i> in the British National Corpus	110
4.2	Semantic feature analysis of <i>man</i> , <i>woman</i> , <i>boy</i> , and <i>girl</i>	112
5.1	Isoglosses between Low and High German consonant shift /p t k/ → /(p) f s x/	138
5.2	Implicational scale of isoglosses between Low and High German consonant shift /p t k/ → /(p) f s x/	139
5.3	Markedness metric for -t/d deletion	141
5.4	Temporal development of varieties for the rule of -t/d deletion	142
6.1	The conceptual field expressed by Tok Pisin <i>gras</i> compared with Warapu	189

Preface

MODERN linguistics has generally taken for granted that grammars are unrelated to the social lives of their speakers. Thus, linguists have usually treated language as an abstract object which can be accounted for without reference to social concerns of any kind. Sociologists, for their part, have tended to treat society as if it could be constituted without language. I have called this book *Language in Society*, which is what sociolinguistics is all about.

The term 'sociolinguistics' was coined in the 1950s to try to bring together the perspectives of linguists and sociologists to bear on issues concerning the place of language in society, and to address, in particular, the social context of linguistic diversity. Although it is still a young field of research, it gathered momentum in the 1960s and 1970s and continues to do so today. Educational and social policies played a role in the turning of linguists' attention to some of these questions, as did dissatisfaction with prevailing models of linguistics. Since the late 1950s mainstream linguistics has been conceived of as a largely formal enterprise increasingly divorced from the study of languages as they are actually used in everyday life.

Sociolinguistics has close connections with the social sciences, in particular, sociology, anthropology, social psychology, and education. It encompasses the study of multilingualism, social dialects, conversational interaction, attitudes to language, language change, and much more. It is impossible to put all the different approaches to the topic into neat pigeon-holes, each of which is distinct in terms of methodology, goals, etc. There is considerable overlap, so that for instance, while dialectologists have studied speech varieties and language change, subjects of paramount interest to many sociolinguists, they have generally employed quite different methods of data collection and concentrated on rural rather than urban speech (see Chapter 5).

Different authors writing about what has now become a very broad field have divided it up in various ways. Some distinguish, for instance, between theoretical and applied sociolinguistics. The former is concerned with formal models and methods for analysing the structure of speech communities and speech varieties, and providing a general account of communicative competence. Applied sociolinguistics deals with the social and political implications of fundamental inequalities in language use in various areas of public life, e.g. schools or courts. A glance at the two-volume work *Sociolinguistics: An International Handbook of the Science of Language and Society* (1987–8, Mouton de Gruyter),

which contains entries for nearly 200 topics, will give an indication of the multifaceted nature of the field.

More often, however, the field is subdivided into two broad headings: macro- and micro-sociolinguistics, with the macro domain sometimes also referred to as the 'sociology of language'. Macro-sociolinguistics takes society as its starting point and deals with language as a pivotal factor in the organization of communities. Micro-sociolinguistics begins with language and treats social forces as essential factors influencing the structure of languages. A recent example of this approach can be found in a two-volume work, one volume of which deals with what is referred to as the 'sociolinguistics of society' and the other with the 'sociolinguistics of language'. In his preface to the second volume, the author says he is not able to see much in common between issues about form and use of language on a small scale and large-scale socio-political issues. Thus, he presents sociolinguistics as a series of unconnected topics because he finds no common theoretical framework within which to link them.

I have always seen this division into two subfields as an artificial and arbitrary division of labor, which leads to a fruitless reductionism. It is no accident in my view that no convincing sociolinguistic theories exist. As long as scholars are prepared to ignore the forest for the trees, no theory is likely to be forthcoming.

Joshua Fishman, whose work is generally thought of as belonging to the sociology of language, said recently that the sociolinguistic enterprise is undergoing a mid-life crisis. Instead of progressing firmly on two legs (one propelled by linguistic matters and the other by sociological matters), it is trying to move ahead primarily on the linguistic front while merely shuffling on the social. He would like to see the 'socio' put into more prominence. One reason why I have called this book *language in society* rather than *language and society* is to emphasize the fact that the study of society must accord a place to language within it at the same time as the study of language must take account of society.

I cannot promise to produce a sociolinguistic theory in this book. Nevertheless, I hope that some of what I say might contribute to such a framework in the long term. What I offer in this short introductory text is an overview of the field by someone who has spent rather more time among the trees, while trying not to lose sight of the forest. The choice of which topics to include in a small survey of what is now a large and diverse field is to a great extent arbitrary. Every book inevitably reflects its author's assumptions about what is most interesting and important. I have made my choices based on those areas where there has been significant growth in terms of research findings, and also those areas where I have first-hand experience myself.

My research over the past fifteen years has involved me in trying to come to grips with problems of societal multilingualism, language change, and

language contact in the broadest sense, initially with respect to the status of the languages spoken by ethnic minorities in the UK and elsewhere in Europe, and more recently, through my research on the pidgin and creole languages of the Pacific, particularly in Papua New Guinea and Hawai'i. My recent work in Papua New Guinea, in particular, has convinced me that there are crucial connections between the large-scale socio-political issues typically addressed by the sociology of language on the one hand, and the forms and uses of language on a small scale dealt with by sociolinguistics on the other. They are manifestations of similar principles, albeit operating at different levels. Variability is inherent in human behavior.

In preparing the second edition of this book, I have benefited from discussions, comments, and reviews, and have incorporated some of the ideas which have emerged. However, the old adage about not being able to please all of the people all of the time, let alone even some of the people some of the time, very much applies to authors and their audiences. What one reviewer or colleague loved about the book, another hated. Chapter 4 on language and gender proved, not surprisingly, to be one of the most controversial. Since then, I have given that topic a book-length treatment, which made it even more difficult for me to confine myself to a chapter. Although some readers would have liked to see additional chapters on discourse and pragmatics, I have kept the same choice of topics. My main aim in this edition is to update the material to take account of works published since I first wrote the book in the early 1990s. Although I had hoped to keep this edition about the same length, it has inevitably ended up slightly longer.

Likewise, some readers liked the system of referencing I adopted with no in-text citations; others abhorred it, with one even considering that it set a bad example for students. This edition, however, does incorporate a general bibliography in addition to the annotated bibliographies at the end of each chapter, which I have correspondingly shortened.

Oxford
2000

S. R.

Acknowledgements

I WROTE the first edition of this book while I was a visiting professor in Sweden in 1991–2. My biggest debt is therefore to my colleagues at FUMS (Avdelningen för forskning och utbildning i modern svenska) in the Institute for Nordic Languages at the University of Uppsala and to the Swedish Research Council for the Humanities and Social Sciences for awarding me the Kerstin Hesselgren visiting professorship which made my stay possible and rewarding both intellectually and personally. My perception of fundamental sociolinguistic problems has been sharpened by my interaction with my colleagues at FUMS, whose interests and expertise span virtually the whole field of sociolinguistics.

I would also like to thank Isabel Forbes for sending me the article on French usage, and my colleague Olle Josephson at FUMS for bringing the case of *nörd* to my attention, both of which provide examples for my discussion of linguistic change in Chapter 6. Thanks also to Gunnel and André Melchers for discussion of the Swedish T/V system. I am also grateful to Nancy C. Dorian for helping me to locate various pieces of information I needed, and to Jim and Lesley Milroy, and John Rickford, for their helpful comments on a first draft of this book. Thanks also to the many readers who provided feedback.

Chapter 1

Language in Society/Society in Language

I NOTED in my Preface how prevailing trends in linguistics have marginalized the study of the social role of language. In discussing the differences between the concerns of sociolinguistics and mainstream linguistics, Noam Chomsky, who is the leading figure in theoretical linguistics, observed that sociolinguistics was not concerned with 'grammar' but with concepts of a different sort, among them perhaps 'language'. To this he added, 'if such an object can become an object of serious study'. Chomsky then goes on to say that questions of language are basically questions of power, but these are not the sorts of issues which linguists should address. He is certainly right about the former. The latter is a matter of opinion. The narrowing of modern linguistics to the study of grammar has ruled out investigation of many interesting questions about how language functions in society. This book is about some of these issues which form the subject matter of sociolinguistics, chief among them being the question of what we mean by a language.

I can't begin to estimate how many times people have asked me questions such as how many languages there are in the world, how many dialects of English there are, and whether American English is a language or a dialect of English. I am sure my answers are generally seen as unsatisfactory because I invariably reply that it depends on what we mean by terms such as 'language' and 'dialect' and that these are not linguistic but rather social matters. It may at first glance seem incredible to non-linguists that linguists cannot define such essential and basic concepts in purely linguistic terms. The purpose of this chapter is to explain why the notions of language and dialect are fundamentally social and not linguistic constructs. I will also introduce other concepts such as 'communicative competence' and say why these too are primary concerns of sociolinguistics because they depend on society in crucial ways.

Language v. dialect

The term 'dialect' has generally been used to refer to a subordinate variety of a language. For example, we are accustomed to saying that the English language has many dialects. These dialects may be of different kinds. A 'regional dialect' is a variety associated with a place, such as the Yorkshire dialect in England or the Bavarian dialect in Germany. Dialects of a language tend to differ more from one another the more remote they are from one another geographically. In this respect the study of dialects or dialectology has to do with boundaries, which often coincide with geographical features such as rivers and mountains (see Chapter 5 for further discussion). Boundaries are, however, often of a social nature, e.g. between different social class groups. In this case we may speak of 'social dialects' (see Chapter 3 for a discussion of these). Social dialects say who we are, and regional dialects where we come from.

The term 'dialect' also has historical connotations. Historical linguists, for instance, speak of the Germanic dialects, by which they mean the ancestors of language varieties now recognized as modern Germanic languages, such as English, Dutch, and German. The entities we label as the 'English language' or 'Flemish dialect' are not, however, discrete. Any variety is part of a continuum in social and geographical space and time. The discontinuities that do occur, however, often reflect geographical and social boundaries and weaknesses in communication networks.

Language and dialect in Papua New Guinea

A preliminary example from north-west New Britain in the Pacific region will illustrate the problems in applying purely linguistic criteria in deciding what counts as a language or dialect. The Pacific is a good place to begin because it is a vast area containing many indigenous languages, whose number must have been even greater before European contact. In many parts of the region there are extensive chains of interrelated varieties with no clear internal boundaries. The greatest concentration of diversity is found in Melanesia (an area comprising the south-west Pacific island nations of Papua New Guinea, the Solomon Islands, Vanuatu, New Caledonia, and Fiji), where up to 1,500 languages are spoken, with as many as half found in Papua New Guinea alone. Most of the languages in Papua New Guinea are spoken by small groups; probably 40 per cent have fewer than 500 speakers. There is a great diversity of language types and only a handful of these languages has been investigated in any detail. New

Britain is one of the larger islands in the Bismarck Archipelago off the north-eastern coast of the island of New Guinea, which lies just 100 miles north of the tip of Queensland, Australia. Politically, the islands are part of Papua New Guinea (independent since 1975), and the island of New Britain is divided into two provinces, East and West New Britain (see Fig. 1.1). In the part of north-west New Britain to be discussed here people live in small villages along the coast and in the interior. All are multilingual and it is not uncommon for people to be able to speak four or five languages.

The following ten examples illustrate how people in different villages would request someone to give them betelnut to chew. For the moment, let's use the term 'variety' as a neutral term which does not commit us to any decision about whether the varieties concerned have the status of language or dialect. The grammar is the same in all cases: first, the item desired is named (in this case, betelnut), then follows a third person singular form of the verb 'come', and finally, a first person verb phrase indicating what the person requesting the item is going to do with it. Literally, the request means 'betelnut, it comes, I chew', or loosely, 'give me some betelnut to chew'. Betelnut is the small green nut of the betel palm, which when chewed is a mild intoxicant (and also carcinogen). It is typically chewed with lime pepper and it turns the mouth a bright reddish-orange. Later, it is spat out. Sharing betelnut and other items such as tobacco or yams is culturally important in north-west New Britain and other parts of Papua New Guinea. Offering these items is a sign of friendliness on the part of those who give them, while accepting or requesting them indicates trust that a spell has not been cast over them.

1. <i>ezim</i>	<i>o-mên</i>	<i>da-kîn</i>
2. <i>eliep</i>	<i>max</i>	<i>nga-ngas</i>
3. <i>bile</i>	<i>me</i>	<i>nge-nges</i>
4. <i>bile</i>	<i>me</i>	<i>nga-nges</i>
5. <i>bile</i>	<i>me</i>	<i>nga-nges</i>
6. <i>vua</i>	<i>i-nama</i>	<i>nga-songo</i>
7. <i>vua</i>	<i>i-nama</i>	<i>nga-songo</i>
8. <i>bua</i>	<i>i-nam</i>	<i>nga-songo</i>
9. <i>vua</i>	<i>i-mai</i>	<i>nga-songo</i>
10. <i>eilep</i>	<i>i-me</i>	<i>a-ngas</i>
<i>betelnut</i>	<i>3 sing. come</i>	<i>1 sing. chew</i>

Let's for the moment try to sort these ten utterances into groups based on how similar they are to one another in terms of the words they use and see if we can make a guess at how many languages and dialects there are here in purely linguistic terms. We would certainly want to recognize the first variety as a separate language since it seems to share none of its vocabulary with any of the

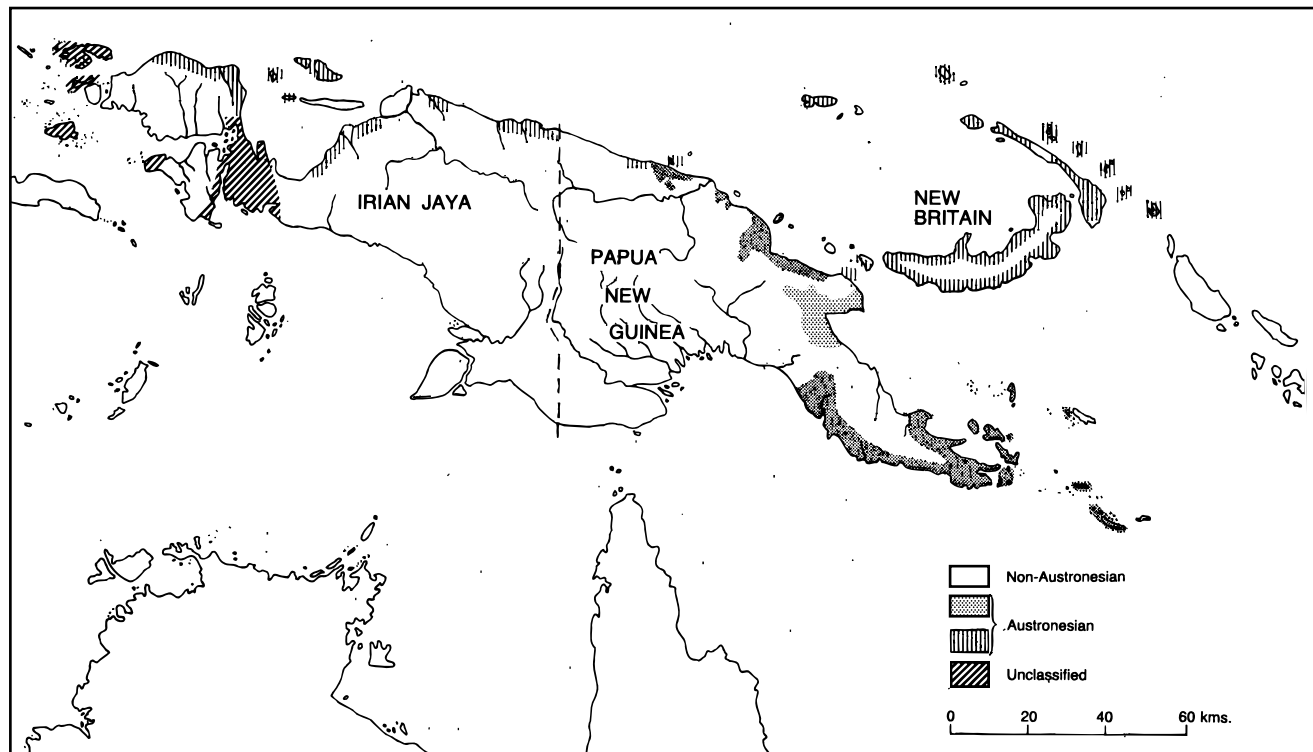


FIG 1.1 Map of Papua New Guinea showing New Britain and the distribution of Austronesian and non-Austronesian languages

other varieties, except possibly some remote similarity in the verb 'to come'. The other varieties, however, obviously have some lexical relationship to one another, though some more so than others. For instance, varieties (6) and (7) are identical, therefore it seems reasonable to suppose that the villages speaking these varieties do not have totally different languages, but rather dialects of the same language, or even the same language. Varieties (8) and (9) are also very similar to (6) and (7), differing only slightly in the pronunciations of the words for 'betelnut' and 'come'. So we might plausibly imagine that these four varieties constitute dialects of one language. Varieties (3), (4), and (5) also show a close relationship, differing only in terms of the vowels in the root and in the prefix for the verb 'chew', so we might consider them dialects of one language. Variety (10) is also not so very different, apart from its use of *eilep* instead of *bile* for 'betelnut' (which is similar to variety 2) and its lack of an initial consonant in the verb prefix for 'chew'.

There are some explicit linguistic procedures we could invoke to back up this impressionistic view. In fact, most of what is known about linguistic relationships in Papua New Guinea has relied on a measure called 'lexico-statistics', a method which still remains extremely popular because it provides a simple means of comparing the speech of different communities. The method relies on counting percentages of apparent cognates, i.e. related forms meaning the same thing, in a word list of 100 or 200 items. Those who use this method generally regard varieties sharing between 81 and 100 per cent cognates as dialects within a language. If there are between 28 and 81 per cent cognates, then the varieties count as languages within a family. Fewer cognates indicate a more distant relationship. These measures of course tell us nothing of what the speakers themselves consider the status of these language varieties to be.

When we ask what varieties the speakers themselves consider to be separate languages, we see that the linguistic evidence is interpreted in another way. We can get an answer to this question by looking at the names given to the ten varieties. In fact, all the varieties are recognized as separate languages each with its own name. They are shown in Fig. 1.2 in a grouping which is based on their linguistic similarities and their supposed historical relationship. The names used by the speakers are given here along with the numbers I used above.

Linguists generally recognize two major language families in Papua New Guinea comprising between 700 and 800 languages, Austronesian and non-Austronesian (or Papuan). We are still a long way from arriving at a generally accepted classification of these languages, particularly the non-Austronesian or Papuan group. It is, however, usually agreed that speakers of the latter group of languages arrived in Oceania long before the speakers of Austronesian languages. The coastal distribution of most of the Austronesian languages, which

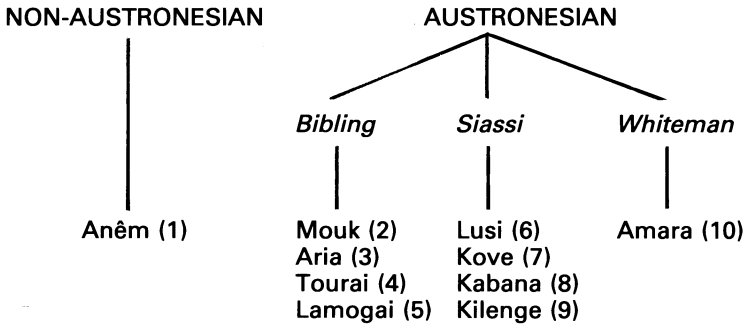


FIG 1.2 Languages of north-west New Britain

can also be seen in Fig. 1.1, is commonly taken to indicate the later arrival of their speakers.

The Anêm language spoken in north-west New Britain is a member of the non-Austronesian family and is not lexically related to any other language known in Papua New Guinea as far as we can tell at the moment. In fact, the Anêm-speaking people claim theirs is the original language of the area. This is probably correct. It is also likely that Anêm is the sole survivor of a group of non-Austronesian languages which were once spread over a wider area of New Britain. Thus, by anyone's criteria, Anêm has to be recognized as a separate language, at least at the level of vocabulary. It is at present the language of four villages, three of which are on the coast and the other in the interior, separated from the coastal speakers by Lusi territory. The interior Anêm speakers have intermarried with Mouk, Aria, and Lusi speakers. The latter speak languages which are all classified as Austronesian, although on the basis of a more precise linguistic analysis we can divide them into three subgroups: Bibling, Siassi, and Whiteman. However, these Austronesian languages are structurally unlike those spoken outside the Melanesian area. Languages which are geographically adjacent are the most similar and those languages that have been in close association the longest exhibit structural similarities which are due to contact rather than inheritance.

Mouk is most closely related to Aria and many Mouk, especially older men, know a fair amount of Anêm. Due to their interaction with Anêm speakers, Mouk and Anêm now share certain linguistic features to the exclusion of the neighboring languages. Within the Bibling group Aria, Lamogai, and Tourai are closely related and the people who call themselves Tourai consider their language to be intermediate between Aria and Lamogai but more closely related to Aria, an impression which is supported by linguistic evidence. Within the Siassi group Lusi and Kove are very similar and some linguists have taken Lusi and Kove to be dialects of one language, ignoring the claims of the speakers. Kilenge

is less closely related to Lusi, Kove, and Kabana. Amara, the only representative of the Whiteman group in this area, is apparently near extinction. The Amara and Kabana interact so closely with each other that Kabana is quickly becoming the language of first choice among young Amara children.

It is obvious that the boundaries reflected by the names given to these varieties are socially rather than linguistically constructed. But how is it that all these varieties emerge as entities autonomous enough to be named separate languages by their speakers when some of them are virtually identical from a linguistic point of view? Not surprisingly, early investigators were puzzled by the complex nature of linguistic relationships in the New Guinea area. Historical connections are complicated by such extensive long-term contacts and movements between Austronesian and non-Austronesian-speaking peoples which has also brought about drastic grammatical convergence. There is still some dispute over the membership of many languages since linguists have been reluctant to accept that a language might be 20 per cent Austronesian and 80 per cent non-Austronesian. Lexical relationships have been taken as primary in historical classification. Moreover, we can see that it is largely on the basis of lexical forms that the people of north-west New Britain see their varieties as being separate languages.

The names given by the early European government patrol officers to villages and census divisions do not always correspond to linguistic differences, although some have made this assumption. Most of the names given on maps are administrative names of subdistricts and cannot be relied upon to yield the linguistic composition of an area. Europeans often arbitrarily chose the name of one of the constituent clans for a group of clusters they took as members of a village. The term 'village', however, suggests a degree of centralization not found in traditional societies, where residential units were not compact. Villages are simply points of contact between the administration and the local populace, who may not actually occupy that particular place. From the European perspective, it seemed natural to expect that a place name, the names of people occupying that place, and the name of a language spoken there should be the same. That is what typically happens in Europe, where there is a close association between the names of countries, peoples, and languages, as can be seen in sets of names such as England/English/English, Germany/German/German, France/French/French.

However, in Papua New Guinea there is no one-to-one mapping between village names, groups, and language names. Sometimes up to four local names apply to what a district officer decided to call 'one' village, and some names are common to three or four sites. In addition, many groups have no special names for their languages. The Sare people of the Sepik, for example, call their language Sare, but this means simply 'to speak or talk'. There are probably many

more languages than the actual names indicate and the boundaries between them are fuzzy. Much of the early information contained in explorers' accounts and the reports of patrol officers is anecdotal. What is now called Kabana by some linguists was earlier called Barriai by one observer, which is a name given by the Kabana and Amara people to their land. Lusi was called Kaliai-Kove in an earlier description in recognition of the near identity of grammar between Kove and Lusi and also because the people of this area are commonly called Kaliai by people outside the region. Kaliai is the name of one of the Lusi clans. The Mouk-speaking village of Salkei was counted as Aria-speaking because fieldworkers unwittingly collected their data from an Aria speaker who happened to be visiting the aid post there rather than from a local resident.

Even if we had descriptions of all the varieties spoken in each village, it would still pose considerable problems to say which constituted languages and which dialects since the point of view of the speakers themselves will differ from the linguistic evidence. Moreover, the views of speakers themselves may vary a great deal. For example, we saw how Tourai and Aria could on linguistic grounds be thought of as dialects of a single language, but Tourai speakers (as well as those of Lamogai and Aria) consider them separate languages. The language spoken in Bolo village is also from a linguist's point of view identical to Aria, but Aria speakers from other villages say it is not Aria. They say Bolo villagers really speak Mouk. However, the people of Salkei village, who speak Mouk, say that Bolo people speak Aria. As for the Bolo themselves, they claim to be Anêm speakers! While traditionally this was true, only a handful of older men now know the language and people generally recognize that Anêm is no longer spoken there. For their part, the Anêm people do not think the Bolo speak acceptable Anêm anymore. Thus, this one village of Bolo, which is said to be Anêm-, Aria-, and Mouk-speaking, speaks a variety which no one else accepts as a legitimate member of their own language group.

The counterpart to this at the level of linguistic form is that there are words widely used in the region for intergroup communication which no one accepts as part of the 'true' vocabulary of their language. Each group speaking what it regards as a separate language has what it considers its own true lexicon, but there are also synonyms, which are identical in form to the words used by others speaking different languages. Here, for example, is a list of words (not complete) meaning 'wallaby' in some of the languages:

Anêm: *apose, gauxu, kis, nautus, zei*

Amara: *natus, kope, *kio*

Mouk: *natus, apose, sokolo*

Aria: *apare, *apose*

Tourai: **apare, apose*

Lamogai: *airok, apare, keneng*

Those marked with an asterisk are words which the people of that particular language consider as their own true word for 'wallaby', while the others are seen as equivalents. The use of such synonyms aids intergroup communication at the same time as it allows each group to claim at least one word as its own. This permits each group to maintain its distinctiveness. There are, however, other words known among the groups which no one claims as a true word of their own language. The extensive overlap in vocabulary from one language to another facilitates what has been called 'dual-lingualism', where speakers communicate with one another while each speaking their own language. It also facilitates active learning of other languages.

The Siassi languages function widely as *lingua francas* (i.e. additional languages used across wide areas for intergroup communication) in this area and most people speak either Lusi, Kove, or Kabana well. The Bibling languages also function dual-lingually so that knowing one gives access to understanding the others. Anêm and Amara, however, have no value as *lingua francas* since neither can be used dual-lingually with any other language. Since European contact in the nineteenth century, another language has been added to the villagers' repertoire, an English-based pidgin called Tok Pisin ('talk pidgin'), a language which developed from a contact language which emerged on plantations in Queensland, where Pacific Islanders worked as contract laborers. Recruiting was particularly heavy in the Bismarck Archipelago; at one time, few men were left in the villages. When the laborers returned, they often taught Tok Pisin to the younger male generation (see Chapter 6 for further information about Tok Pisin and other pidgins and creoles). Tok Pisin eventually became such a useful additional language that it spread throughout mainland New Guinea and today it is the most widely spoken *lingua franca* in the country. It has been used in New Britain for at least eighty years and all the languages of north-west New Britain have now incorporated Tok Pisin words into their vocabulary.

A number of explanations have been proposed for the existence of such extreme linguistic fragmentation of New Guinea. First, a span of 40,000 years of human habitation affords sufficient time-depth for natural processes of change and diversification to produce a multitude of languages. Secondly, the rugged nature of the terrain poses physical barriers to human social interaction. Thirdly, cultural attitudes play an important role in fostering and maintaining diversity. Diversity is cultivated in Melanesia as a badge of identification and is largely a conscious reaction. However, even more important, in my view, is that none of the pressures towards convergence found for a long time in Europe and elsewhere, such as literacy, standardization, centralized administrative control, schooling, media, was present to any great degree in precolonial days. These factors have, however, become relevant since European contact and have

avored the spread of *lingua francas* such as Tok Pisin and metropolitan languages such as English.

Traditional dialectologists believed that isolation led to linguistic diversity, while mixing of populations created uniformity. Yet it is clear that geography alone is not sufficient to explain diversity, nor is the concentration of people in one area or isolation a guarantee of uniformity or conservatism, as research on urban social dialects has shown (see Chapter 3). Compare, for instance, one of the areas of greatest diversity, the easily navigable Sepik River, with Enga in the interior Highlands, which has some of the most rugged terrain in the country and is an area with much less diversity. The north coastal areas, where there are more small, unrelated language groups, are more linguistically diverse than the more isolated areas of the interior.

The distribution of linguistic diversity implies demographic factors as additional contributory causes. It seems likely that the existence of endemic diseases such as malaria in the coastal lowlands restricted not only mobility, but also population growth. In the Highlands, which are largely malaria-free, language groups are larger than on the coast. Compare Enga which has over 150,000 speakers with Erima, a language spoken in only four villages, with only 400 speakers. The threat of disease therefore probably limited the possibility for groups to expand much beyond their immediate territory, and thus impeded the spread of any one language or group of languages.

Since the imposition of colonial administration in various Pacific islands such as Papua New Guinea, there has no doubt been a decline in the number of languages. Moreover, dramatic changes have taken place in those which survive, often through contact with major metropolitan languages. New *lingua francas* arose as a result of contact with Europeans; the most important of these were languages like Kâte and Yabem, spread by missionaries as church *lingua francas*, and Tok Pisin and Hiri Motu.

Any estimate of the number of languages spoken in an area like Papua New Guinea is fraught with difficulties due to the problems inherent in defining terms such as 'language' and 'dialect'. The very concept of discrete languages is probably a European cultural artefact fostered by processes such as literacy and standardization. Any attempt to count distinct languages will be an artefact of classificatory procedures rather than a reflection of communicative practices. Lexico-statistics will not yield any non-arbitrary technical definition of terms such as 'language', 'dialect', or 'family'.

Language and dialect in Europe

Other examples from Europe can be taken to illustrate the arbitrariness of linguistic criteria, and the importance of social factors in deciding what counts as a language or dialect. Some classic cases are the West Romance and Germanic dialect continua. The West Romance dialect continuum stretches through rural communities from the Atlantic coast of France through Italy, Spain, and Portugal. Mutual intelligibility exists between adjacent villages, although speakers of the standard varieties of French, Italian, Spanish, and Portuguese find one another mutually unintelligible to varying degrees. Similarly, the Germanic dialect continuum connects a series of historically related varieties that differ from one another with respect to one or more features.

Degree of mutual intelligibility is greatly affected by the extent of social and other contact between the groups concerned as well as their attitudes to one another and does not necessarily have much to do with lexico-statistical relationships. In Scandinavia, for instance, if a traveller knows Danish, Swedish, or Norwegian, it is possible to communicate across language boundaries. Certainly, linguistically the languages are very close, in fact close enough from a linguistic point of view to be considered dialects of one language. Indeed, structurally they form a nice parallel to the linguistic situation in parts of north-west New Britain because their grammar is very similar and most of the distinctive differences lie in vocabulary, and pronunciation, although the differences here are also not very great in many cases. Danish and Norwegian have a great deal of vocabulary in common, but differ in pronunciation, while Swedish and Norwegian differ more in vocabulary, but have a more similar pronunciation. In a sample of 1,000 words from Norwegian, as many as 50 per cent of the words are identical and another 25 per cent are variants of the same form. A further 15 per cent have the same pronunciation but are spelled differently, while only 10 per cent are essentially different. Some of the superficial similarities can be seen in these examples.

Danish: *Hun sidder i vinduet og ser ud over gaden.*

Norwegian: *Hun sitter i vinduet og ser ut over gatan.*

Swedish: *Hon sitter i fönstret och ser ut över gatan.*

The modern languages are derived historically from a common Nordic ancestor and their increasing fragmentation reflects political history. It is largely for political reasons that they are regarded as separate languages. By 1700 Swedish and Danish standards were firmly established, but Norway was still under Danish rule. When these languages were standardized, differences between



FIG 1.3 Dialect continua in Europe

them were consciously exaggerated. For instance, before 1906 all three languages wrote the word meaning 'what' unphonetically as *hvad*. Now only Danish does. Swedish spells it as *vad* and Norwegian as *hva*. Thus, orthographic differences now disguise what is a similar pronunciation and make the languages look more different in their written form than they are when spoken.

In studies of mutual intelligibility some interesting asymmetries emerge. Danes claim to understand Norwegians much better than Norwegians claim to understand Danes. The poorest understanding is between Danes and Swedes and the best between Norwegians and Swedes. However, Norwegians and

Danes claim to understand Swedes better than Swedes claim to understand either Norwegians or Danes. How can this be? Studies of mutual intelligibility are not really about linguistic relationships between varieties, but about social relationships since it is people and not the varieties who understand or do not understand one another. More Norwegians and Danes have been in Sweden than Swedes have been in the other two countries. Only one-quarter of Swedes claim to read anything in the other two languages. While 41 per cent of Danes and 52 per cent of Norwegians listen to Swedish radio, only 9 per cent of Swedes listen to Norwegian or Danish radio. It is obvious that more accommodation is made towards Sweden and Swedish by Danes and Norwegians because Sweden is a larger and wealthier country. It is much more self-contained both economically and intellectually. Not surprisingly, Swedes show the least interest in Nordic cooperation, while the Danes favor it most, no doubt because they stand to gain the most from it. They are generally less well understood than Norwegians or Swedes.

A Danish school principal told the story of how she gave a lecture to an audience in Stockholm from a manuscript which had been translated into Swedish. She said, 'They understood me very well. Then I fumbled for an expression, and the audience cried out, "just talk Danish, you are so easy to understand". I switched to Danish, to the great surprise of the Swedes, who understood nothing! They had thought I was talking Danish all along.' Under the present political circumstances, convergence towards a common inter-Scandinavian speech form could be brought about only by conscious language planning and increased social contact.

The dividing line between the languages we call Swedish, Norwegian, and Danish is linguistically arbitrary but politically and culturally relevant. Max Weinreich's often quoted dictum, 'a language is a dialect with an army and a navy', attests the importance of political power and the sovereignty of a nation-state in the recognition of a variety as a language rather than a dialect. Situations in which there is widespread agreement as to what constitutes a language arise through the interaction of social, political, psychological, and historical factors, and are not due to any inherent properties of the varieties concerned. In China, a range of mutually unintelligible varieties which a linguist would certainly call separate languages are nevertheless considered dialects of Chinese because they are linked by a common writing system. While speakers of varieties such as Cantonese and Mandarin would not be able to communicate in the spoken language, they would share the same writing system. Each would write the same symbol for the 'same' words, even though their spoken forms would be completely different. This is really in a sense the opposite of what exists in Scandinavia, where differences in writing obscure basic similarities in pronunciation. Many speakers of what is sometimes called

Serbo-Croatian, the standard language of the former Yugoslavia, and current standard of the Federal Republic of Yugoslavia, say it is really two languages rather than one because Croats write their variety in Roman script, while Serbs write theirs in Cyrillic.

Certain varieties of the West Germanic dialect continuum are considered to be dialects of Dutch and others dialects of German because of the relationship these varieties have to their respective standard languages. The process of standardization is connected with a number of socio-historical factors such as literacy, nationalism, and cultural and ethnic identity. It results in the selection and fixing of a uniform norm of usage, which is promoted in dictionaries, grammars, and teaching. A standard language is a variety that has been deliberately codified so that it varies minimally in linguistic form but is maximally elaborated in function. Most of the European languages became standardized under periods of intense nationalism. The standardization and promotion of a common language was seen as an important symbol of the process of political unification. Some of the consequences of this will be discussed in Chapters 3 and 7.

Some linguists have found the terms 'autonomous' and 'heteronymous' speech varieties useful as alternative labels to language and dialect. Thus, we can say, for instance, that the Dutch dialects are dependent on or 'heteronymous' with respect to standard Dutch, German dialects to standard German, etc. This means that because speakers of German watch German TV, are taught standard German in school, read in standard German etc., they look to standard German as a reference point. There will be more linguistic similarities between the varieties of German and Dutch spoken close to the border between those two countries than there will be between standard German and standard Dutch. Nevertheless, a standard language serves to create a feeling of unity among the speakers who take it as a reference point, particularly those who speak varieties far removed from one another geographically. Thus, for speakers of Cockney English in London and speakers of local dialects in Tyneside, in the north-east of England, the linguistic differences may be so substantial as to prevent easy spoken communication, but both groups would say in a larger sense they 'speak the same language' since both have standard English as a superordinate variety.

The term 'language' is employed for a variety that is autonomous, together with all those varieties that are heteronymous upon it. Because heteronymy and autonomy reflect political and cultural rather than purely linguistic factors, they can change. Often due to political developments formerly heteronymous varieties can achieve autonomy, as is the case with Afrikaans in South Africa, which was standardized in the 1920s and recognized as a language and not a dialect of Dutch. Conversely, autonomous varieties may lose their

autonomy, as Scots English did when it ceased to function as the language of the Scottish court after the Union of Crowns in 1603. There is nothing inherently better about a variety which achieves autonomy, and autonomy can always be challenged. Political and social factors are responsible for the selection of one out of many varieties which could have been candidates for standardization (see further in Chapter 3).

Disputes about the status of a variety are often used to bolster claims about the ethnic membership of the speakers or political status of territory with which it is associated. The present-day nation calling itself Macedonia was once the heart of an empire that stretched from Gibraltar to the Panjab, and its 2 million people live on land that until 1991 was part of Greece, Yugoslavia, Bulgaria, and Serbia, all of whom have taken chunks of its territory. When Macedonia was one of five constituent republics of Yugoslavia (along with Serbia, Montenegro, Croatia, Bosnia-Herzegovina, Slovenia) from 1944 to 1991, Macedonian was subordinate to Serbo-Croatian. Extremist Serbian politicians in now independent Serbia have publicly denied the legitimacy of Macedonian language and nationality and called for the reabsorption of Macedonia into Serbia. Although Bulgaria recognized the independent republic of Macedonia in 1992 after the break-up of the Socialist state of Yugoslavia, it still insists that Macedonian is a dialect of Bulgarian.

Meanwhile, neighboring Greece, which claims Alexander the Great (b. 356 BC), probably the most famous Macedonian, as the founding father of Greek nationhood, has always officially denied both Macedonian nationality and language, insisting instead that Ancient Macedonian was a dialect of Greek. It regards the existence of an independent Macedonia as an assault on Greek sovereignty. When Macedonia became independent in 1991, Greece blocked its entrance to the United Nations under its own name because Greece uses the same name to refer to a province of its own territory and claims exclusive right to the name *Makedonia*. The extension of Greek national sovereignty over parts of the territory of Makedonia was accompanied by particularly aggressive measures aimed at Hellenization of the Slavic-speaking population, among them the prohibition of the use of any language but Greek in public. People were fined, sent to prison, or forced to drink castor oil, and children thrashed at school, if they were caught speaking their own language.

In 1993 newly independent Macedonia joined the UN under the rather unwieldy euphemism 'Former Yugoslav Republic of Macedonia'. Greece imposed an economic blockade that was not lifted until Macedonia agreed to alter a disputed symbol in its flag. Greek authorities also grounded a Macedonian passenger aircraft until the word Makedonia on its fuselage was painted over.

What light can historical and linguistic evidence shed on these competing claims from Greece, Serbia, and Bulgaria? The Macedonians are descended from



FIG 1.4 Map of the south Balkan region

Slavs who settled in the Balkan Peninsula from north of the Carpathian range around AD 550 to 630 and linguistic evidence indicates that Ancient Macedonian was in fact separate from what later became Greek. At the time of the Slavic invasions, the territory of Macedonia was part of the Byzantine Empire. Thereafter, it shifted between Greek and Slavic domination until the end of the fifteenth century when it became part of the Ottoman Empire. It then passed from Ottoman to Serbian control as a result of the Balkan wars of 1912 and 1913 until it became a republic of Yugoslavia in 1944 under the Communist regime.

Macedonian is part of the South Slavic dialect continuum at the ends of which are Serbian and Bulgarian, both of which served at one time as languages of education for many Macedonians. Longstanding contact and multilingualism in the Balkan region between Slavic languages such as Bulgarian, Macedonian, and Serbian and non-Slavic languages such as Greek, Turkish, and Albanian have led to linguistic convergence of the type found in north-west New Britain. Turkish, for instance, contributed a significant amount of vocabulary to Macedonian, because it was the language of administration in Macedonia from the fourteenth century to the beginning of the twentieth, as well as a language of considerable economic and cultural prestige. Macedonian shares some similarities with Serbian and some with Bulgarian. The variety of Macedonian spoken in the capital, Skopje, for example, shares some of its most salient features with Serbian. This allows some Serbian authors to cite certain linguistic similarities between Serbian and Macedonian to bolster their territorial claims, at the same time as Bulgarians can use other linguistic features shared between Bulgarian and Macedonian to support their view that Macedonian is a dialect of Bulgarian. Not surprisingly, given the territorial ambitions of their Greek, Bulgarian, and Serbian neighbors, many Macedonians perceive themselves as linguistically and culturally threatened even in their own country.

English: language and dialect

So the answer to one of the questions I am often asked about whether American English is a dialect of English or a separate language depends on your point of view. George Bernard Shaw characterized England and America as two nations divided by a common language. When H. L. Mencken decided to call his book *The American Language* rather than *The English Language in America*, he was making a political statement. Similar sentiments had been expressed earlier by Noah Webster when he authored his dictionary of American