



CONCISE ENCYCLOPEDIA OF LANGUAGES OF THE WORLD

KEITH BROWN • SARAH OGILVIE



CONCISE ENCYCLOPEDIA OF
LANGUAGES
OF THE WORLD

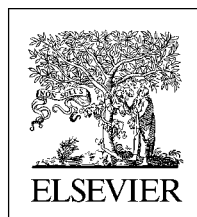
CONCISE ENCYCLOPEDIA OF LANGUAGES OF THE WORLD

COORDINATING EDITOR

KEITH BROWN
University of Cambridge
Cambridge, UK

CO-EDITOR

SARAH OGILVIE
University of Oxford
Oxford, UK



Amsterdam Boston Heidelberg London New York Oxford
Paris San Diego San Francisco Singapore Sydney Tokyo

Elsevier Ltd., The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK

© 2009 Elsevier Ltd.

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic, or mechanical, including photocopy, recording, or any information storage and retrieval system, without permission in writing from the publishers.

Permissions may be sought directly from Elsevier's Rights Department in Oxford, UK: phone (+44) 1865 843830; fax (+44) 1865 853333; e-mail permissions@elsevier.com.

Requests may also be completed online via the homepage (<http://www.elsevier.com/locate/permissions>).

First edition 2009

Library of Congress Control Number: 2008934269

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

ISBN 978-0-08-087774-7

09 10 11 12 13 10 9 8 7 6 5 4 3 2 1

This book is printed on acid-free paper
Printed and bound in China

Cover image: Adapted from *Orbis Terrarum Nova* (1594) by Petro Plancio

Working together to grow
libraries in developing countries

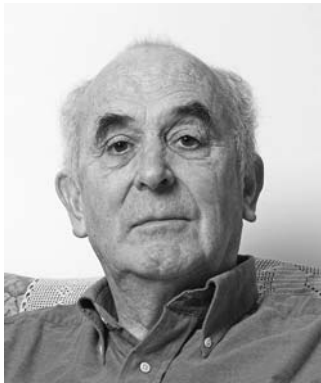
www.elsevier.com | www.bookaid.org | www.sabre.org

ELSEVIER

BOOK AID
International

Sabre Foundation

THE EDITORS



Keith Brown was Editor-in-Chief of the second edition of the *Encyclopedia of Language and Linguistics* (Elsevier, 2006). He is now an Associate Lecturer in the Faculty of English at Cambridge. From 2007 he has been President of the Philological Society. From 1990 to 1994 he was President of the Linguistics Association of Great Britain, and he has been a Member of Council of the Philological Society since 1998. He is author of *Linguistics Today* (Fontana, 1984) and co-author, with Jim Miller, of *Syntax: A Linguistic Introduction to Sentence Structure* and *Syntax: Generative Grammar* (Hutchinson, 1981). Keith was joint editor of *Concise Encyclopedia of Linguistic Theories* and *Concise Encyclopedia of Grammatical Categories* (Pergamon Press, 1997 and 1998), *Common Denominators in Art and Science* (Aberdeen University Press, 1983) and *Language, Reasoning and Inference* (Academic Press, 1986).



Sarah Ogilvie, Trinity College, Oxford, is a linguist and lexicographer who specializes in words that enter English from non-European languages. She was Languages of the World section editor of the second edition of the *Encyclopedia of Language and Linguistics* (Elsevier, 2006), a former editor of the *Oxford English Dictionary*, and was Etymologies Editor of the *Shorter Oxford Dictionary* (6th ed., 2007).

Albanian

B Demiraj, University of Munich, Munich, Germany
A Esposito, Oxford English Dictionary, Oxford, UK

© 2006 Elsevier Ltd. All rights reserved.

Linguistic Type

Albanian constitutes a single branch of the Indo-European family of languages. It is often held to be related to Illyrian, a poorly attested language spoken in the western Balkans in classical times, but this has not yet been proved conclusively. Although as a people the Albanians have been known since the 2nd century A.D., the earliest surviving records of the Albanian language date only from the 15th century. In its grammar Albanian displays several characteristic features of Indo-European languages, such as declension of nouns by means of case endings and conjugation of verbs by means of personal endings; in its lexicon it preserves a considerable number of words of inherited Indo-European stock.

Albanian may further be characterized as a member of the Balkan Sprachbund. During the many centuries of their evolution the languages of the Balkans (several languages not directly related and belonging to different branches of Indo-European) have come to share certain linguistic features with each other that they do not share with other non-Balkan languages to which they are ostensibly more closely related. Albanian displays several of these features, for example: postposition of the definite article, analytic formation of the future tense (in Albanian with the semiauxiliary verb *dua* 'to want' in the fossil form *do*), substitution of the infinitive by subjunctive clauses, pronominal doubling of objects.

In addition to features shared respectively with other Indo-European languages and with other Balkan languages, Albanian also displays several innovative features, in phonology, in morphosyntax, and in lexis, which mark it out from other European languages.

The phonemic inventory of standard Albanian comprises 7 vowels and 29 consonants, and is remarkable for the way that phonetically similar consonants (including plosives, affricates, fricatives, and liquids) have formed phonemic pairs. The phonological system also reveals the operation of umlaut in former times (with which compare the Germanic languages). As regards morphosyntactic structure, may be mentioned the development, alongside the postpositive definite article, of a proclitic article with indefinite function, which, in turn, has given rise to further innovations: the creation of a special class of adjectives and the reformation of ordinal numerals and of the genitive case. Another important

innovation is the development of the admirative mood in the verbal system, used to express surprise, disagreement, etc.

Present-day Albanian may be categorized as a partly synthetic, partly analytic language, which, alongside synthetic features (both inherited and innovatory), has also developed several analytic features, such as the formation of the perfect and future tenses with auxiliary verbs and the frequent use of prepositions with inflected forms of nouns and pronouns.

The vocabulary of Albanian is notable for the high level of borrowing it shows from different neighboring and influential languages over the course of many centuries, for example: ancient Greek and Latin, the Slavic languages of the Balkans, Turkish, medieval and modern Greek, and (in our own times) French, Italian, and English.

Geographic Spread

Today Albanian is spoken by a population of about 6 500 000 native speakers in a compact ethno-linguistic area in the western Balkans, which comprises:

1. Albania;
2. almost the whole of Kosovo;
3. a broad band of northwestern Macedonia (the former Yugoslav republic) from Kumanovo to Struga;
4. the districts of Medveda, Preševo, and Bujanovac in southern Serbia;
5. the southern and southwestern part of Montenegro;
6. the region of Chameria in northwestern Greece.

Albanian is the official language of the Republic of Albania, and one of the official languages of Kosovo (U.N. administration) and the Republic of Macedonia; it is a national minority language in the Republic of Montenegro.

Outside this compact ethno-linguistic area Albanian is also spoken today in a considerable number of linguistic pockets in the Balkans and beyond. These have arisen as a result of continuing economic and political migrations over the last 700 years. The descendants of the earliest attested diaspora of Albanian-speakers live in scattered communities in southern Greece (the Peloponnese, Attica, and the Aegean islands); the original migration dates from the 14th and 15th centuries, and its cause appears to have been chiefly economic (see Jochalas, 1971). Further scattered communities of Albanian-speakers are to be found in southern Italy and Sicily, where their ancestors settled during the 15th and 16th centuries for political and religious reasons after the

occupation of the western and southern Balkans by the Ottoman Turks. The exact number of Albanian-speakers in these linguistic pockets is difficult to determine, as many of them, especially the younger generation, have abandoned their ancestral language, and speak Greek or Italian, respectively. Those who still retain Albanian (all of whom are bilingual) speak an archaic variety heavily influenced by the superstrate language.

Other linguistic pockets, which, however, are now in danger of being completely assimilated, exist in Serbia (the Sanjak), Croatia (Zadar), central Macedonia, south-eastern Bulgaria (Mandrica), Turkey, and the Ukraine.

During the 20th century emigration of Albanian speakers has continued, especially at the beginning and end of the century from Albania to the United States, Canada, Italy, Greece, and the United Kingdom, and from Yugoslavia (and its successor states) and northern Greece to Turkey, Germany, Switzerland, and Sweden.

Dialects

Within the compact ethno-linguistic area in the western and central Balkans, Albanian is spoken in two main dialects, Gheg and Tosk, each of which may be further divided into several subvarieties. The River Shkumbin in central Albania historically forms the boundary between these two dialects, with the population to the north speaking varieties of Gheg and the population to the south varieties of Tosk (see Gjinari, 1989).

Gheg and Tosk are distinguished from one another chiefly by several important phonological developments. For example, in Tosk /a/ before a nasal has become a central vowel (schwa), and intervocalic /n/ has become /r/. These two sound changes have affected only the old pre-Slav stratum of the Albanian lexicon, that is, native words and loanwords from ancient Greek and Latin. The only important dialectal difference in grammatical structure is the loss of the infinitive in Tosk, in which constructions with the subjunctive predominate just as in all other Balkan languages (with the exception of Serbian and Croatian). However, these innovations, as those that are also evident in different varieties of Gheg, are not such as to impede communication between speakers of the two dialects. Furthermore, the major part of the Albanian lexicon is common to the two dialects.

Of the two main varieties of Albanian spoken outside the ethno-linguistic area, Arvanitika (spoken by the descendants of the ancient migration to Greece) and Arbëresh (spoken by the descendants of the ancient migration to Italy), both preserve archaic features characteristic of varieties of southern Tosk. (The majority of emigrants in these historical

migrations were from southern Albania.) The archaic dialectal features and the separate development of these varieties under the powerful influence of superstrate languages (Greek and Italian) make communication between speakers of the diaspora and those of the ethno-linguistic homeland almost impossible. This differentiation, conditioned by time and space, has caused several specialists to treat these varieties as separate languages (see Sasse, 1991).

Overlying the dialectal diversity of Albanian are different religious (Catholic, Orthodox, Muslim), cultural, and political allegiances that over time have also greatly influenced linguistic developments.

Codification

Up until the early 20th century Albanian was written in a variety of scripts (Roman, Greek, Arabic, Cyrillic), depending on local influences. In 1908 the Congress of Monastir decided on the adoption of the Roman alphabet. The use of Albanian as an official language first became possible after the proclamation of independence of Albania in 1912. However, the emergence of an agreed standard language took time; competing local standards continued to be used until well into the second half of the 20th century. Modern standard Albanian (largely Tosk-based), which is today the accepted standard throughout the whole ethno-linguistic area, did not gain its final sanctioning until 1972 at the Orthographic Congress of Tirana, organized by the Albanian Academy of Sciences, in which linguists and writers from Yugoslavia and the Albanian diaspora also participated.

Present and Future Trends

The decade of the 1990s saw great upheavals in the western Balkans (the fall of communism in Albania, the dismemberment of Yugoslavia, and the war in Kosovo) that radically affected the lives of Albanian speakers. One consequence has been a dramatic increase in the influence of foreign languages on Albanian. A flood of loanwords, especially from English and Italian, is pouring into both the colloquial and the standard language. There exists an unofficial movement opposed to the use of 'unnecessary' foreign words, but attempts to engage the interest of the state in support of its efforts have so far proved unsuccessful.

Bibliography

Camaj M (1984). *Albanian grammar*. Wiesbaden: Otto Harrassowitz.

- Demiraj S (1988). *Gjuha shqipe dhe historia e saj*. Tirana: Shtëpia botuese e librit universitar.
- Demiraj S (1998). 'Albanian.' In Giacalone Ramat A & Ramat P (eds.) *The Indo-European languages*. London: Routledge.
- Gjinari J (1971). 'De la continuation de l'Illyrien en Albanais.' In Korkuti M, Anamali S & Gjinari J (eds.) *Les Illyriens et la genèse des Albanais*. Tirana: Université de Tirana. 173–181.
- Gjinari J (1989). *Dialektet e gjuhës shqipe*. Tirana: Akademia e shkencave të republikës popullore socialiste të Shqipërisë.
- Hamp E P (1972). 'Albanian.' In Sebeok T A (ed.) *Current trends in linguistics*, 9. The Hague: Mouton. 1626–1692.
- Jochalas T (1971). 'Über die Einwanderung der Albaner in Griechenland; eine zusammenfassende Betrachtung.' In *Dissertationes Albanicae: in honorem Josephi Valentini et Ernesti Koliqi septuagenariorum*. Munich: Trofenik.
- Hetzer A (1985). *Geschichte des Buchhandels in Albanien*. Wiesbaden: Otto Harrassowitz.
- Hetzer A (1996). 'Die vereinheitliche albanische Schriftsprache.' In Schaller H (ed.) *Sprache und Politik: die Balkansprachen in Vergangenheit und Gegenwart*. Munich: Südosteuropa-Gesellschaft.
- Newmark L (1998). *Albanian–English dictionary*. Oxford: Oxford University Press.
- Raka F (1997). *Historia e shqipës letrare*. Prishtina: Rilindja.
- Sasse H J (1991). *Arvanitika*. Wiesbaden: Otto Harrassowitz.

Algonquian and Ritwan Languages

D H Pentland, University of Manitoba, Winnipeg, MB, Canada

© 2006 Elsevier Ltd. All rights reserved.

More than 30 languages of the Algonquian family were formerly spoken along the east coast of North America from about 34°N (Cape Fear, North Carolina) to about 56°N (Davis Inlet, Labrador), around the upper Great Lakes, and west to the foothills of the Rocky Mountains. They were the first North American languages encountered by French and English explorers; by the end of the 17th century several languages had already been described in detail. Three centuries later, however, two-thirds of the languages are no longer spoken, with only English loanwords such as *moccasin*, *skunk*, and *squaw* to reflect their former existence. The 'Ritwan' languages (Wiyot and Yurok) of California are distantly related. Pilling (1891) provides a nearly exhaustive inventory of the earlier sources; later publications are listed by Pentland and Wolfart (1982), but the only comprehensive bibliography of the most recent literature is in Nichols (1981–).

Classification

The only widely accepted genetic subgroup within the Algonquian family is Eastern Algonquian, consisting of the languages which descended from Proto-Eastern Algonquian (Goddard, 1978b). It includes the languages of the Maritime provinces, southern Quebec, and the northern New England states – Micmac (several dialects), Malecite-Passamaquoddy, Etchemin, Eastern and Western Abnaki (two languages, each

with several dialects), and Pocumtuck or 'Loup B' – and those formerly spoken in the Hudson and Delaware River basins of New York, Pennsylvania, and New Jersey – two dialects of Mahican, and the two 'Delaware' languages, Munsee (including the divergent Wappinger dialect) and Unami (three dialects). The languages of southern New England and Long Island – Nipmuck ('Loup A'), Massachusetts (Wampanoag), Narragansett, Pequot-Mohegan-Montauk, and Quiripi-Unquachog – and those of the southeastern states – Nanticoke, Conoy (Piscataway), Powhatan (Virginia Algonquian), and Roanoke-Pamlico (Carolina Algonquian) – may also be part of the Eastern subgroup, but since all are extinct, the crucial phonological details depend on interpretations of early written records.

The so-called 'Central' languages were located between Hudson Bay and the Ohio River valley; each shares many features with its neighbors, but there are no ancient subdivisions.

Cree-Montagnais-Naskapi is a dialect chain extending across central Canada from Labrador to Alberta, conventionally subdivided according to the reflex of Proto-Algonquian **l*: Plains Cree (Nêhiyawêwin), the dialect with *y* < **l*, in Alberta and Saskatchewan; three varieties of Woods Cree (with *ð*) in northern Saskatchewan and Manitoba, one of which probably continues the extinct Missinipi dialect (with *r*; cf. Pentland, 2003); three or more varieties of Swampy Cree (with *n*) in Manitoba and northern Ontario; Moose Cree (with *l*) on the southwest coast of James Bay; and Atikamekw (or Tête de Boule, with *r*), in southwestern Quebec, cut off from the others by a dialect of Ojibwa. In the eastern dialects Proto-Algonquian **k* has