

Mikael Parkvall

Limits OF Language

almost everything
you didn't know
you didn't know
about language
and languages

WILLIAM JAMES & CO.



8536 SW St. Helens Drive, Ste. D · Wilsonville, Oregon 97070 · USA · www.wmjasco.com



United Kingdom and Sri Lanka · www.battlebridge.com · pb@battlebridge.com

Publisher—US edition Jim Leisy
Editor and Publisher—UK edition Philip Baker
Production Editor—US edition Tom Sumner

Printed in the U.S.A.

© 2008 Battlebridge Publications; William, James & Co.

This U.S. edition is published by William, James & Company in association with Battlebridge Publications. All rights are reserved, including the right of reproduction in whole or in part in any form.

For information, contact William, James & Company:

Rights and Permissions
William, James & Co.
8536 SW St. Helens Drive, Ste. D
Wilsonville, Oregon 97070

Cataloging-in-Publication information is available from William, James & Co.

First published in the United Kingdom in 2006.

isbn (paperback): 978-1-59028-198-7

isbn (hardback): 978-1-59028-210-6

BRIEF CONTENTS

LANGUAGE IN SOCIETY	
LANGUAGE VARIATION AND REGISTERS FOR SPECIAL OCCASIONS	1
LANGUAGE GEOGRAPHY	7
LANGUAGE AS A LEGAL MATTER	17
LANGUAGE PLANNING	34
USING MORE THAN ONE LANGUAGE	38
LANGUAGES IN CONTACT	50
LANGUAGE DEATH AND REVIVAL	57
BIG AND SMALL LANGUAGES	
BIG AND SMALL LANGUAGES	65
LANGUAGE CHANGE	
LANGUAGE CHANGE	88
LANGUAGE FAMILIES	99
SPECTACULAR SPREAD	105
LANGUAGE IN ALTERNATE HISTORY	111
WRITTEN LANGUAGE	
WRITTEN LANGUAGE	114
SIGNED LANGUAGES	
SIGNED LANGUAGES	125
LANGUAGE LEARNING	
LEARNING YOUR FIRST LANGUAGE	133
SECOND-LANGUAGE LEARNING	136
POLYGLOTS	143
NON-HUMANS AND LANGUAGE	
NON-HUMANS AND LANGUAGE	148
ARTIFICIAL LANGUAGES	
ARTIFICIAL LANGUAGES	157
FORENSIC LINGUISTICS	
FORENSIC LINGUISTICS	165
POPULAR CONCEPTIONS	
THE SUBJECTIVE APPROACH	170
LANGUAGE MYTHS	174
LANGUAGE AND IDENTITY	
PRAGMATICS	180
LANGUAGE AND GENDER	185
LEXICON	
HOW MANY WORDS ARE THERE?	189
LONG AND SHORT WORDS	192
HAVING A WORD FOR IT	195
DICTIONARIES	201
COMMON WORDS	203
COINING NEW WORDS	208
COLOR TERMS	214
PERSONAL NAMES	217
PLACE-NAMES	223
GLOSSONYMS	234
SPEECH SOUNDS	
SPEECH SOUNDS	241
CONSONANTS	244
VOWELS AND DIPHTHONGS	249
SYLLABLE STRUCTURE	254
SUPRASEGMENTALS	258
ARCHITECTURE OF LANGUAGE	
PARTS OF SPEECH	265
VERBS	266
NOUNS	268
ADJECTIVES	270
ADVERBS	271
ADPOSITIONS	273
ARTICLES	275
DEMONSTRATIVES	277
PERSONAL PRONOUNS	280
NUMERALS	288
NEGATION	296
CASE	298
GENDER AND NOUN CLASSES	301
NUMBER	305
TENSE, MOOD AND ASPECT	309
VOICE	316
MORPHOLOGY	318
SUPPLETION	322
SENTENCES	324
LINGUISTS AND THEIR DISCIPLINE	
LINGUISTICS DEPARTMENTS	328
MOST STUDIED LANGUAGES	332
EXPERIMENTATION IN LINGUISTICS	333
SOME WORK THAT DIDN'T STAND THE TEST OF TIME	338
LINGUISTS	347
LINGUISTIC TERMINOLOGY	361
CLASSIC EXAMPLE SENTENCES	366
THE LINGUIST'S GUIDE TO THE GALAXY	378

DETAILED CONTENTS

FOREWORD	XII
----------------	-----

LANGUAGE VARIATION AND REGISTERS FOR SPECIAL OCCASIONS—PAGE 1

MOST DIALECTALLY HOMOGENOUS LANGUAGES	1	OTHER UNUSUAL REGISTERS.....	4
MAKING FUN OF THE HANDICAPPED	3	MOST ELABORATE STATUS DIFFERENTIATIONS	4
MOST RECHERCHÉ REGISTER.....	3	AMERICAN INDIAN CODE TALK	4

LANGUAGE GEOGRAPHY—PAGE 7

MOST LINGUISTICALLY DENSE COUNTRIES	7	ODDEST GEOGRAPHICAL DISTRIBUTION.....	12
LEAST LINGUISTICALLY DENSE COUNTRIES	7	ON THE WRONG SIDE OF THE BORDER	12
MOST LINGUISTICALLY DIVERSE COUNTRIES	8	MAN ODD GEOLINGUISTIC SITUATION	13
CONTINENT WITH THE FEWEST LANGUAGES.....	9	MAKING THE MOST OUT OF A SINGLE ISOGLOSS	14
MOST ISOLATED LANGUAGES	10	DENSEST ISOGLOSSES	14
MOST WIDESPREAD LANGUAGES	10	VOLUNTARY LINGUISTIC SEGREGATION IN SOUTH DAKOTA.....	15

LANGUAGE AS A LEGAL MATTER—PAGE 17

THE WORLD'S MOST DANGEROUS LANGUAGE	17	DISCREPANCY BETWEEN NUMERIC AND OFFICIAL STATUS.....	25
LANGUAGE CONFLICT INVOLVING ONLY ONE ETHNIC GROUP	17	AN UNEXPECTED MAJORITY LANGUAGE	25
EARLIEST DOCUMENTED LANGUAGE-BASED DISCRIMINATION.....	18	UNEXPECTED OFFICIAL LANGUAGES.....	27
THE INFAMOUS LANGUAGE LEGISLATION IN FRANCE	18	GASTRONOMIC LANGUAGE PURISM	29
MOST REPRESSIVE MINORITY LANGUAGE POLICY	18	RETROACTIVE ASSIMILATION	30
LINGUISTIC RIGHTS IN THE "CRADLE OF WESTERN DEMOCRACY"	20	FROM DIALECT TO LANGUAGE OVERNIGHT	30
MOST GENEROUS MINORITY LANGUAGE POLICY.....	20	AN INDIAN TOWER OF BABEL.....	30
SMALL MAJORITY LANGUAGES	22	FIRST KNOWN LANGUAGE OF INTERNATIONAL DIPLOMACY	30
SMALL OFFICIAL LANGUAGES	22	MOST LINGUISTIC CONSTITUTION.....	31
MOST OFFICIAL LANGUAGE.....	23	OWNING A LANGUAGE	31
HIGHEST NUMBER OF OFFICIAL LANGUAGES	24	LANGUAGE LEGISLATION IN THE UNITED STATES.....	31

LANGUAGE PLANNING—PAGE 34

PLANNED KOINÉIZATION.....	34	A HEROINE FOR BREAKFAST	35
PLANNED PIDGINIZATION	34	LINGUISTIC ENGINEERING IN MELANESIA	36
PURISM IN REVERSE.....	35	NOUN CLASS PRESCRIPTIVISM	36

USING MORE THAN ONE LANGUAGE—PAGE 38

MOST MULTILINGUAL CITIZENS	39	MOST BILINGUAL PUBLISHING	46
GROUPS	41	BIGGEST SHARE OF FOREIGN-LANGUAGE PUBLISHING	46
DO WE REALLY LIVE ON THE SAME PLANET?	42	MOST TRANSLATED LANGUAGES	46
WHERE FAMILIES ARE BILINGUAL BY DEFINITION	42	MOST TRANSLATED TEXTS	47
MOST AND LEAST MULTILINGUAL CITIES.....	43	LARGEST TRANSLATION SERVICE	47
MONOLINGUAL COUNTRIES.....	45	FIRST MACHINE TRANSLATION	48
CORPORATE MONOLINGUALISM	45	FIRST SIGN LANGUAGE MACHINE TRANSLATION.....	48

LANGUAGES IN CONTACT—PAGE 50

MOST LOANWORDS	50	LARGE NUMBER OF BORROWED PHONEMES	52
MOST FAMOUS SPRACHBUND	51	STARTING AT THE WRONG END.....	53
MOST INTERNATIONAL WORDS.....	51	MOST STABLE DIGLOSSIA	53
AN UNBORROWED BODY-PART?	52		

LANGUAGE DEATH AND REVIVAL—PAGE 57

BIGGEST MORIBUND LANGUAGES	57	SLOW LANGUAGE REPLACEMENT.....	63
THE PACMANS OF LANGUAGES.....	58	AN UNUSUAL CASE OF SHIFTING LANGUAGE LOYALTY.....	63
RAPID LANGUAGE SHIFT	59	LATIN IS DEAD—BUT THE KING LIVES!	63
RAPID LANGUAGE DEATH	61	SUCCESSFUL LANGUAGE REVIVAL	63

BIG AND SMALL LANGUAGES—PAGE 65

SMALL LANGUAGES.....	66	OFFICIAL LANGUAGES OF INTERNATIONAL ORGANIZATIONS	76
BIG LANGUAGES.....	67	LANGUAGE SIZE IN THE PAST.....	78
“RICH” LANGUAGES	70	FUTURE SIZE OF LANGUAGES	79
GLOBAL INFLUENCE	71	GLOBAL ESTIMATE FOR 2050:.....	80
LANGUAGES IN SCIENCE	72	LANGUAGE-BASED INTERNATIONAL ORGANIZATIONS	81
MEDIA REPRESENTATION: LANGUAGES IN THE PRESS	72	GROSS OVERESTIMATIONS OF THE OWN LANGUAGE’S SIZE	82
LANGUAGES IN BOOKS	73	WORLD LANGUAGES.....	83
LANGUAGES IN FILMS	74	PERCEIVING ENGLISH AS THE ONE	
LANGUAGES ON THE RADIO	75	AND ONLY FOREIGN LANGUAGE	85
LANGUAGES ON THE WEB	75	WORLD LANGUAGE OR LOCAL LINGO?.....	85

LANGUAGE CHANGE—PAGE 88

YOUNGEST LANGUAGES.....	88	SOME INDIVIDUALS WHO HAVE HAD DISPROPORTIONATE	
DEVELOPING INTO		INFLUENCE ON LANGUAGE DEVELOPMENT.....	93
YOUR OWN OPPOSITE	89	JOHANNES AAVIK	93
BIZARRE SOUND CHANGE	90	BROR REXED.....	94
MOST STABLE		ADOLF HITLER	94
VOCABULARY ITEMS	90	EDWARD YOUNG AND JOHN ADAMS.....	95
NAUGHTIEST GRAMMATICALIZATION	90	GOD AND ADAM	96
RAPID LANGUAGE CHANGE.....	90	LUDWIG ZAMENHOF	96
CONSERVATIVE LANGUAGES.....	91	JOHN II OF PORTUGAL.....	96
TELLY-INDUCED LANGUAGE CHANGE	91	MUHAMMAD	97
REDUCTION	92	ELIEZER BEN-YEHUDA.....	98

LANGUAGE FAMILIES—PAGE 99

BIGGEST LANGUAGE FAMILY	99	SIMILAR AND DISSIMILAR LANGUAGES	101
OLDEST LANGUAGE FAMILY	99	MEASURING DISTANCE IN DAYS.....	102
BIGGEST ISOLATE	100	MOST AND LEAST REPRESENTATIVE LANGUAGES.....	102
LANGUAGE WITH THE MOST PROPOSED RELATIVES.....	100	BEFORE SIR WILLIAM	103

SPECTACULAR SPREAD—PAGE 105

ENGLISH	105	INTERETHNIC LANGUAGES OF	
OTHER IMPERIAL LANGUAGES	105	PAPUA NEW GUINEA.....	107
RENNELLESE SIGN	106	KRIO.....	108
ESPERANTO.....	106	FRENCH SIGN LANGUAGE.....	108
AFRIKAANS.....	106	SPECTACULAR SPREAD OF A SINGLE LINGUISTIC FEATURE.....	110

LANGUAGE IN ALTERNATE HISTORY—PAGE 111

IT JUST MIGHT HAVE HAPPENED.	112
--------------------------------------	-----

WRITTEN LANGUAGE—PAGE 114

HIGHEST AND LOWEST LITERACY RATES	114	LITERACY—GENDER DIFFERENCES	115
MOST POPULAR WRITING SYSTEMS.....	115	MOST COMPLEX WRITING.....	115

THE UNIQUE FOURTH TYPE	116	THE IMPORTANCE OF	
A LOST WRITING SYSTEM	116	ALPHABETICAL ORDERING.....	120
MOST AND FEWEST CHARACTERS	117	MISINTERPRETING EGYPTIAN HIEROGLYPHS	120
MOST ELABORATE CHARACTERS	117	MAKING THE MOST OF WHAT ONE HAS	121
WORST SPELLING	117	CHROMATOGRAPHIC WRITING.....	122
FREQUENT ALPHABET CHANGES	118	SPEED READING	123
OLDEST WRITING SYSTEM.....	119	SPEED WRITING	123

SIGNED LANGUAGES—PAGE 125

SIGN LANGUAGES OF THE WORLD	125	SIGN SINGING	128
THE MOST COMMON SIGN LANGUAGE MYTHS	125	WHEN YOU CAN'T HEAR VOICES IN YOUR HEAD	129
THE MOST WIDELY USED SIGN LANGUAGE.....	127	EARLY SIGN RECORDINGS.....	129
MOST PROLIFEROUS SIGN LANGUAGE.....	127	SOME INTERESTING SPECIFIC SIGNS	129
SIGN LANGUAGE AS A MEDIUM		A REMARKABLE TACTILE SIGNER.....	130
OF HIGHER EDUCATION.....	128	FIRST SCIENTIFIC STUDY OF TACTILE SIGN	131
LOCALLY WIDESPREAD SIGNING COMPETENCE	128	SHE TOOK THE TRUTH INTO HER OWN HANDS	131

LEARNING YOUR FIRST LANGUAGE—PAGE 133

OLDEST FIRST LANGUAGE LEARNER	133	EARLIEST ADPOSITIONS.....	134
BEST STUDIED FERAL CHILD.....	134	LANGUAGES WITHOUT MOTHERESE.....	134

SECOND-LANGUAGE LEARNING—PAGE 136

LARGEST FOREIGN LANGUAGE INSTITUTE.....	136	MOST POPULAR FOREIGN LANGUAGES	137
WORLD'S BIGGEST LANGUAGE LESSON	137	LESS POPULAR FOREIGN LANGUAGES.....	138
THE SECRET BEHIND		MOST AVID LANGUAGE-LEARNERS IN EUROPE.....	138
SPEAKING ENGLISH.....	137	¿HABLA USTED PHRASE-BOOKISH?	139

POLYGLOTS—PAGE 143

FEMALE POLYGLOTS.....	145	POLYGLOT LINGUISTS.....	146
-----------------------	-----	-------------------------	-----

NON-HUMANS AND LANGUAGE—PAGE 148

LARGEST VOCABULARY ACQUIRED BY AN APE.....	148	TALKATIVE TREES.....	153
A CHAT WITH KOKO.....	148	CONVERTING WOOFs INTO WORDS.....	153
BEST KNOWN SUBJECTS OF APE LANGUAGE RESEARCH	149	TALKING TO ALIENS	154
BEST COMMUNICATIVE PERFORMANCE BY AN APE	151	FIRST CHATTERBOT	155
LARGEST DOG VOCABULARY.....	151	FIRST BOOK OF COMPUTER-GENERATED POETRY	155
MOST HUMAN-LIKE COMMUNICATION (WELL, IN A WAY....)	151	FIRST ACCEPTED COMPUTER-	
BEST TALKING BIRD	152	GENERATED CONFERENCE TALK	155

ARTIFICIAL LANGUAGES—PAGE 157

MOST SUCCESSFUL ARTIFICIAL LANGUAGE	157	AGLOTTIC POP—TAKING NONSENSICAL	
AN ARTIFICIAL LANGUAGE GOING NATIVE	158	POP LYRICS TO THEIR LOGICAL EXTREME	160
A BIT OF GEEKOLOGY:		SOLRÉSOL—THE “MUSICAL LANGUAGE”	161
WHO ON EARTH SPEAKS KLINGON?.....	159	MOST ECCENTRIC ESPERANTO CLONE	162
BRITISH ROMANCE	159	MOST RHYMING CONLANG	163

FORENSIC LINGUISTICS—PAGE 165

IDENTIFYING A RECLUSE BILLIONAIRE	165	USING LINGUISTICS TO	
THE MYSTERY OF THE PERSIAN MUMMY.....	165	FRAME THE UNABOMBER	168
THE HELANDER DEBÂCLE	166	INCLUSIVITY IN COURT	168
POSTHUMOUS PARDON THANKS TO A LINGUIST	167	INCLUSIVITY	168

THE SUBJECTIVE APPROACH—PAGE 170

AESTHETICS	171
----------------------	-----

LANGUAGE MYTHS—PAGE 174

THE LEGENDARY SNOW HOAX	174	THE SECRET MEANINGS OF COCA COLA	176
THE GERMAN VOTE	175	THE CHAMBER OF HORRORS: ASSORTED ALL-TIME CLASSICS	177
IF IT WAS GOOD ENOUGH FOR JESUS	175		

PRAGMATICS—PAGE 180

VERBAL INTERACTION IN ABORIGINAL AUSTRALIA	180	GREETING IN MANINKA	182
TRUE OR FALSE—OR SOMETHING IN BETWEEN?	181	POLITENESS IN KOREAN	183
CHAOTIC CARIBBEAN CONVERSATIONS	181	A PRAGMATIC UNIVERSAL?	183
BEING POLITE IN WOLOF	182	BAD TURNTAKING SKILLS	184
		USE OF PERSONAL NAMES	184

LANGUAGE AND GENDER—PAGE 185

LANGUAGES WITH THE GREATEST GENDER DIFFERENCES	185	AN EXCLUSIVELY FEMININE SCRIPT	187
NON-STANDARD FEMALE SPEECH	186	THE QUEST FOR A GENDER-NEUTRAL PRONOUN	187
		GENDERED COMPUTERS	188

HOW MANY WORDS ARE THERE?—PAGE 189

SOME REASONS WHY WE CANNOT ESTIMATE THE NUMBER OF WORDS IN A LANGUAGE	189
---	-----

LONG AND SHORT WORDS—PAGE 192

SHORTEST WORDS	192	YES, THERE ARE LONGER WORDS THAN “ANTIDISESTABLISHMENTARIANISM”	193
LANGUAGES WITH EXCEPTIONALLY SHORT WORDS	192		

HAVING A WORD FOR IT—PAGE 195

EXPRESSING AN ODD FOOD PREFERENCE	195	KOREAN KIN AND OTHER RELATIVES	197
MACROPOD LOCOMOTION IN ARNHEM LAND	196	MISCELLANEOUS GOODIES	197
MOST DIFFICULT-TO-TRANSLATE WORD?	196	DEBUNKING A POPULAR MYTH	198
SOUTH SEA CLOUDS	196	IMPROVING YOUR LANGUAGE BY MEANS OF LIFFS	199

DICTIONARIES—PAGE 201

MOST PROCRASTINATED DICTIONARY PUBLICATION	201	AN ONOMATOPOEIC DICTIONARY	202
WORST DICTIONARY	201	DIRTY DICTIONARIES	202

COMMON WORDS—PAGE 203

MOST FREQUENT WORDS IN ENGLISH	203	JANUS WORDS	205
MOST FREQUENT WORDS IN OTHER LANGUAGES	204	SOME ENGLISH JANUS WORDS	205
FREQUENCIES IN SPECIFIC GENRES	204	PSEUDO-LOANS	207

COINING NEW WORDS—PAGE 208

A NEW CONJUNCTION	208	COINING A WORD BY MISTAKE	209
COINING WORDS FOR MONEY	208	THE TEN SILLIEST NAMES OF MOLECULES (IN NO PARTICULAR ORDER)	210
CORPORATE PRESCRIPTIVISM	208	BIOLOGICAL NAMES	211

COLOR TERMS—PAGE 214

LARGEST NUMBER OF COLOR TERMS	214	UNUSUAL AND	
SMALLEST NUMBER OF COLOR TERMS.....	215	PROBLEMATIC CASES	215

PERSONAL NAMES—PAGE 217

TOP TEN SURNAMES IN SELECTED COUNTRIES.....	217	MOST COMMON FIRST NAME.....	220
TOP TEN SURNAMES IN SOME COUNTRIES.....	218	EUROPEAN NAMES IN AFRICA.....	221
DIVERSITY OF SURNAMES IN CERTAIN TERRITORIES	220	CHANGING POLITICS, CHANGING NAMES	221
AN UNEXPECTED FLIPINO NAME	220	WHERE THE PEOPLE HAVE NO NAMES	222

PLACE-NAMES—PAGE 223

NAMES OF COUNTRIES	223	LUNAR TOPONYMY	229
A NEW TOPONYMIC TRADITION?	224	OLD NAMES FOR A NEW WORLD.....	230
ALEXANDER THE PROLIFEROUS AND HIS SUCCESSORS	226	MOST COMMON PLACE-NAMES	231
RADICAL RENAMING BEFORE LENIN	228	LONG AND SHORT PLACE-NAMES	232

GLOSSONYMS—PAGE 234

SHORTEST GLOSSONYMS	234	PARLEZ-VOUS BUCKET?.....	235
LONGEST GLOSSONYMS	234	THEY SPEAK WHAT?!	235
GLOSSONYMS WITH VOWELS ONLY	235	ORIGINS OF LANGUAGE NAMES.....	236
DIFFICULT-TO-PRONOUNCE		USELESS BONUS INFORMATION 1.....	239
LANGUAGE NAMES.....	235	USELESS BONUS INFORMATION 2.....	239

SPEECH SOUNDS—PAGE 241

LARGEST PHONEMIC INVENTORY	242
SMALLEST PHONEMIC INVENTORY	242
MOST COMPLEX PHONEME	242

CONSONANTS—PAGE 244

CONSONANTS	244	ODD TRILL	
PLOSIVES.....	244	REALIZATION	246
FRICATIVES.....	245	BILABIALS	246
AFFRICATES	245	NASALS	247
CLICKS	245	PHARYNGEALS	247
LIQUIDS	246	AMAZONIAN ODDITIES	247

VOWELS AND DIPHTHONGS—PAGE 249

VOWELS	249	MOST DIFFICULT-TO-DETERMINE	
NASAL VOWELS.....	249	VOWEL SYSTEM	251
DIPHTHONGS	250	FRONT AND BACK VOWELS	251
VOICELESS VOWELS.....	250	THREE DEGREES OF LIP ROUNDING.....	252
LANGUAGES WITHOUT VOWELS?	251	RETROFLEXED VOWELS	252

SYLLABLE STRUCTURE—PAGE 254

SIMPLE PHONOTACTICS	254	VOWEL-LESS WORDS	255
COMPLEX PHONOTACTICS	255	PREFERRING THE UNUSUAL	256

SUPRASEGMENTALS—PAGE 258

LENGTH	258	BOTH TONE AND STRESS	259
TONE.....	259	SINGING IN TONES.....	260

WHISPERING AND SHOUTING SUPRASEGMENTALS.....	260	ASPIRATION IN PALULA.....	261
SUBTLE SUPRASEGMENTAL DIFFERENCES.....	261	NASALIZATION.....	262
SENTENCE INTONATION.....	261	MOST HIGH-PITCHED LANGAUGE.....	262
		LOUDEST LANGUAGE.....	263

PARTS OF SPEECH—PAGE 265

LANGUAGES WITH NO PARTS OF SPEECH AT ALL?	265
SOURCES NOT MENTIONED IN THE TEXT	265

VERBS—PAGE 266

FEW VERBS	266
EXTENSIVE VERBAL SERIALIZATION	266
UNUSUALLY VERBY LANGAUGES	267

NOUNS—PAGE 268

NOUNS IN DIFFERENT LANGUAGES AND GENRES	268
MOTHERING AND FATHERING	269

ADJECTIVES—PAGE 270

LANGUAGES WITH FEW ADJECTIVES	270
-------------------------------------	-----

ADVERBS—PAGE 271

FEWEST ADVERBS	271	SIMPLE DEICTIC SYSTEM.....	271
ADVERBS BUT NO ADJECTIVES	271	HOW DO YOU FEEL?	272
INTERROGATIVE PRO-VERBS	271	TELLING TIME IN PIRAHÁ	272

ADPOSITIONS—PAGE 273

FEWEST ADPOSITIONS	273
FLEXIBLE ADPOSITIONS.	273
PREPOSITION STRANDING	274

ARTICLES—PAGE 275

TENSED ARTICLES	275
THREE DEGREES OF DEFINITENESS	275
COMPLEX ARTICLE SYSTEMS.	276

DEMONSTRATIVES—PAGE 277

HIGHEST NUMBER OF DEGREES OF DEIXIS.....	277	NO DEMONSTRATIVES AT ALL?	278
ELABORATE DEMONSTRATIVE SYSTEMS.....	277	AN UNUSUAL DICHOTOMY.....	278

PERSONAL PRONOUNS—PAGE 280

COMPLEX PERSONAL PRONOUN SYSTEMS.....	280	LONG PRONOUN.....	284
SOMEWHAT SIMPLER PRONOUN SETS.....	282	LANGUAGES NOT USING PERSONAL PRONOUNS	284
FOURTH PERSON.....	282	FURTHER ELABORATIONS	
SEX CONTRASTS.....	283	OF THE PRONOUN SET	285
ANALYTIC AND PORTMANTEAU PRONOUNS.....	283	GOD'S OWN PRONOUN	286

NUMERALS—PAGE 288

NO CATEGORY OF NUMERALS	288	SEVERAL PARALLEL NUMERAL SYSTEMS.....	289
MORE ORDINALS THAN CARDINALS	288	SMALLEST NUMERAL SYSTEM	289

x DETAILED CONTENTS

LARGEST NUMERAL SYSTEM	290	MORE ON FINGER COUNTING	292
UNUSUAL BASES	290	LONGEST NUMERALS	293
BASE 64 AS A ROYAL DECREE.....	291	SHORTEST NUMERALS	294
INTERESTING ORIGINS OF NUMERALS	291	DISCONTINUOUS NUMERALS	294

NEGATION—PAGE 296

WHERE LESS IS MORE.	296
LEAVING NEGATION TO CONTEXT	296
MULTIPLE NEGATION	297

CASE—PAGE 298

TONAL CASE.....	298	UNUSUAL CASE.....	299
CASES ON CASES.....	298	AN ODD MARKEDNESS PREFERENCE	299
HIGHEST NUMBER OF CASES	299	SPLIT ERGATIVITY THE HARD WAY	299

GENDER AND NOUN CLASSES—PAGE 301

HIGHEST NUMBER OF NOUN CLASSES	301	UNEXPECTED GENDER AGREEMENT	303
SOME FAVORITE NOUN CLASSES.....	302	ETHNOCENTRIC GENDER ASSIGNMENT.....	303
MOST FAMOUS NOUN CLASS.....	302	CLASSIFYING HANDLING	
DO-IT-YOURSELF CLASSIFIERS.....	302	ACTIONS IN NORTH AMERICA.....	303
DOUBLE GENDER SYSTEMS.....	303	GENDER IN UNEXPECTED PLACES.....	304

NUMBER—PAGE 305

COMPLEX NUMBER SYSTEM	305	NO NOMINAL NUMBER.....	306
NUMBERS ON NUMBERS	305	NO NUMBER AT ALL.....	307
INVERSIVE NUMBER MARKING	305	UNEXPECTED	
GREATEST NUMBER OF NUMBERS	306	NUMBER AGREEMENT	307
AMBAL	306	VERBAL NUMBER	307
EVEN MORE UNUSUAL NUMBERS.....	306	PLURAL AS THE DEFAULT NUMBER	307

TENSE, MOOD AND ASPECT—PAGE 309

TENSED NOUNS.....	309	COMPLETE TENSELESSNESS	311
TMA (TENSE, MODALITY, AND ASPECT MARKERS) ON PRONOUNS.....	310	ELABORATE ASPECT.....	311
TENSED DETERMINERS.....	310	RAREST ASPECT?	311
TENSED PREPOSITIONS.....	310	EVIDENTIALS GALORE.....	312
TENSED IMPERATIVES	310	PLENTY OF TENSES	313
TONAL TMA	311	PLURALITY OF MODALITY	313
		LANGUAGES WITH COMPLEX TMA SYSTEMS IN GENERAL.....	314

VOICE—PAGE 316

ODD VOICE FORM	316
LARGE NUMBER OF VOICES	316
VOICE AGREEMENT	317

MORPHOLOGY—PAGE 318

COMPLEX AND NOT SO COMPLEX MORPHOLOGY	318	PERSON MARKING ON NOUNS.....	320
ANOTHER WAY OF MEASURING THINGS	319	PERSON MARKING ON NOUNS	320
MOST DADAISTIC VERB MORPHOLOGY	319	METATHESIS AS A MORPHOLOGICAL DEVICE.....	320

SUPPLETION—PAGE 322

MOST VERBAL SUPPLETION	322
SOME UNUSUAL SUPPLETIVES	322

SENTENCES—PAGE 324

MOST COMMON AND		LONG SENTENCES	325
LEAST COMMON BASIC WORD ORDER	324	SHORT SENTENCES	326
FREE WORD ORDER AGAINST ALL ODDS	325	A LANGUAGE WITHOUT EMBEDDING?	326

LINGUISTICS DEPARTMENTS—PAGE 328

BIGGEST DEPARTMENT	328	NORTHERNMOST AND SOUTHERNMOST DEPARTMENT	329
SMALLEST DEPARTMENTS	328	MOST RENOWNED FACULTY	330
LARGEST NUMBER OF		OLDEST LINGUISTICS DEPARTMENT	330
LINGUISTICS DEPARTMENTS	328	LARGEST LINGUISTIC ASSOCIATION	330

MOST STUDIED LANGUAGES—PAGE 332

BEST DESCRIBED LANGUAGES.	332
-----------------------------------	-----

EXPERIMENTATION IN LINGUISTICS—PAGE 333

TESTING INNATENESS BEFORE CHOMSKY	333	CAN THE IMPOSSIBLE BE LEARNED?	334
EARLY SPEECH SYNTHESSES	333	THE MONSTER STUDY	335
BEING MEAN TO LITTLE CHILDREN		BLUSHING FOR NO APPARENT REASON	335
AT THE WORKPLACE	334	THE ULTIMATE TEST: MIDWIFING A NEW LANGUAGE	336

SOME WORK THAT DIDN'T STAND THE TEST OF TIME—PAGE 338

AND THEY HUFFED AND THEY PUFFED	338	TUNE IN, TURN ON, BEGIN TO SPEAK	340
A FOGGY THEORY	339	PLAGIARISM IN SOCIOLINGUISTICS	341
SPEECH FROM A RAY OF LIGHT	339	IN SEARCH OF THE PROTO-LANGUAGE	341
DAVID OATES' SPEECH REVERSAL LAB	339	THE BEAUTY AND LOGIC OF THE FRENCH LANGUAGE	344
MAKE LOVE, NOT WAR—BY AVOIDING VOICED FRICATIVES!	340	ENGLISH—A DEADLY LANGUAGE	345
EXAMINE YOUR LIMBS AND FIND OUT IF YOU CAN SPEAK	340	MARRISM	345

LINGUISTS—PAGE 347

MOST CITED LINGUISTS	347	YOUNGEST LINGUIST	356
TOP 20 (WITH A CONSERVATIVE DEFINITION OF "LINGUISTICS")	347	FIRST LINGUISTS	356
LINGUISTS IN POLITICS	348	LINGUISTS WITH LUXURIANT FLOWING HAIR	356
PART-TIME LINGUISTS	349	LANGUAGES SPOKEN BY LINGUISTS	356
LINGUISTS WHO HAVE SPENT TIME IN PRISON	350	MEN AND WOMEN IN LINGUISTICS	357
SINISTER LINGUISTS	352	NAMES OF LINGUISTIC WORKS	357
LINGUISTS' TIME OFF	352	BEST APOLOGETIC ENDNOTE	358
REVIEWING YOUR OWN WORK	353	BEST AMBIGUITY	358
DEAD LINGUISTS SOCIETY	353	PRESTIGE IN LINGUISTICS	358
SOME LINGUISTS WHO LIVED QUITE A BIT LONGER	355	HOW TO RECOGNIZE A REAL LINGUIST	359

LINGUISTIC TERMINOLOGY—PAGE 361

(LIST OF LINGUISTIC VOCABULARY TERMS)	361
---	-----

CLASSIC EXAMPLE SENTENCES—PAGE 366

CITATIONS CONCERNING THE GENERAL NATURE OF LANGUAGE366	HISTORICAL LINGUISTICS373
GARDEN-PATH SENTENCES368	SOCIOLINGUISTICS AND PRAGMATICS374
SEMANTICS368	LANGUAGE ACQUISITION BY
THE STRUCTURE OF LANGUAGE IN GENERAL	INFANTS AND ANIMALS.....374
OR ENGLISH IN PARTICULAR369	MISCELLANEOUS375
THE STRUCTURE OF SPECIFIC LANGUAGES OTHER THAN ENGLISH...371	OTHER MEMORABLE EXAMPLES376

THE LINGUIST’S GUIDE TO THE GALAXY—PAGE 378

LINGUISTIC FURNITURE.....378	LINGUISTIC NAMING PRACTICES383
LINGUIST COOKING378	FOR THE LINGUIST TOURIST383
IN THE LINGUIST’S RECORD COLLECTION379	WHERE TO STAY.....385
LINGUISTIC MONEY381	AUTHORIZED TYPEFACES.....385
COLLECTING STAMPS LINGUISTICALLY381	DRESSING DEPARTMENT-SMART.....386
WHAT LINGUISTS VOTE FOR382	FURTHER TRAVEL TIPS: LINGUISTIC MONUMENTS386
A LINGUIST’S NIGHT OUT382	WHAT LINGUISTS READ.....390
LINGUISTS AND SPORTS.....382	LINGUISTICS IN FILMS.....392

THE LINGUIST’S CALENDAR.....	396
PICTURE CREDITS.....	426
REFERENCES.....	427
INDEX.....	449

FOREWORD

This book was born, as I recall, during a stroll along the Regent's Canal in London in the spring of 2001. My friend Philip Baker who had founded the Battlebridge publishing company had asked me about ideas concerning a book that might appeal to a wider audience than the titles they had thus far produced. I was myself very happy with my idea, and had expected Philip to excitedly start jumping up and down in his seat. When I carefully expressed my disappointment with what I perceived as lack of appropriate enthusiasm, I realized I had forgotten that he is British. "I am *wildly* enthusiastic", he calmly stated, in a tone you would only hear from a Briton.

During the following five years, *Limits of Language* grew at rather uneven intervals, but finally—here it is. It was originally conceived as a kind of *Guinness Book of World Records* for languages, but as languages do not easily lend themselves to quantifying, it would have been impossible to write a book consisting only of non-subjective records. Therefore, not only the *Guinness*, but also a few other books which formed my childhood view of the world served as models: *The Book of Lists* and the publications by the Diagram Group. Drawing on these sources of inspiration, I hope that *Limits of Language* can show the uninitiated some of the incredible aspects that linguistics and human languages have to offer, teach beginners some of the basics of linguistics, but also to serve as a reference book for experienced linguists—here, the linguist can identify the extremes, and thereby judge to what extent his or her own language is "normal".

This book contains data on almost 1000 language varieties, and a book with such a broad scope is bound to contain some errors. It would be impossible to actually master—or even have considerable experience of—more than a fraction of these languages. So, like any linguist dealing with more than a handful of languages, I have had to rely on descriptions and analyses made by others, the quality of which I am not always competent to assess. It would thus be surprising indeed if there were absolutely no inaccuracies to be found among these pages. While I have endeavoured to consult people with a solid competence in the languages concerned, if there are any cases where I have not been careful enough, I would very much appreciate having these drawn to my attention. In that way, any errors or shortcomings can be corrected in future editions of *Limits of Language*. Also, of course, readers of this book may know of additional material worthy of consideration for inclusion in subsequent editions, and I would be very grateful to receive such suggestions. So, for these reasons and for feedback of any kind, please contact me at <beyondlimits@battlebridge.info>. It would be much appreciated if you could also indicate the source of your information (or alternatively indicate your competence as a speaker of the language in question).

Some initial errors have in fact already been spotted and corrected by the helpful friends and colleagues who have gone through parts of the manuscript, or who have otherwise contributed significantly. These include Bjørn Bojesen, Östen Dahl, Eskil Dalenius, Niklas Edenmyr, Robert Eklund, Stéphane Goyette, Anthony Grant, Harald Hammarström, Bernd Heine, Samuel Härgestam, Tore Janson, and Jonas Söderström. In particular, Peter Bakker, and Daniel Bunčić have been involved in the making of this book to an extent just short of listing them as co-authors.

Olle Engstrand, Jussi Karlgren, Ulrika Kvist Darnell (who suggested the very title), Zita Lekeberg, Anna-Lena Nilsson, and Ljuba Veselinova have also been helpful in various ways, as has Päivi Juvonen, who was at an early stage intended to co-author the book. While this did not happen, she has remained the best of friends throughout the process. I am also grateful to all the following people who provided me with information or assisted in other ways: Inger Ahlgren, Keira Ballantyne, Hendrik Biebeler, Mickey Blake, Lars Borin, Ross Clark, Mark Donohue, James Emmett, Anders Eriksson, Anders G Eriksson, Bill Foley, Colette Grinevald, M J Hardman, Ellen Jernström, Niklas Jonsson, Axel Kristensen, Diana Krull, Mai Kuha, Henrik Liljegren, Eva Lindström, Mike MacMahon, Matti Miestamo, Bill Morris, John Ohala, Nicholas Ostler, Jens Östlund, Roz Prickett, David Quinto, Dónall ó Riagáin, Jørgen Rischel, Sarah Roberts, Bernhard Wälchli, Ashley Williams, Janet Wilson. In addition, Philip Baker has engaged himself in this enterprise to an extent way beyond what a publisher/editor normally does.

The Max-Planck-Institut für evolutionäre Anthropologie in Leipzig, Germany, had the kindness to invite me to spend several weeks there as a visiting researcher in the spring of 2006. As this happened to take place just a few weeks before the original publication of this book, some of my time there was spent on making final preparations for that. I am grateful to both the institute and its staff for the hospitality they showed me.

Let us not speculate about any possible causality, but two great linguists sadly passed away shortly after having provided comments on this book—Larry Trask and Peter Ladefoged. I had no idea how serious Larry’s condition was when I sent him a draft copy, but realized afterwards that he was in fact reading the manuscript on his deathbed. Having long been impressed by his works, I was nevertheless amazed by the exceptional detail of his comments. His last message simply finished with the line “I’ve been in poor health recently, and reading this stuff has cheered me up no end”. I loved hearing that, of course, but those words took on a new meaning when he died not long afterwards. For his impressive contributions to linguistics, and for his helpfulness in the making of this book, I would like to dedicate this book to the memory of Larry Trask. Peace be with him, and may he enjoy himself whatever he may be doing these days.

A NOTE ON SOME CONVENTIONS ADOPTED IN THIS BOOK

Although written in English, this book is international in using metric units throughout. No bushels, long tons, furlongs, fluid ounces, fathoms, pounds per square inch, hogsheads, or stones’ throws in sight anywhere. For the same reason, it also uses the decimal comma instead of the decimal point, and gives dates in accordance with international standards, i.e. in the format YYYY-MM-DD.

Although published in the US, the punctuation follows the norms of UK English in that punctuation marks are placed outside of quotation marks when the meaning demands it.

For language names, an effort has been made to make these conform to the practices of the Ethnologue database (www.ethnologue.com), which should facilitate the search for additional information on these languages. In some cases, this was not possible, and in a few, a label thought to be more widely known was—for precisely that reason—preferred over the Ethnologue counterpart (e.g. “Cantonese” rather than “Chinese, Yue”).

For phonetic representation, this book uses the International Phonetic Alphabet (IPA). For information on how to interpret the IPA, go to <http://www.arts.gla.ac.uk/ipa/ipachart.html>. An easy-to-use interactive introduction to the IPA for only the sounds of English is here: <http://www.arts.gla.ac.uk/ipa/ipachart.html>.

LANGUAGE VARIATION AND REGISTERS FOR SPECIAL OCCASIONS

IN MOST, possibly even all societies, people speak differently on different occasions. In western cultures, for instance, we usually express ourselves differently while having a drink at the pub with our close friends than we would at an employment interview—we use a different *register*. In many societies, this register differentiation is far more institutionalized than in English. Among the most commonly cited examples are Aboriginal languages in Australia, many of which contain subsystems often referred to as “mother-in-law” languages or “avoidance styles”. These must be used in the presence of certain relatives (though not necessarily mothers-in-law). In some languages, the entire lexicon (except function words) is replaced, while keeping the grammatical structure intact. Very similar phenomena are also encountered elsewhere in the world, e.g. in some languages of southern Africa.

MOST DIALECTALLY HOMOGENOUS LANGUAGES

Deciding which language is the most dialectally homogenous all boils down to what you choose to label “a language”. If Indo-European is a language, then it sure is heterogeneous, but if what you and your buddies—but nobody else—speaks is a language, then it probably is pretty homogenous. Among more widespread areas where languages (as customarily defined) display remarkably minor differences with regard to geography, those of Russian, American English, Icelandic, Hakka, and Hausa may be mentioned. “Surely”, I can hear someone say, “you can hear the difference

in speech between a person from New York and one from Boston?” Well, although some Bostonians have more or less expected me to perceive differences in pronunciation between people living on opposite sides of the same street, the fact remains that there are hardly any areas of the world where you could travel a distance corresponding to that between, say, Miami and Seattle, or Boston and Los Angeles, without having more difficulties in making yourself understood than you would in the USA. Similarly, the differences between Moscow and St Petersburg, although (barely)

perceivable to native speakers, are virtually negligible when you consider that speakers of regional varieties of Bengali, Khanty, Italian, Fijian, and Japanese would have a hard time (to say the least) understanding each other unless they adapted to the standard. (And note that these languages are spoken in far smaller geographical areas than are Russian and American English).

In general, the relative lack of dialectal variation applies to recently colonized territories, and thus to all European languages in the New World and Oceania. Interestingly, however, the first ever description of the English of the Falkland Islands (Sudbury 2000:85-86) documents a certain dialectal

variation, with the speech varieties of Hill Cove (of West Falkland) and Darwin and Walker Creek (on southern East Falkland) being the most noticeable. There are no readily available population figures for these settlements, but given that all of West Falkland has a total population of 144 (excluding the far more numerous sheep and penguins), the Hill Cove variety would also seem to be a good candidate for one of the smallest dialects of English.

The recent settlement situation also applies to Israel, where the geographical differences among native speakers are small indeed. (The Hebrew of individuals may be influenced by the other languages they speak but this has little to do with geography.)

MAKING FUN OF THE HANDICAPPED

Like its neighbor Nootka, Quileute has a number of odd speech registers. When speaking to cross-eyed persons, for instance, every word must be prefixed by /tə-/. Similarly,

words directed to 'funny persons' must be made to begin in /tʰ-/. Prefixes used when addressing small-sized men, hunchbacks and lame persons are /s-/, /tʰʰ-/ and /tʰχ-/ respectively.

MOST RECHERCHÉ REGISTER

The Lardil people of Mornington island off Australia's north coast until recently had a particularly peculiar speech register, called *Damin*. It is remarkable in several ways, and it is difficult to avoid the impression that it was more or less consciously created precisely in order to be spectacular.

The vocabulary consisted only of a couple of hundred words, and the affixes were identical to those of ordinary Lardil. Its flamboyant character was manifested mainly in its phonology. First, it is the only language known outside of southern Africa to include clicks (see p 245), and it even had a couple of click sounds not attested in any other language in the world. Furthermore, it has an exceptional variety of phonation types,

including a pulmonic ingressive and velaric and glottalic egressives.

Other extravagant features of *Damin* include a pronoun inventory with a distinction only between first person singular and everything else (Lardil has 19 different personal pronoun forms). Similarly remarkable is the regular formation of opposites by means of a prefix, where a negative, rather than a positive adjective, is the base form (e.g. *tjitjuu* 'small' versus *kuritjituuu* 'large').

Damin was taught by older men in initiation rites, but ceased to be used during the 20th century. Now, Lardil itself, of which there are only a couple of dozen speakers left, is suffering the same fate.

OTHER UNUSUAL REGISTERS

It is reported that speakers of Kalam (New Guinea) must use a special register when collecting pandanus nuts. Unless they do so, the spirits inhabiting the nuts would realize that the nut is about to be picked, and would cause it to rot instantly.

From Africa, a “royal accent” has been reported among speakers of Akan. This is not unusual in itself, since nobility is often associated with specific speech styles, but the Akan variety stands out in including (among other things) conscious stuttering.

In the same part of the world, Itsekiri and Okpe have developed partially different vocabularies to be used depending on the time of the day. Thus, concepts such as ‘blood’,

‘fire’ and ‘firewood’ are labelled differently by day and by night.

In Buru (eastern Indonesia) taboos are related not primarily to activities or interlocutors, but particularly to geographical areas. In certain areas, specific words must not be used, and in others, particular subjects (e.g. hunger and thirst) must be avoided. The uninhabited interior of Buru island constitutes a special case. There, any language can be spoken, provided it is *not* Buru. Malay would do fine (as would Cockney, Basque or Klingon), but a specific language, called Li Garan, has also developed. Li Garan is only used by persons who for one reason or another need to visit the area concerned.

MOST ELABORATE STATUS DIFFERENTIATIONS

Remarkably complex systems of speech levels, where your speech makes far-reaching adjustments to the social standing of the person you’re talking to, are conspicuous in East Asia. One of the most ornate and oft-cited examples is that of Javanese, which sports a high (*krama*), a

neutral (*madya*), and a low (*ngoko*) speech level, with a further distinction into two sub-levels of each variety. As can be seen from the examples below, virtually all words are differentiated, including function words, numerals, and basic items such as ‘yes’ and ‘no’.

Highest level	<i>Menapa</i>	<i>nandalem</i>	<i>mundhut</i>	<i>sekul</i>	<i>semanten?</i>
	<i>Menapa</i>	<i>panjenengan</i>	<i>mendhet</i>	<i>sekul</i>	<i>semanten?</i>
	<i>Napa</i>	<i>sampéyan</i>	<i>mendhet</i>	<i>sekul</i>	<i>semonten?</i>
	<i>Napa</i>	<i>sampéyan</i>	<i>njupuk</i>	<i>sega</i>	<i>semono?</i>
	<i>Apa</i>	<i>sliramu</i>	<i>mundhut</i>	<i>sega</i>	<i>semono?</i>
Lowest level	<i>Apa</i>	<i>kowé</i>	<i>njupuk</i>	<i>sega</i>	<i>semono?</i>
Literal translation	<i>question marker</i>	<i>you</i>	<i>take</i>	<i>rice</i>	<i>that much</i>

‘Did you take that much rice?’

The differences between the various styles are primarily lexical, and almost every single word must be substituted when switching from one style to another. Some analysts, however, have also claimed to have identified prosodic and morphophonemic differences, although these remain poorly understood.

Korean features seven degrees of politeness for the 2nd person pronouns. In addition to that, Korean used to have a 'superpolite' level, which is now archaic.

The choice of appropriate pronoun forms is also a delicate matter in Sinhala. As well as five levels of politeness, there are special ones for addressing Buddhist monks. On top of this, regional variation makes the friendly/egalitarian form of one part of Sri Lanka vulgar or insulting in another. One teaching manual advises Sinhala-learners to avoid using 2nd person pronouns until they have lived in the country for at least six months!

AMERICAN INDIAN CODE TALK

Following a suggestion from a Philip Johnston, who had grown up in a Navajo reservation in Arizona, the American military in early 1942 hired a number of Navajo speakers for use as radio-telegraphists. Even if Japanese cryptographers could crack an ordinary code, chances were slim that they would be in a position to first crack the code, and then find somebody to translate the message from Navajo to Japanese.

The Navajo code-talkers were first used in battle on Guadalcanal in August of 1942. Not only was the code more difficult to crack for the enemy, but the Navajos were also more efficient, as they had the knowledge of Navajo in their heads rather than in a code book. In addition, even in the unlikely case that the Japanese should learn to understand Navajo, it would be impossible for them to imitate it with a native-speaker accent, and thus the risk of receiving fake messages was minimised.

The success was such that the code-talkers had their own body-guards, assigned to protect the valuable pool of competence.

In order to adapt the Navajo lexicon to the realities of modern warfare, words for various birds and fish were drafted in to serve as designations for aircraft and warships. Names for foreign nations were also coined,

and 'Australia' was *Cha-yes-desi*, that is 'rolled hat', while 'Britain' was *Toh-ta* or 'between waters', and France *Da-gha-hi*—'beard'. Other items for which Navajo lacked words were conveyed letter by letter in English, because Japanese overhearing the communication might get a clue as to the content of the message by understanding the English words. The spelling, however, was such that Navajo words were used for the initial letter of the *corresponding English words*. Thus, *dze* stood for the letter *e*, as the Navajo word means 'elk'. In order to avoid repetitions, which might facilitate things for enemy code-crackers, the most common letters had two or three Navajo counterparts. In this way, the name Guadalcanal, which contains four instances of a could be rendered as *klizzie, shi-da, wol-la-chee, lha-cha-eh, be-la-sana, dibeh-yazzie, moasi, tse-nill, tsah, tse-nill, ah-jad*, with only one repetition. The whole chunk translates into English as *goat, uncle, ant, dog, apple, lamb, cat, axe, needle, axe, leg*—in other words 'Guadalcanal'. The Japanese never did succeed in cracking the code.

Much of the code-talker activities were classified for a long time after the war, but they did get a late recognition as the *National Navajo Code Talkers Day* was proclaimed on



Photo of Navajo Code Talkers in formation at Camp Pendleton, California

August 14, 1982 by then American president Ronald Reagan (a local *Navajo Code Talkers Day* had been declared in New Mexico already on April 10). In 2002, the Hollywood film *Windtalkers*, starring Nicholas Cage, was made on the theme, and a monument to the Navajo language warriors has been erected in Phoenix, Arizona. Perhaps the surest sign of the veterans' popularity is that the very phrase "Navajo Code Talkers" is now a registered trademark of the *Navajo Code Talkers' Association*.

It should be noted, finally, that American Indians (in that case Choctaws) had been used for these purposes already in the First World War, though their history is less well documented. Also, even though the Navajos were the most numerous, members of other nations, such as the Comanches, also saw action in this role during World War Two.

A similar practice has been documented in South America. During the Chaco war against Paraguay in the early 1930s, the Bolivians used Chiquitano as a secret code.

SOURCES NOT MENTIONED IN THE TEXT

- Australian "mother-in-law" languages: Dixon (1980:58-65).
- Dialectal differences in Russian: Comrie (1990c:66), Östen Dahl (p c).
- Dialectal differences in Hausa: Newman (1990:215).
- Dialectal differences in Hakka: Li & Thompson 1979:298.
- Dialectal differences in Khanty: Comrie (1981:7).
- Lack of dialectal variation in Hebrew: McWhorter (2001:86).
- Special handicapped-directed speech varieties in the American north-west: Mithun (1999:275).
- Damin: Hale (1973), Dixon (1980:66-67).
- Pandamus register in Kalam: Lynch (1998:259).

Itsekiri and Okpe: Wolff (2000:305).

Akan: Wolff (2000:306).

Buru: Grimes & Maryott (1994).

Elaborate register differentiations in Javanese: Irvine (2001:29-30).

Politeness in Korean: Sohn (1999:207, 269, 272).

Politeness in Sinhala: Philip Baker (p c)

American Indian code talk: Singh (1999:221-30) and the *Navajo Code Talker Association*.

Comanche code-talkers mentioned in the *Shawnee News-Star* 2002-06-12 and the *Daily Oklahoman* 2002-06-12.

Bolivian use of Chiquitano as a secret war language: Adelaar (2004:478).

LANGUAGE GEOGRAPHY

WHEREVER THERE ARE PEOPLE, there are languages. Language geography is therefore necessarily a product of human settlement. Yet, languages are not necessarily dispersed in the same way as people are.

MOST LINGUISTICALLY DENSE COUNTRIES

The most linguistically dense countries are microstates or other very small countries which have attracted immigrants for various reasons (often fiscal ones). Thus, apart from the dubious case of the Vatican State, Monaco may be the linguistically most diverse sovereign country in relation to its area (1,95 km²). Almost half of its population speaks French, while a fifth speaks Italian. English, German, Dutch, Portuguese and Spanish all account for more than one percent of the population, closely followed by smaller minorities speaking Swedish and Greek.

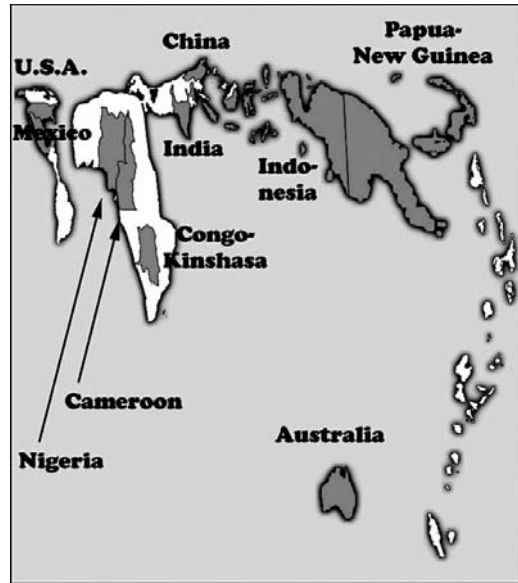
Among somewhat larger countries,

Singapore has an unusually complex linguistic situation—the *Ethnologue* lists 21 languages for this city state of 637,5 km². Excluding countries for which four languages or less are listed in the *Ethnologue* (for which one language more or less would yield an enormous difference), the 17 languages listed for the Federated States of Micronesia puts that country in the world record class with more than 24 languages per 1 000 km². It is followed by Vanuatu which, with its 109 languages has a language density of 8,9 languages/1 000 km². Most other competitors are found in the notoriously heterogeneous South-East Asia/Pacific area.

LEAST LINGUISTICALLY DENSE COUNTRIES

The countries with the smallest number of languages in relation to their respective areas are mostly territories which contain vast inhospitable deserts. In this category, we find Greenland, Kazakhstan, Saudi Arabia and Turkmenistan, all with a language density well below 0,005 languages/km².

It is worth noting that linguistically rich countries such as China, Brazil, USA, and Australia are found in this rather than the “dense” end of the spectrum. They are in other words exceptionally multilingual more by virtue of their size than anything else.



Languages are very unevenly distributed among the countries of the world. The map tries to capture this fact by rendering each country in a size corresponding to the number of languages spoken in it. (Because of the inherent problems in accomplishing this, sizes are rather approximate). The ten shaded countries are those in which more than 200 languages are in use.

MOST LINGUISTICALLY DIVERSE COUNTRIES

There is little doubt that the island of New Guinea, divided between Papua New Guinea and the Indonesian-occupied West Papua, is an area which has an unparalleled wealth of languages. On this island, which is neither

geographically nor demographically significant on a global level, something like a sixth of all the world's languages are spoken. Again according to the numbers supplied by *Ethnologue*, the most linguistically diverse countries are as follows:

Country	No. of languages	Country	No. of languages
Papua New Guinea	823	Australia	235
Indonesia	726	Congo-Kinshasa	218
Nigeria	505	China	201
India	387	Brazil	192
Mexico	288	United States	176
Cameroon	279	Philippines	169

The *Ethnologue* also uses another interesting way of measuring linguistic diversity—“Greenberg’s diversity index”. It indicates “the probability that any two people in the country picked at random will have different mother tongues”, and has a theoretical range of variation between 0 (for a completely

homogenous country) and 1 (for a hypothetical country where every single citizen has his own exclusive language). Most countries which score zero are island microstates, some of the few larger countries with a near-complete linguistic homogeneity being Jamaica, Tunisia, El Salvador, Madagascar,

Burundi and the Koreans. The most populous almost-monolingual countries are Japan and Brazil.

Most countries with exceptionally high diversity values, not surprisingly, are located in Melanesia, an area notorious for its exceptional linguistic diversity. Papua New Guinea scores highest, with 0,99, followed by its neighbors Vanuatu and the Solomon Islands, which, along with Cameroon and Tanzania, all in the 0,96-0,97 range. Even the *one hundred* most widely spoken languages of Papua New Guinea account for only about two thirds of the country's population.

There is a general tendency for poor and underdeveloped countries to be more diversified than first world states (or, to put it another way, high-tech civilizations are more or less artificially cleansed in that nation-states and education tend to more or less push citizens into the same linguistic fold). The western country (apart from tax haven Moncao and

the Vatican State) ranking the highest is Israel (0,73), followed by Belgium and Canada, both around 0,59-0,60.

Some countries show high diversity figures (above 0,70) not so much due to the presence of a plethora of indigenous languages, but rather because of contemporary immigration. The most obvious case in this category is Israel, but we also find oil-rich Gulf States of United Arab Emirates, Qatar and Bahrain, with their large contingents of South and East Asian labourers, the European microstates of Monaco, the Vatican State and Andorra, as well as American and French overseas territories such as Guam and French Guiana, where artificial economies based on metropolitan government spending attract job-seekers from neighboring third world countries.

The average value is about 0,40, and in all, as defined by the *Ethnologue*, the average country houses 44 languages—no doubt more than most people would expect.

This diversity index was first proposed by Greenberg (1956), and is defined as $DI=1-\sum(P_n)^2$, where P_n is the fraction of the total population made up by speakers of the n-th language of the country in question.

CONTINENT WITH THE FEWEST LANGUAGES

According to the *Ethnologue* statistics, Europe is by far the linguistically poorest continent—only 3% of the world's languages are spoken there. The remaining 97% are relatively equally distributed across the planet. On the other hand, it is possible to travel from the Rio Grande to northern Alaska without ever passing a language boundary, while a trip of comparable length in Europe would encompass a great deal of linguistic variation. Once aboriginal languages are extinct (and the vast majority of those that haven't already

passed into oblivion are acutely threatened, and usually on the verge of extinction), North America and Australia are going to be the most monolingual continents on earth—both of them dominated by the same language. Even now, you can (perhaps allowing for a few detours) go from Tierra del Fuego to Alaska without really crossing more than a single language boundary, which is unparalleled elsewhere in the world, except perhaps for the case of Russian, which will take you from St. Petersburg to Kamchatka. Allowing one language boundary crossing,

Mandarin would then take you into the tropics of south-east Asia.

In another sense, however, Europe is less linguistically poor than it would seem. Applying Joseph Greenberg's diversity index to entire continents, Europe is linguistically considerably more varied than Oceania or the Americas. The

explanation, of course, is that while the latter two continents have plenty of indigenous languages, these are for the most part very small indeed, and are vastly outnumbered by a small number of originally European languages. In Europe, on the other hand, the largest languages are relatively equal in size.

Africa	Asia	Europe	North America	Oceania	South America
0,96	0,93	0,92	0,62	0,60	0,56

Continent-wide diversity figures

MOST ISOLATED LANGUAGES

Of course, the most isolated language in the world would be one that we have never heard of (and there may still be a few left). On the other hand, lack of contact with western civilization need not imply that a language is isolated from its neighbors. On the contrary, in, for instance, remote areas of New Guinea which few westerners have visited, multilingualism in a number of local languages is often the norm. The most isolated languages appear to be found in the same areas where linguistic fragmentation is the highest, that is in the rainforests of the Amazonas, Indonesia

and along the border between Nigeria and Cameroon. Languages specifically mentioned as exceptionally isolated in the Ethnologue include Hixkaryána, Avá-Canoeiro and Yanomamö of Brazil (the latter also spoken in Venezuela) and Pisabo in neighboring Peru, Tugutil of Indonesia, and Akum, Beezen, Tibeá and Acipa of the Cameroon/Nigeria area. These languages are spoken by small groups of people (Avá-Canoeiro only has about 50 speakers) in remote areas with virtually no roads, and their users are for the most part not bilingual in other languages.

MOST WIDESPREAD LANGUAGES

A few languages have spread to such an extent that they have circumnavigated the globe in a rather literal sense. English is of course the most obvious case, with English-speaking New Zealand being located almost as far from Britain as one can possibly travel (hence the name Antipodes for a group of islands off New Zealand). Spanish comes close in that parts of South America are almost on the other side of the globe vis-à-vis the Philippines, which is still home to a tiny

Spanish-speaking minority. Also close, but not quite there, are Tahiti and the Seychelles, and Vanuatu and the Ivory Coast city Abidjan, all partially French-speaking. Three languages compete for the largest continuous area. At least 15 million continuous square kilometers each are dominated (not only officially, but by actual speakers) by English, Russian and Arabic. English is of course unique in completely dominating two entire continents, while English, French and

Spanish are probably the only ones having non-immigrant native language communities on all the world's inhabited continents. 28% of the world's land area is occupied by countries having English as their official language. French follows quite a bit behind, with 14%, as do Russian (12%) and Arabic (10%). Among the others, only Spanish, Portuguese, and Mandarin account for more than 5%. Romani is quite possibly the world's

most cosmopolitan language in the sense that if you are a speaker of this language, there is a 95% chance that another speaker will be from another country. For Spanish and Arabic, this figure is 88%, and for Yiddish about 78%. Despite the world-wide spread of English, the demographic weight of the United States (which is home to two thirds of all native speakers of English) reduces the figure for English to 52%.

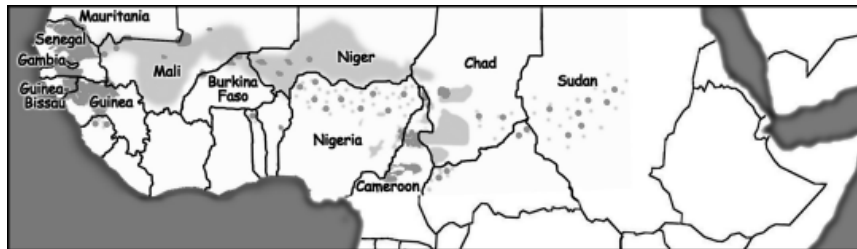
Top-10 languages by continent

Europe	North America	South America	Asia	Africa	Oceania
Russian 16,6%	English 49,7%	Portuguese 49,6%	Mandarin 22,6%	Arabic 19,0%	English 63,3%
German 12,8%	Spanish 36,7%	Spanish 43,5%	Hindi 6,8%	Hausa 4,0%	Fijian 1,5%
English 8,5%	French 2,1%	Quechua 2,6%	Bengali 5,4%	Yoruba 3,1%	Hindi 1,3%
French 8,1%	French Creoles 1,9%	Guaraní 1,4%	Japanese 3,2%	Fula 3,1%	Italian 1,2%
Italian 8,0%	English Creoles 1,1%	Aymara 0,5%	Panjabi 2,9%	Igbo 2,6%	Samoaan 1,0%
Polish 5,5%	German 0,9%	Italian 0,5%	Arabic 2,5%	Oromo 2,5%	Cantonese 0,9%
Ukrainian 4,6%	Sinitic lgs 0,7%	German 0,3%	Wu 2,4%	Amharic 2,5%	Greek 0,9%
Spanish 4,5%	Italian 0,4%	Japanese 0,2%	Javanese 2,2%	Malagasy 2,0%	French 0,9%
Romanian 3,3%	Tagalog 0,3%	English Creoles 0,2%	Telugu 2,1%	Somali 1,4%	Enga 0,8%
Dutch 2,9%	Náhuatl 0,3%	Mapudungun 0,1%	Marathi 2,0%	Ciluba 1,2%	Arabic 0,7%

Here is a way of making Italian appear as the most wide-spread language on earth: it is the only language that makes it into a top-ten list of languages on four of the world's six permanently inhabited continents. English, French, German, Spanish and Arabic only do so for three continents each (as does Chinese, if you combine the Sinitic languages). With the exception of Arabic and Italian, these are also among the ten most widely spoken languages among the population of transient scientists on the seventh continent, Antarctica. Note that, contrary to popular belief, there are more speakers of Portuguese than of Spanish in South America.

ODDEST GEOGRAPHICAL DISTRIBUTION

The Fula language is thought to have originated somewhere in the area of present-day Senegal and Guinea. Beginning in the 12th century, however, the Fulas began an eastward expansion, so that they now occupy a band stretching for about 5000 kilometers from the Atlantic to the Nile. As the pious Moslems that they generally are, the Fulas have thus striven ever closer to Mecca for more than eight hundred years!



The expansion is in part due to a successful adaptation to cattle raising in the arid Sahel climate, and in part through sheer conquest. In the eastern parts of the area, Fula has managed to establish itself as an important regional lingua franca in parts of Cameroon and Nigeria (see pp 66–68).

Known under a variety of names (e.g.

Fulani, Fulbe, Fulfulde, Peul, Pulaar, and Toucouleur), Fula now has 26 million speakers. It is the biggest language of Guinea, the second in Senegal, Cameroon, Guinea-Bissau, Burkina Faso, Mauritania, and the Gambia, and third in Niger, Mali and Nigeria. Close to two million additional speakers are scattered over eight other West African countries.

ON THE WRONG SIDE OF THE BORDER

The Bering Strait divides Siberia and Alaska from one another. On both sides, people speak Yup'ik—the only language indigenous to both the New World and the Old. The official language of the Siberian side is Russian, while that of the Alaskan side is English. However, when the Russians arrived in the region, they came overland (except for the crossing of the strait, of course), and took more interest in the American side than in the areas which today remain Russian. The English-speakers, on the other hand, came by sea, and had partly different commercial interests, for which reason they had more contact with the Asian coast. A couple of centuries later, traces of this situation are

still linguistically visible—Siberian Yup'ik has more loan words from English than has its Alaskan sibling, which in turn has incorporated more Russian terms. Thus, for instance, the word for 'cow' in Siberian Yup'ik is derived from English *cow*, while the Alaskan variety uses a form derived from Russian *корова* (*korova*).

Another case of language material winding up in unexpected places is provided by Dutch-lexicon creoles. Three such languages survived in the Caribbean area into the 20th century: Negerhollands, Berbice, and Skepi. Similarly, three areas remained under Dutch rule: the ABC islands (Aruba, Bonaire, and Curaçao), the SSS islands (Saba, St

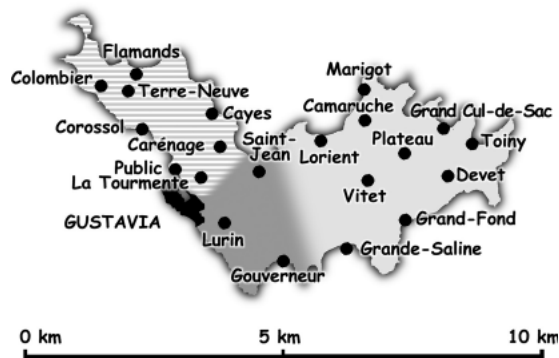
Eustatius, a k a ‘Statia’, and the southern half of St Martin (see pp 25–26) and Dutch Guiana (now independent under the name Surinam). Creole languages were spoken in all three areas, but none was lexically based on Dutch. Papiamentu, the main language of the ABC islands, draws its lexicon first and foremost from Spanish, while an assortment of English-lexicon creoles and creoloids are used on the

SSS islands and in Surinam. Meanwhile, the three Dutch creoles were all used in countries where English is the official language: Negerhollands was spoken on the U S Virgin Islands, and the other two in British Guiana (now Guyana *tout court*). In other words, at least in the 20th century Caribbean, Dutch rule and the use of a Dutch-derived language were in a perfectly complementary distribution.

AN ODD GEOLINGUISTIC SITUATION

The tiny Caribbean island of Saint-Barthélemy, or St Barth, presents a remarkable linguistic fragmentation. Despite being a mere 10 kilometers across, and home to little more than 3 000 people, it has traditionally had at least four distinct languages. In the north-western part, a Norman patois has some 500-700 speakers, and in the east, a French-lexicon Creole is used by 600-800 people. In

the middle region, an archaic variety of French is now on the verge of extinction. In addition to this, in the administrative center of Gustavia, a black population, some 100 strong, speaks a local and slightly creolized variety of English. On top of this, standard French is the official language, and is increasingly used. (Members of all groups born after about 1970 usually have standard French as their mother tongue).



St Barth and its language boundaries

While this situation is not particularly uncommon in the world as a whole, it is rare in the western world, especially in a recently colonized area—such linguistic fragmentation usually requires far more time than just a couple of centuries to develop.

Understandably, this diversity requires isolation, and as late as in the 1940s, there were people in the northwestern village of Flamands who only visited Gustavia once a year. And yet, the distance between the two is about two kilometers!

MAKING THE MOST OUT OF A SINGLE ISOGLOSS

Particularly since the Napoleonic Wars, language has been seen as a basis for nationality, and thus for drawing country boundaries. For instance, one of the reasons for recognizing Rhaeto-Romance in Switzerland as a separate language was to counter Mussolini's claims to the southern parts of the country.

In continuum situations, however, language boundaries are difficult to establish, and at the most, isoglosses can be identified. Thus, the fact that Serbian and Macedonian—

as opposed to Bulgarian—used [c] or [tʃ] as a reflex of proto-Slavic *tj was exploited by Serbian nationalist Aleksandar Belić in 1919 to support Serbia's claim to most of what is today Macedonia. On the other hand, the use of a suffixed definite article in Macedonian (as well as in Bulgarian, but not in most varieties of Serbian) was sufficient for some Bulgarian writers to demonstrate that not only Macedonia, but also the Timok-Morava valley in southern Serbia should really be under Bulgarian jurisdiction.

DENSEST ISOGLOSSES

Baarle-Hertog is a Belgian enclave in the Netherlands. On most maps detailed enough to show the community in the first place, it simply figures as a dot.



Baarle presents an intricate patchwork of Dutch and Belgian territory, even including enclaves-within-enclaves. In the sketch above, the shaded territory is Belgian

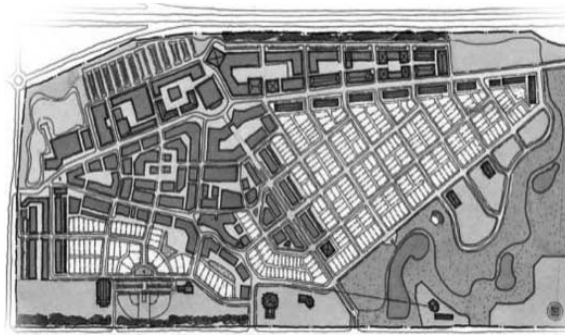
However, on a really detailed map, the enclave turns out to consist of innumerable little unconnected patches. Not only one, but a multitude of frontiers run through the town, and some houses are actually divided between the two nations (in which case the owner is considered a citizen of the country in which his front door happens to be). This division goes back to 1190, when the duke of Brabant ceded the territory to the lord of Breda, with the exception of the land which was built-up at the time.

On both sides of the intricate line, Dutch is spoken. A Dutch dialectologist who visited the area in 1956 was therefore first sceptical about the locals' claims that the utterly complex state border was actually also a dialect border. But when he questioned people, however, it turned out that the Belgians indeed did say *gazet*, *fakteur*, *madam* and *sjuust* ('newspaper', 'post', 'mrs' and 'just') where the Dutch would use *krant*, *post*, *mevrouw* and *precies*. The differences even included certain phonetic features.

An isogloss is a line on a map indicating the geographical division between two comparable linguistic features. Such features can include different words for the same thing (for example, the locally preferred word for 'a stream' is generally *beck* in some of the most northerly English counties and *brook* elsewhere), different pronunciations of the same word (e.g. there are three widespread regional pronunciations of 'last' in England: [læst], [last] and [la:st]), or different uses of verb forms (e.g. *instead* of standard 'we are' the dialectally preferred form in an area south and west of Worcestershire and Oxfordshire (but excluding Devon and Cornwall) is *we be*). Isoglosses for English are often based on the usage of older people to reveal dialectal differences which may now be declining.

VOLUNTARY LINGUISTIC SEGREGATION IN SOUTH DAKOTA

In 2005, the Michigan-based firm Nederveld Associates, Inc. announced their plan to construct an entirely new town in South Dakota.



The Laurent town plan

The name chosen for the community is Laurent, after Laurent Clerc, who introduced French Sign to the United States (see p 127). Located 50 km west of Sioux Falls, Laurent is to receive the first of its envisaged 3 000 residents in 2006.

While this is unusual in itself (especially since no new settlements had been established in that state for a century), the truly unique feature of the town is that it is to be designed specifically for sign language users.



A street view in Laurent, South Dakota

One design feature intended to cater for the future deaf residents is a fully bilingual (English and American Sign) school system. Also, public spaces are to be designed with maximal visibility in mind, to facilitate signing. Shop assistants and officials would of course be bilingual, and emergency vehicles would have more lights and fewer sirens. Internet connections are to pay special attention to the needs of the hearing-impaired, so that signing can be easily transmitted over video.

SOURCES NOT MENTIONED IN THE TEXT

Unless otherwise mentioned, the number of speakers of various languages, in this and the following sections, are from Parkvall (n d), while population figures are from the 2005 edition of the *CIA World Factbook*.

Languages in Monaco: based on birthplace and citizenship data in the 2000 Monaco census. The semi-official indigenous Monégasque language is assumed to be extinct (cf Shaul 2002).

Fula's geographical distribution: Olson (1996:181-82), Arnott (1974), Parkvall (n d).

Russian and English loans in Yup'ik: Comrie (1981:258).

Presence and absence of Dutch creoles: Holm (1989).

Linguistic situation in St Barth: Maher (1997b), Calvet & Chaudenson (1998).

Making the most of an isogloss: Friedman (1996).

Baarle-Hertog: Weijnen (1957:4-5), Thijs (2000).

Laurent, South Dakota: Press Release from Nederveld Associates Inc., Official web site of Laurent at <http://www.laurentsd.com> (2005-03-30).

LANGUAGE AS A LEGAL MATTER

MOST COUNTRIES ON EARTH have a linguistic legislation of one kind or the other. About two thirds specify in the constitution that this or that is the official language, but in most cases, it is hastily added that all languages are equal. Upon closer inspection, however, it is painfully obvious that some languages are more equal than others.

THE WORLD'S MOST DANGEROUS LANGUAGE

Although many, and perhaps even most countries have exerted repression on minority languages, it is less common that entire languages have been declared illegal, as has been the case with Kurdish in Turkey. To be sure, children have been punished for speaking the wrong language in school (Basque in Franco's Spain, Welsh in Britain, French in the USA, Aboriginal languages in Australia, Saami in Sweden, etc.), or laws have been adopted to reduce the public use of undesired languages (such as the *Loi Toubon* in France). Though extreme, the Turkish example is not unique.



Probably no other language has invoked such fear with authoritarian regimes as has Esperanto.

There is another language, however, which has been banned as such, and in more countries than any other—Esperanto. There is no complete list of countries where Esperanto has been banned, but quite a few are mentioned in Ulrich Lins' book *La Danĝera Lingvo* ("the dangerous language") including Nazi Germany, Stalin's Soviet Union, China, Yugoslavia, Austria, Japan, Bulgaria, Portugal, and Spain. As late as in 1990, an Iraqi Esperantist was let out of prison only on the condition that he never teach the language again.

LANGUAGE CONFLICT INVOLVING ONLY ONE ETHNIC GROUP

In Greece, *katharevousa* (literally 'pure language'), was the official language from independence in the 1820s until the end of military rule in 1976. Yet, most people spoke

dimotiki ('people's language'), which had not been subject to the artificial cleansing of foreign (in particular Turkish) loans that *katharevousa* had. The lower variety subsequently became

instrumentalized, but a classical diglossic situation still persisted. Although only one ethnic group was involved, this diglossia (see p 53) was not unproblematic. The translation

of the gospels and the writings of Aischylos into *dimotiki* around 1900 led to riots, resulting in several deaths and the turnover of the government.

EARLIEST DOCUMENTED LANGUAGE-BASED DISCRIMINATION

The biblical Book of Judges provides some details on the battles between Gileadites and Ephraimites. Both peoples spoke Hebrew, but with dialectal differences. A good way, then, to distinguish members of one group from the other was to ask them to pronounce the word *shibboleth* ('torrent'). Apparently,

Ephraimites would normally depalatalize this word into *sibboleth*. Doing so at a Gileadite checkpoint, however, had fatal consequences, as it revealed the ethnicity of the speaker. If we are to believe the Book of Judges (chapter 12, verses 5–6), 42 000 Ephraimites were thus slain and tossed into the Jordan.

THE INFAMOUS LANGUAGE LEGISLATION IN FRANCE

In 1994, the Mitterrand government passed the *Loi no. 94-665 du 4 août 1994 relative à l'emploi de la langue française*, better known as the *Loi Toubon*, after Jacques Toubon, then Minister of Culture. Contrary to popular belief, it does not really outlaw the use of languages other than French, but it states, among other things, that public signs "should be in French", although the French text may be accompanied by a "translation into a foreign language", provided that the French version is "equally legible, audible or intelligible"

(articles 3 and 4).

More seriously, the law forbids the public use of "foreign terms and expressions" whenever there is a French counterpart. Even more worryingly, "foreign languages" in this context refers to anything other than French, that is, it includes languages such as Breton, Basque, Occitan, and Creole, some of which are in fact majority languages in parts of the French Republic.

Similar laws exist, by the way, in a number of other democracies, e.g. Poland and Mexico.

MOST REPRESSIVE MINORITY LANGUAGE POLICY

Even though many atrocities in world history have affected certain linguistic groups more than others, most ethnicity-based human rights violations have more complex roots, often also involving religion and other factors.

Nevertheless, most countries have a history of repression of languages perceived

as a threat to the ethnicity from which the élite caste is recruited. Even in states with relatively old democratic traditions, bans on using languages other than the official one in public milieux persisted well into the 20th century.

In recent years, much of this hostility has been toned down, but there are some remarkable cases where exceptionally

oppressive language policies are still in force. Chief among these is probably Turkey.

Speakers of Turkish form a massive majority of the country's population, but between a fifth and a tenth are believed to be speakers of Kurdish.

According to official Turkish policy, however, they did not exist until recently as a separate ethnicity, and were until 1991 referred to as "mountain Turks". Kurdish schools, associations, names, and even music were banned.

"No spitting on the floor or speaking Breton"



In the bad old days, not too long ago, even countries now regarded as relatively civilized, had laws banning one language or another. The Americans forbade French in Louisiana and Maine, while the French forbade Breton in France.

The very constitution of the country makes the point quite clear in stating that "No language prohibited by law shall be used in the expression and dissemination of thought" (article 26), that "Publication [of newspapers] shall not be made in any language prohibited by law" (article 28), and that "No language other than Turkish shall be taught as mother tongue to Turkish citizens at any institutions of training or education" (article 42).

Specific laws, such as Law 2932, passed in 1983, further stipulate that the conveying of any idea in languages which are not official languages of other nations (which, of course, Kurdish is not) be punished with imprisonment. Amazingly, Law 2932 also states contrafactually that the mother tongue of all Turkish citizens is Turkish (!!!).

There may be light at the end of the tunnel, however. Due mainly to pressure from the European Union, which Turkey is eager to join, some of the above laws were repealed in the 1990s, and constitutional changes were announced in October 2001 and again in August 2002. June 2004 even saw Turkish national television broadcast a short program in Kurdish for the first time ever.

Yet, in terms of minority language policy (and also respect for human rights in general), Turkey has little to be proud of.

To be sure, there are other contenders for the title. Leclerc (2006; in my translation) describes the language policy of Syria as "unacceptable", as it "ignores the most fundamental rights" of the nation's inhabitants, while that of Turkmenistan is characterized as "pulverization of the linguistic rights of all minorities". He does not give much in the way of examples from Burma, but nevertheless summarizes the situation in that country in a way which makes further analysis redundant. Burma, Leclerc concludes, is "at the very bottom of the abyss".

Yet another case of an authoritarian regime and an assimilatory language policy is Malaysia. In this country, the policy enforced—consisting essentially in the promotion of Malay and acceptance of English—is beyond discussion. The Sedition Act of 1948 and the Constitution (Amendment) Act of 1971 explicitly state that the status of Malay "may no longer be questioned".