THE LANGUAGES OF THE SOVIET UNION
CAMBRIDGE LANGUAGE SURVEYS


This new series will offer general accounts of all the major language families of the world. Some volumes will be organised on a purely genetic basis, others on a geographical basis, whichever yields the most convenient and intelligible grouping in each case. Sometimes, as with the Australian volume, the two in any case coincide.

Each volume will compare and contrast the typological features of the languages it deals with. It will also treat the relevant genetic relationships, historical development, and sociolinguistic issues arising from their role and use in the world today. The intended readership is the student of linguistics or general linguist, but no special knowledge of the languages under consideration is assumed. Some volumes will also have a wider appeal, like those on Australia and North America, where the future of the languages and their speakers raises important social and political issues.

Already published:
The languages of Australia R. M. W. Dixon

Forthcoming titles include:
Japanese/Korean M. Shibatani and Ho-min Sohn
Chinese J. Norman and Mei Tsu-lin
S. E. Asia J. A. Matisoff
Dravidian R. E. Asher
Austronesian R. Blust
Afro-Asiatic R. Hetzron
North American Indian W. Chafe
Mesoamerican Indian J. A. Suarez
Slavonic R. Sussex
Germanic R. Lass
Celtic D. Macaulay et al.
Indo-Aryan C. P. Masica
Balkans J. Ellis
Creole languages I. F. Hancock
Romance languages R. Posner
THE LANGUAGES
OF THE SOVIET UNION

BERNARD COMRIE
Associate Professor of Linguistics
University of Southern California

CAMBRIDGE UNIVERSITY PRESS
Cambridge
London New York New Rochelle Melbourne Sydney
CONTENTS

Tables viii
Preface xi
Abbreviations xv
Notes on transcription xvi
Map of the peoples of the U.S.S.R. xviii
Index to map xxii

1 Introduction 1
1.1 Language and ethnic group: census statistics 2
1.2 Classification of the languages of the U.S.S.R 9
1.2.1 Genetic classification 10
1.2.2 Typological classification 11
1.2.3 Sociological classification 15
1.2.4 Historical background 19
1.3 Language policy 21
1.3.1 Pre-Revolutionary policy 21
1.3.2 Soviet policy 22
1.3.3 Autonomous administrative units 29
1.4 The influence of Russian 31
Further reading 37

2 Altaic languages 39
2.1 The individual languages 42
2.1.1 Turkic languages 42
2.1.2 Mongolian languages 54
2.1.3 Tungusic languages 57
2.2 Phonology 59
2.3 Morphological structure and categories 71
2.4 Syntax 77
Texts in Altaic languages 85
## Contents

Tatar text 85
Buryat text 87
Orok text 89
Further reading 90

3 Uralic languages 92
3.1 The individual languages 95
3.1.1 Balto-Finnic languages 95
3.1.2 Lapp 101
3.1.3 Volgaic languages 102
3.1.4 Permic languages 103
3.1.5 Ugric languages 105
3.1.6 Samoyedic languages 107
3.2 Phonology 109
3.3 Morphological structure and categories 117
3.4 Syntax 120
   Texts in Uralic languages 136
   Estonian text 136
   Khanty text (Vakh dialect) 138
   Selkup text 139
   Further reading 140

4 Indo-European languages 142
4.1 The Indo-European family in the U.S.S.R. 142
4.2 Balto-Slavonic languages 142
4.2.1 Slavonic languages 144
4.2.2 Baltic languages 146
4.3 Indo-Iranian languages 155
4.3.1 Indic (Indo-Aryan) languages 155
4.3.2 Iranian languages 158
4.4 Armenian 179
4.5 Germanic languages 184
4.6 Romance languages 187
   Texts in Indo-European languages of the U.S.S.R. 189
   Lithuanian text 189
   Tadzhik text 191
   Armenian text 192
   Further reading 194

5 Caucasian languages 196
5.1 The individual languages and their subgrouping 196
5.2 Phonology 200
Contents

5.3 Morphology  207
5.4 Syntax  
   Texts in Caucasian languages  231
   Georgian text  231
   Abkhaz text  233
   Avar text  235
   Further reading  236

6 Paleosiberian and other languages  238

6.1 Chukotko-Kamchatkan languages  240
6.2 Eskimo-Aleut languages  252
6.3 Yukagir  258
6.4 Ket  261
6.5 Nivkh  266
6.6 Aramaic (Assyrian)  272
6.7 Dungan  273
   Texts in Paleosiberian languages  274
   Chukchi text  274
   Yukagir text  275
   Nivkh text  276
   Further reading  278

Appendix 1 Ethnic and linguistic composition of the U.S.S.R. according to the 1970 census  279
Appendix 2 Ethnic administrative areas of the U.S.S.R.  282
Appendix 3 Alphabets of the languages of the U.S.S.R.  286

References  290
Addendum: Recent statistics and literature  301
Language index  304
Subject index  315
TABLES

1.1 Morphological typology of Lithuanian and Tatar 13
2.1 Singular personal pronouns in Altaic, Uralic, Indo-European, and Chukotko-Kamchatkan 41
2.2 Genetic classification of modern Turkic languages 46
2.3 Genetic classification of the Tungusic languages 58
2.4 Kirgiz vowel quality system 60
2.5 Illustrations of Kirgiz vowel harmony 61
2.6 Illustrations of Azerbaydzhan vowel harmony 62
2.7 Even vowel system 70
2.8 Evenki vowel system 71
2.9 Illustration of Evenki declension 72
2.10 Illustrations of Evenki conjugation 72
3.1 Genetic classification of the Uralic languages 94
3.2 Consonant gradation in Karelian 114
3.3 Subject agreement in Estonian 125
3.4 Indefinite object and definite object conjugations in Hungarian 126
3.5 Definite and indefinite conjugations of Mansi tee- ‘eat’ 126
3.6 Erzya-Mordva verb forms (past tense of kund- ‘catch’) showing subject and object agreement 127
4.1 Branches of the Indo-European family (modern languages) 143
4.2 Stress and tone in Lithuanian declension 150
4.3 European dialects of Romany 158
4.4 Genetic classification of modern Iranian languages 161
4.5 Declensional forms of Ossete (Iron) san ‘head’ 172
4.6 Indicative verb forms in Kurdish of Azerbaydzhan 174
4.7 Declensional forms of Armenian ban ‘word, thing’ 181
5.1 Genetic classification of the Caucasian languages 197
5.2 Consonant system of Georgian 201
5.3 Consonant system of Avar 201
<table>
<thead>
<tr>
<th>Tables</th>
<th>ix</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4 Consonant system of Ubykh</td>
<td>202</td>
</tr>
<tr>
<td>5.5 Consonant system of the Bzedug dialect of Adyge</td>
<td>204</td>
</tr>
<tr>
<td>5.6 Vowel system (including diphthongs) of Lowland Chechen</td>
<td>205</td>
</tr>
<tr>
<td>5.7 Locative cases in the Kubachi dialect of Dargva</td>
<td>210</td>
</tr>
<tr>
<td>6.1 Consonant system of Chukchi</td>
<td>243</td>
</tr>
<tr>
<td>6.2 Verb forms in Aleut dialects of the U.S.S.R.</td>
<td>253</td>
</tr>
<tr>
<td>6.3 Verb agreement and focus in Yukagir (past tense forms)</td>
<td>259</td>
</tr>
<tr>
<td>6.4 Obstruent system of Nivkh (Amur dialect)</td>
<td>267</td>
</tr>
<tr>
<td>A.1 Russian (Cyrillic) alphabet</td>
<td>286</td>
</tr>
<tr>
<td>A.2 Chukchi text in current (Cyrillic) orthography</td>
<td>287</td>
</tr>
<tr>
<td>A.3 Modern Armenian alphabet</td>
<td>288</td>
</tr>
<tr>
<td>A.4 Armenian text in current orthography</td>
<td>289</td>
</tr>
<tr>
<td>A.5 Modern Georgian alphabet (<em>mxedruli</em>)</td>
<td>289</td>
</tr>
<tr>
<td>A.6 Georgian text in current orthography</td>
<td>289</td>
</tr>
</tbody>
</table>
This book has two main aims: first, to introduce the reader to some of the salient linguistic features of the various languages and languages-families of the U.S.S.R.; secondly, to give some indication of how the various languages of the U.S.S.R. interact in a multilingual society, especially of how they interact with Russian. Chapter 1 and the early parts of subsequent chapters and sections deal primarily with sociolinguistic background, while the body of the book deals with structural features.

In the sociolinguistic discussion I have taken for granted the existence of the U.S.S.R. as a multinational and multilingual entity, and have discussed the language policies that have been adopted in the light of this situation, including actual or implied comparisons (on the whole, favourable) with policies of the Tsarist government and policies adopted in other multilingual countries. A different, though, in the current political situation, unrealistic perspective, would have been gained by comparing how the languages in question might have fared if their speakers had formed independent nation-states outside the U.S.S.R.

The selection of linguistic (phonological, morphological, and syntactic) topics has necessarily been very restricted by the nature of the volume. Rather than attempt to give a superficial survey of every aspect of language structure, I have preferred to concentrate on those aspects of individual languages and language-families that strike me as particularly interesting and important from a general linguistic viewpoint. For this reason considerable space is, for instance, devoted to phonetics in Chapter 5 (Caucasian languages), to vowel harmony in Chapter 2 (Altaic languages), and to case-marking and negation in Chapter 3 (Uralic languages).

Likewise, the list of references has been considerably more restricted than would have been the case in a strict introduction to the comparative philology of one or other of the language families discussed. In general, no reference has been given for statements about individual languages or language-families that are uncontroversial; likewise, straightforward illustrative examples have often been taken or adapted from standard handbooks listed in the references without explicit statement
Preface

of the source. Apart from references cited in the text, other references (i.e. in the Further reading sections at the end of each chapter) are designed, primarily, to direct the reader to one or more comprehensive grammars of each of the languages concerned, preference being given to those written in more accessible languages. In addition, references have been given to general works dealing with languages and language policies in the U.S.S.R., and to handbooks dealing with the language families covered. But the list of references is explicitly not claimed to be comprehensive, in particular in not including lexicographical materials or monograph studies on specific aspects of individual languages.

At the end of each chapter there is a sample of texts in the languages discussed, to give a more global, functioning picture of these languages. The texts have been selected solely for their illustrative value, from a range of styles (traditional stories, Soviet fiction, scientific writing), and the views expressed as the content of these texts are not necessarily my own.

Substantial portions of the text were written by two of my colleagues: J. R. Payne (University of Birmingham) wrote section 4.3.2 (Iranian languages) and also prepared the Tadzhik text; B. G. Hewitt (St John’s College, Cambridge) wrote Chapter 5 (Caucasian languages) and collaborated with me in the writing of section 4.4 (Armenian). All other chapters and sections were written by me, and I also took overall responsibility for editing the whole volume. The glossing and translation of texts, and the translation of citations from foreign languages, were done by the author of the respective section/chapter.

Examples from languages of the U.S.S.R. are presented in accordance with the transcription system discussed on pp. xvi–xvii. Where relevant, hyphens are used in examples to indicate morpheme boundaries, though not all morpheme boundaries are so indicated (e.g. where they are not relevant to the point at issue). A morpheme-by-morpheme gloss is provided into English (with glosses of bound morphemes, in block capitals, abbreviated according to the list on p. xv); where necessary, this is accompanied by a freer translation.

The transliteration into Latin script of names of people, places, and languages in the U.S.S.R. presents immense problems, and no one system devised has met with universal acceptance. In the present work the following practices have been adopted. Russian family names have been transliterated according to the International system (see p. 286), family names of other peoples of the U.S.S.R. according to the transliteration of the Russian form of their name; where this differs from the form actually given on the title page of a book or article, the actual form, or a transliteration thereof, is appended in square brackets. Where the only reference by an author uses the Latin script, then this spelling is retained. Russian book and article titles and common nouns are also cited in the International system.
Preface

Place-names in the U.S.S.R. have been given in the form used by the *Times atlas of the world*, though with the alternatives Moskva/Moscow I have preferred the latter. For names of administrative units, however, where this atlas simply transliterates the Russian name, I have preferred to translate, e.g. South-Ossete Autonomous Oblast rather than Yugo-Osetinskaya Avtonomnaya Oblast; Appendix 2 includes the appropriate name in the form found in the *Times atlas of the world* where this departs radically from the form used here.

For language (and ethnic) names, accepted English terms have been used for those few languages of the U.S.S.R. whose names are in frequent currency among English-speaking linguists, such as Georgian (Russian: грузинский), Lapp (Russian: саамский). For all others, I have used the standard Soviet Russian name, shorn of Russian-specific suffixes, in the usual nontechnical transliteration (i.e. as used in the *Times atlas of the world*). In addition to removing Russian suffixes, I have tried not to add English-specific suffixes, to make the stem of the language name quite transparent, i.e. Vot (rather than Votic); wherever possible, even the plural suffix has been omitted. This policy has been dictated by the orientation of the book as a volume on the languages of the U.S.S.R. rather than as an introduction to the philologies of individual language-families. No-one will be happy with all the forms we have chosen, which often do not correspond to those used by non-Soviet specialists (most noticeably for Uralic languages), or fail to make phonetic distinctions that specialists prefer to maintain in language names (as with many Caucasian languages). As small consolation to anyone offended by any of the spellings adopted, I have had to sacrifice my own favourite spelling of Chukchee for Chukchi. The language index gives the Russian form (transliterated) for each of the languages of the U.S.S.R. in parentheses after the English form adopted in this work.

In citations from other authors, differences concerned solely with transliteration have been removed in favour of the principles above. In the list of references, original spellings are, of course, retained.

In preparing this volume, the latest detailed statistical materials available to me were from the census of 1970. These statistics are already somewhat dated, in that a further census was carried out in 1979, but as the detailed breakdown of the returns of this census are unlikely to be available for some years, I have continued to use the present tense to refer to the situation as mirrored in the 1970 returns.

My interest in writing this book arose largely from a project on the linguistic typology of the non-Slavonic languages of the U.S.S.R., supported during 1975–8 by the Social Science Research Council, London: B. G. Hewitt and J. R. Payne also participated in this project. Although the present volume is not directly part of that project, the material we worked on during that period has been of immense help in preparing this volume. The administrative and scientific staff of various institutes in
Preface

the U.S.S.R. have helped us in gathering and analysing material on the languages of the U.S.S.R., and we would like to express our thanks to members of the following institutes: Academy of Sciences of the U.S.S.R. (Linguistics Institute, Leningrad Section of the Institute of Linguistics, Leningrad Section of the Oriental Institute), Tbilisi State University, Sukhumi University and Research Institute, Academy of Sciences of the Georgian S.S.R. (Linguistics Institute), Academy of Sciences of the Tadzhik S.S.R. (Language and Literature Institute). Our participation in exchange programmes with these institutes has been facilitated by the cooperation of the British Academy, the British Council, and the University of Birmingham (Centre for Russian and East European Studies). I am also grateful to R. Austerlitz and R. Hetzron for comments on the manuscript.

The editing of the final manuscript was carried out while I was a Visiting Fellow in the Department of Linguistics, Faculty of Arts, the Australian National University.

Bernard Comrie
September 1979
# Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABS</td>
<td>Abessive ('without X')</td>
</tr>
<tr>
<td>ABL</td>
<td>Ablative ('from X')</td>
</tr>
<tr>
<td>ABS</td>
<td>Absolute, Absolute</td>
</tr>
<tr>
<td>ACC</td>
<td>Accusative</td>
</tr>
<tr>
<td>ACT</td>
<td>Active</td>
</tr>
<tr>
<td>ADJ</td>
<td>Adjective</td>
</tr>
<tr>
<td>ADV</td>
<td>Adverbial</td>
</tr>
<tr>
<td>AFF</td>
<td>Affective</td>
</tr>
<tr>
<td>AFFIRM</td>
<td>Affirmative</td>
</tr>
<tr>
<td>ALIEN</td>
<td>Alienable possession</td>
</tr>
<tr>
<td>ALL</td>
<td>Allative ('to(wards) X')</td>
</tr>
<tr>
<td>AOR</td>
<td>Aorist</td>
</tr>
<tr>
<td>AUG</td>
<td>Augmentative</td>
</tr>
<tr>
<td>CAUS</td>
<td>Causative</td>
</tr>
<tr>
<td>COL</td>
<td>Column</td>
</tr>
<tr>
<td>COMIT</td>
<td>Comitative ('together with X')</td>
</tr>
<tr>
<td>COMPAR</td>
<td>Comparative</td>
</tr>
<tr>
<td>COND</td>
<td>Conditional</td>
</tr>
<tr>
<td>CONT</td>
<td>Continuous</td>
</tr>
<tr>
<td>COP</td>
<td>Copula</td>
</tr>
<tr>
<td>DAT</td>
<td>Dative</td>
</tr>
<tr>
<td>DEF</td>
<td>Definite</td>
</tr>
<tr>
<td>DIM</td>
<td>Diminutive</td>
</tr>
<tr>
<td>DO</td>
<td>Direct Object</td>
</tr>
<tr>
<td>DU</td>
<td>Dual</td>
</tr>
<tr>
<td>DYN</td>
<td>Dynamic (Nonstatic)</td>
</tr>
<tr>
<td>ELAT</td>
<td>Elative ('out of X')</td>
</tr>
<tr>
<td>EMPH</td>
<td>Emphatic</td>
</tr>
<tr>
<td>ERG</td>
<td>Ergative</td>
</tr>
<tr>
<td>ESS</td>
<td>Essive ('as X')</td>
</tr>
<tr>
<td>EXCL</td>
<td>Exclusive</td>
</tr>
<tr>
<td>FEM</td>
<td>Female, Feminine</td>
</tr>
<tr>
<td>FIN</td>
<td>Finite</td>
</tr>
<tr>
<td>FOC</td>
<td>Focus</td>
</tr>
<tr>
<td>FREQ</td>
<td>Frequentative</td>
</tr>
<tr>
<td>FUT</td>
<td>Future</td>
</tr>
<tr>
<td>GEN</td>
<td>Genitive</td>
</tr>
<tr>
<td>GER</td>
<td>Gerund (Verbal adverb)</td>
</tr>
<tr>
<td>HABIT</td>
<td>Habitual</td>
</tr>
<tr>
<td>HUM</td>
<td>Human</td>
</tr>
<tr>
<td>ILL</td>
<td>Illative ('into X')</td>
</tr>
<tr>
<td>IMPER</td>
<td>Imperative</td>
</tr>
<tr>
<td>IMPERF</td>
<td>Imperfect</td>
</tr>
<tr>
<td>INAN</td>
<td>Inanimate</td>
</tr>
<tr>
<td>INCH</td>
<td>Inchoative</td>
</tr>
<tr>
<td>INCL</td>
<td>Inclusive</td>
</tr>
<tr>
<td>INDEF</td>
<td>Indefinite</td>
</tr>
<tr>
<td>INESS</td>
<td>Inessive ('in X')</td>
</tr>
<tr>
<td>INFER</td>
<td>Inferential</td>
</tr>
<tr>
<td>INFIN</td>
<td>Infinitive</td>
</tr>
<tr>
<td>INSTR</td>
<td>Instrumental</td>
</tr>
<tr>
<td>ITER</td>
<td>Iterative</td>
</tr>
<tr>
<td>LOC</td>
<td>Locative</td>
</tr>
<tr>
<td>MASC</td>
<td>Masculine</td>
</tr>
<tr>
<td>N</td>
<td>Noun</td>
</tr>
<tr>
<td>NARR</td>
<td>Narrative</td>
</tr>
<tr>
<td>NEG</td>
<td>Negative</td>
</tr>
<tr>
<td>NEUT</td>
<td>Neuter</td>
</tr>
<tr>
<td>NOM</td>
<td>Nominaive</td>
</tr>
<tr>
<td>OBJ</td>
<td>Object</td>
</tr>
<tr>
<td>OBL</td>
<td>Oblique</td>
</tr>
<tr>
<td>OPT</td>
<td>Optative</td>
</tr>
<tr>
<td>PART</td>
<td>Participle</td>
</tr>
<tr>
<td>PARTIT</td>
<td>Partitive</td>
</tr>
<tr>
<td>PASS</td>
<td>Passive</td>
</tr>
<tr>
<td>PERF</td>
<td>Perfect</td>
</tr>
<tr>
<td>PL</td>
<td>Plural</td>
</tr>
<tr>
<td>PLUPERF</td>
<td>Pluperfect</td>
</tr>
<tr>
<td>POSS</td>
<td>Possessive</td>
</tr>
<tr>
<td>POT</td>
<td>Potential</td>
</tr>
<tr>
<td>PRES</td>
<td>Present</td>
</tr>
<tr>
<td>PRESUP</td>
<td>Presuppositional</td>
</tr>
<tr>
<td>PREV</td>
<td>Preverb</td>
</tr>
<tr>
<td>PROG</td>
<td>Progressive</td>
</tr>
<tr>
<td>PROHIB</td>
<td>Prohibitive</td>
</tr>
<tr>
<td>PROL</td>
<td>Prolative ('along X')</td>
</tr>
<tr>
<td>PROTR</td>
<td>Protracted</td>
</tr>
<tr>
<td>Ptcl</td>
<td>Particle</td>
</tr>
<tr>
<td>PURP</td>
<td>Purposive</td>
</tr>
<tr>
<td>Q</td>
<td>Question</td>
</tr>
<tr>
<td>REC</td>
<td>Recent</td>
</tr>
<tr>
<td>RECIP</td>
<td>Reciprocal</td>
</tr>
<tr>
<td>REFL</td>
<td>Reflexive</td>
</tr>
<tr>
<td>REL</td>
<td>Relative</td>
</tr>
<tr>
<td>REM</td>
<td>Remote</td>
</tr>
<tr>
<td>RESULT</td>
<td>Resultative</td>
</tr>
<tr>
<td>SG</td>
<td>Singular</td>
</tr>
<tr>
<td>SIMUL</td>
<td>Simultaneous</td>
</tr>
<tr>
<td>SU</td>
<td>Subject</td>
</tr>
<tr>
<td>SUBJ</td>
<td>Subjunctive</td>
</tr>
<tr>
<td>SUPRL</td>
<td>Superlative</td>
</tr>
<tr>
<td>THEM</td>
<td>Thematic suffix</td>
</tr>
<tr>
<td>TRANSFORM</td>
<td>Transformative ('made of X')</td>
</tr>
<tr>
<td>TRANSL</td>
<td>Translative ('becoming X')</td>
</tr>
<tr>
<td>V</td>
<td>Verb</td>
</tr>
<tr>
<td>VERS</td>
<td>Version</td>
</tr>
<tr>
<td>VN</td>
<td>Verbal noun</td>
</tr>
<tr>
<td>1</td>
<td>First person</td>
</tr>
<tr>
<td>2</td>
<td>Second person</td>
</tr>
<tr>
<td>3</td>
<td>Third person</td>
</tr>
<tr>
<td>X→Y</td>
<td>Subject (agent) X acting on object (patient) Y</td>
</tr>
</tbody>
</table>
NOTES ON TRANSCRIPTION

Examples from languages of the U.S.S.R. that use the Latin alphabet (i.e. Estonian, Latvian, Lithuanian) are presented in the current orthography, for which see below. Russian examples are presented in accordance with the International system of transliteration (see p. 286). Unassimilated loans from Russian into other languages of the U.S.S.R. are also presented in accordance with the International system, and are set in block capitals to distinguish them from the general transcription system discussed below.

Although specialists in Turkic, Finno-Ugric, Iranian, etc., languages use their own transcription systems, we have opted here for a uniform system of broad transcription throughout the book. This system is based on that of the Association Phonétique Internationale (I.P.A.), but with some deviations for the following reasons: the I.P.A. system is not well adapted to transcribing affricates, some of which are extremely frequent in the languages of the U.S.S.R.; in some instances, absence of detailed phonetic information has led us to use a somewhat less specific transcription (e.g. in not distinguishing palatal from palatalised alveolar); we have used some symbols to bring the transliteration more into line with orthographies of the U.S.S.R. and specialist transcriptions of the language-families discussed. The following departures from I.P.A. should be noted.

More front values of vowels (i.e. front, or central relative to back) are indicated by a diaeresis (i.e. ꞻ = I.P.A. [y], and likewise Ꞻ for [u]). The I.P.A. central vowel symbols [i], [a] are used also for back-of-central rounded vowels where the backcentral opposition is nondistinctive. The symbol ꞻ represents a low back rounded vowel, and ꞻ a vowel between [o] and [u]. Vowel length is indicated by doubling the vowel-sign, stress by an acute accent on the stressed vowel. Ꞻ indicates a reduced vowel. Pharyngisation is indicated by a subscript, e.g. ꞻ. Vowels with retracted tongue root take a subscript point, e.g. ꞻ. Closing diphthongs have their second component represented as ꞻ or Ꞻ rather than ꟙ or ꟛ.

For consonants, we use ꟙ for I.P.A. [ʝ] and ꟙ for I.P.A. [ʒ]. Coronal affricates are indicated as follows: ꟙ for [ts], ꟙ for [dz], ꟙ for [tʃ], ꟙ for [dʒ]. Palatalisation is
Notes on transcription

indicated by an acute accent over or to the right of the consonant symbol (e.g. ķ, ũ); the same notation is used to indicate palatal or alveolo-palatal articulation where the difference among these is nondistinctive (e.g. ų for I.P.A. [ŋ], š for I.P.A. [ɛ], ķ for I.P.A. [k]). Labialisation is indicated by a superscript circle to the right of the consonant symbol (e.g. ř), and pharyngalisation by a subscript, as in ţ. Retroflex consonants are shown by a subscript point under the corresponding alveolar (e.g. ť for I.P.A. [t]). A macron over a consonant indicates an intensive (see p. 200), e.g. š. In Armenian and the Kurdish dialect spoken in Armenia, ţ indicates a more intense rhotic (usually trilled) than r (usually continuant). Other specific conventions are discussed in the text where they occur.

The Latin orthographies of the languages of the U.S.S.R. differ from the above system in the following ways. In Lithuanian, ė is mid [ɛ] while e is open [ä]; y and į are the long equivalents of i and u; ų, etymologically indicating nasalisation, now simply indicates a long vowel (thus Ĳ and y represent the same sound); closing diphthongs have i, u as their second component, ie and uo are opening diphthongs; for the representation of tone, see p. 148–51; the affricates [ʒ] and [ʐ] are represented as dz, dź respectively; consonants are automatically palatalised before front vowels, while before back vowels a nonpalatalised consonant is indicated as CV, a palatalised consonant C袆.

Latvian has a distinction between [ɛ] and [ä], long and short, though the orthography underdifferentiates qualitatively with the one symbol ē; ā is an opening diphthong [uo], except in some loans; long vowels are marked by a macron (e.g. ā); closing diphthongs have i, u as their second component; for the representation of tone, see p. 151; dz and dź represent the affricates [ʒ] and [ʐ] respectively; palatal consonants are indicated by a cedilla, i.e. ķ, ģ, ņ, ķ, ţ (the last not in the current Soviet orthography).

In Estonian, ņ represents I.P.A. [ũ]; closing diphthongs have their second component symbolised ĵ, ū; palatalisation is (rather marginally) phonemic for dentals/alveolars, though not represented in the orthography; the orthography also underdifferentiates between long and overlong segments (see pp. 115–16).
Map of the peoples of the U.S.S.R.
(Not shown on this map are peoples occupying a very small geographical area, e.g. users of North-East Caucasian languages restricted to a single village or group of villages, or peoples with extensive geographical distribution as a minority, e.g. Gypsies.)
## INDEX TO MAP

**Indo-European**

**SLAVONIC**
1. Russian
2. Ukrainian
3. Belorussian
4. Polish
5. Bulgarian

**BALTIC**
6. Lithuanian
7. Latvian

**GERMANIC**
8. German

**ROMANCE**
9. Moldavian
10. Romanian

**HELLENIC**
11. Greek

**IRANIAN**
12. Tadzhik
14. Beludzh
15. Kurdish
16. Tat
17. Ossete

**ARMENIAN**
18. Armenian

**Caucasian**
20. Georgian
21. Abkhaz
22. Abaza
23-4 Kabard-Cherkes:
23. Kabard
24. Cherkes
25. Adyge
26. Chechen
27. Ingush
28. Avar
29. Lak
30. Dargva
31. Tabasaran
32. Lezgi
33. Agul

**Uralic**
BALTO-FINNIC
34. Estonian
35. Karelian
36. Veps

LAPP
37. Lapp

PERMIC
38. Komi-(Zyryan)
39. Komi-Permyak

VOLGAIC
40. Udmurt

MORDVA
41. Mari
42. Mordva

UGRIC
43. Khanty
44. Mansi

HUNGARIAN
45. Hungarian

SAMI
46. Nenet
47. Nganasan
48. Selkup

**Altaic**

TURKIC
49. Chuvash
50. Tatar
51. Bashkir
52. Nogay
53. Kumyk

KARACHAY-BALKAR:
54-5 Karachay-Balkar:

54. Karachay

55. Balkar
56. Kazakh
57. Kirgiz
58. Karakalpak
59. Uzbek
60. Uyghur
61. Turkmen
62. Azerbaydzhan
63. Gagauz
64. Altay
65. Khakass
66. Tuva
67. Shor
68. Yakut

MONGOLIAN
69. Buryat
70. Kalmyk

TUNGUSIC
71. Evenki
72. Even
73. Nanay
74. Ulch
75. Oroch
76. Udege

**Paleosiberian** (including Eskimo-Aleut)
77. Chukchi
78. Koryak
79. Helmen
80. Yukagir
81. Nivkh
82. Eskimo (Siberian Yupik)
83. Aleut
84. Ket

**KOREAN**
85. Korean