

CONCISE

COMPENDIUM
OF THE
WORLD'S
LANGUAGES

•

GEORGE L. CAMPBELL

ROUTLEDGE

CONCISE COMPENDIUM OF THE WORLD'S LANGUAGES

BY
GEORGE L. CAMPBELL

江苏工业学院图书馆
藏书章



LONDON AND NEW YORK

First published 1995
by Routledge
11 New Fetter Lane, London EC4P 4EE

Simultaneously published in the USA and Canada
by Routledge
29 West 35th Street, New York, NY 10001

First published in paperback 1998

© 1995 George L. Campbell

Printed and bound in Great Britain by
T J International Ltd, Padstow, Cornwall

All rights reserved. No part of this book may be reprinted or reproduced or utilized in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

British Library Cataloguing in Publication Data

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

Library of Congress Cataloguing in Publication Data

A catalogue record for this book is available from the Library of Congress

ISBN 0-415-11392-X (hbk)

ISBN 0-415-16049-9 (pbk)

INTRODUCTION

The *Concise Compendium of the World's Languages* derives from the two-volume *Compendium of the World's Languages*, published by Routledge in 1991. 'Concise' in the title refers to the reduced number of languages here covered – about a hundred as compared with some 300 in the larger work. Many of the articles selected for inclusion in the present book have, in fact, been substantially amplified. In the concise edition the purpose remains the same: to provide brief descriptions, phonological, morphological, lexical snapshots, in non-technical language, of a fairly wide cross-section of contemporary natural-language systems.

Articles were selected in the light of three criteria:

1. Number of speakers: most languages spoken at present by at least 5–10 million people are included.
2. Socio-political interest and importance: under this heading, certain minor languages have been included whose speakers have, for whatever reason, acquired a heightened political profile and topical relevance: Basque, Macedonian, Kurdish, and Zulu have been chosen simply as representatives of this class, certainly not as its only members.
3. Initially, it had been hoped to illustrate most of the world's language families by at least one member, but constraints of space ruled this out. However, space has been found for four minor languages of great intrinsic interest: Navajo, Lappish, Nama, and Nivkh.

By the same rationale, dead languages have been excluded. Here, the retention of Classical Chinese (Wenli) and Sanskrit may appear anomalous. The explanation is that both of these are, in a sense, living languages. Classical Chinese on two counts: the language of the Yuan drama and the Ming/Qing novel does not differ greatly from literary Modern Standard Chinese; and by virtue of the Chinese script even Tang poetry and the Confucian classics are on visual hold, as it were, and readily accessible to literate Chinese of latter-day generations. As for Sanskrit, it has consistently acted, and continues to act, as an inexhaustible reservoir of living tissue for the new Indo-Aryan languages, by supplying them with straight implants (*tatsama* 'just as it was', scil. in Sanskrit) or its genetic progeny (*tadbhava* 'derived from that', scil. from Sanskrit).

As a result of the greatly reduced coverage, much of the Bibliography which rounded off the *Compendium* became irrelevant, and has been discarded. The remainder has been re-arranged by language. As in the *Compendium*, verses 1–8 of the first chapter of St John's Gospel have been appended as an illustrative example to most articles. Where no translation of this gospel is available, an

alternative passage from the New Testament has been substituted and identified. Alone in the present book, Nivkh, which was unwritten until the 1930s, has no illustrative text.

The relevant script tables are collected at the end of the book.

In conclusion, I would like to express my gratitude to Simon Bell and Claire Trocmé for their invaluable help in the production of this book.

ABBREVIATIONS

abl.	ablative	fam.	familiar
abs.	absolute	fem.	feminine
acc.	accusative	Fin.	Finnish
adess.	adessive	Fr.	French
adit.	aditive	fut.	future
affirm.	affirmative		
Afr.	Afrikaans	gen.	genitive
anim.	animate	Gk	Greek
Ar.	Arabic	Gm.	German
Assyr.	Assyrian		
aux.	auxiliary	hon.	honorific
Av.	Avestan		
Azer.	Azerbaijani	IE	Indo-European
		illat.	illative
Bel.	Belorussian	imper.	imperative
BI	Bahasa Indonesia	imperf.	imperfect
BP	Brazilian Portuguese	impers.	impersonal
		inanim.	inanimate
Ch.	Chinese	incl.	inclusive
Ch.Ap.	Chiricahua Apache	Ind.	Indonesian
Chip.	Chipewyan	indef.	indefinite
cl.	class	iness.	inessive
CL	Classical Latin	instr.	instrumental
com.	comitative	It.	Italian
comp.	comparative		
conj.	conjunctive	Jap.	Japanese
D	Dutch	K	Kramā (Javanese)
dat.	dative		
def.	definite	Lat.	Latin
dep.	dependent	lk	link
		loc.	locative
EArm.	Eastern Armenian	masc.	masculine
EP	European Portuguese	ME	Middle English
erg.	ergative	MF	Middle French
ESlav.	Eastern Slavonic	MSC	Modern Standard
excl.	exclusive		Chinese

ABBREVIATIONS

N	Ngoko (Javanese)	pron.	pronoun
Nav.	Navajo	prox.	proximate
neg.	negative		
neut.	neuter	refl.	reflexive
nom.	nominative	rel.	relative
		Romy	Romany
O/obj.	object	Russ.	Russian
obl.	oblique		
obv.	obviative	S/sbj.	subject
OE	Old English	SC	Serbo-Croat
OF	Old French	sing.	singular
OI	Old Irish	Skt	Sanskrit
Old Ch.	Old Church	Sl.	Slovene
Slav.	Slavonic	Slav.	Slavonic
opt.	optative	Som.	Somali
Osc.	Oscan	Sp.	Spanish
		SSlav.	Southern Slavonic
p.	person	subj.	subjunctive
part.	partitive		
pass.	passive	trans.	transitive
perf.	perfective		
pl.	plural	Ukr.	Ukrainian
Pol.	Polynesian		
poss.	possessive	V	verb
pp.	past participle	Viet.	Vietnamese
prep.	preposition(al)	VN	verbal noun
pres.	present	voc.	vocative
pret.	preterite		
prog.	progressive	WArm.	Western Armenian

CONTENTS

Introduction	vii	Guaraní	204
Abbreviations	ix	Gujarati	208
Afrikaans	1	Hausa	213
Akan or Twi	6	Hebrew	219
Albanian	11	Hindi	225
Amharic	18	Hungarian	230
Arabic	24	Icelandic	236
Armenian, Modern Standard	33	Igbo	242
Assamese	39	Indo-European Languages	247
Balinese	44	Indonesian	249
Baltic Languages	48	Italian	254
Bantu Languages	49	Japanese, Modern Standard	260
Basque	58	Javanese	267
Belorussian	64	Kanarese <i>see</i> Kannada	
Bengali (<i>bānlā</i>)	69	Kannada (Kanarese)	272
Berber	75	Kashmiri	277
Breton	80	Khmer <i>see</i> Cambodian	
Buginese	87	Korean	282
Bulgarian	92	Kurdish	288
Burmese	98	Lahndā	293
Cambodian (Khmer)	103	Lao	298
Catalan	108	Lappish	303
Chinese, Classical (Wenli)	113	Latvian	309
Chinese, Modern Standard	118	Lithuanian	314
Czech	124	Macedonian	319
Danish	130	Malayalam	324
Dravidian Languages	135	Maori	329
Dutch	137	Mapudungu	333
English	143	Marathi	339
Estonian	157	Mongolian, Modern	345
Finnish	163	Nama	350
French	170	Navajo	354
Fulani (Fulbe)	178	Nepali	360
Fulbe <i>see</i> Fulani		Nivkh	365
Georgian	183	Norwegian	369
German	189	Panjabi	374
Greek, Modern Standard	197	Pashto	379
		Persian	384

CONTENTS

Polish	389	Tagalog	512
Polynesian Languages	396	Tamil	518
Portuguese	401	Telugu	525
Quechua	407	Thai	531
Romanian	413	Tibetan	536
Romany	419	Turkic Languages	543
Russian	425	Turkish	547
Samoan	435	Twi <i>see</i> Akan	
Sanskrit	440	Ukrainian	554
Semitic Languages	448	Uralic Languages	560
Serbo-Croat	454	Urdu	562
Sindhi	461	Uzbek	563
Sinhala <i>see</i> Sinhalese		Vietnamese	567
Sinhalese (Sinhala)	465	Welsh	573
Slavonic Languages	470	Wenli <i>see</i> Chinese, Classical	
Slovak	474	Yiddish	582
Slovene	478	Yoruba	588
Somali	484	Zulu	593
Spanish	490		
Sundanese	497		
Swahili	501	Appendix of Scripts	601
Swedish	507	Bibliography	655

AFRIKAANS

INTRODUCTION

Afrikaans belongs to the West Germanic branch of Indo-European, and is derived from the same sixteenth-century Dutch dialect, Frankish in origin, which underlies modern Dutch. It took shape first in Cape Colony, where Jan van Riebeeck had arrived in 1652, and spread to the rest of South Africa, from the seventeenth century onwards. It has been a literary language for a little over a century. On both counts, it is the youngest of the Germanic languages. At present, it is spoken by about 4 million people.

The language was originally known as Kaaps-Hollands or Plat-Hollands. The designation 'Afrikaans' was adopted towards the end of the nineteenth century. From 1910 till 1925, Dutch and English were the joint official languages of the Union of South Africa; in 1925 Afrikaans replaced Dutch. The use of Afrikaans is mainly characteristic of the Cape Province, the Orange Free State, and the Transvaal.

SCRIPT

Roman, 26 letters as in English; *c, q, x, z*, are rarely used. The circumflex is used to mark the long open *ê, ô*: *lê* = /lɛ:/.

PHONOLOGY

Consonants

stops: *p, b, t, d, k, g, ʔ*; palatalized *k'*;

affricates: *ts, tʃ*

fricatives: *f, v, s, ʃ, j, x, h*

nasals: *m, n, ŋ, ŋ*

lateral and flap: *l, r*

/z/ occurs in a few loan-words.

[*g*] and [*x*] are allophones: cf. *berg* 'mountain', /bɛrx/, pl. *berge*, /bɛrgə/. /*p, t, k*/ are non-aspirate, /*b, d*/ in final position are unvoiced → [*p, t*]. Dutch /*sx-*/ = Afrikaans /*sk-*/, and Dutch final *-t* drops: e.g. *nacht* > *nag* /nax/ 'night'; *nest* > *nes*, 'nest'.

Vowels

front: i, i:, e, ɛ, ɛ:, y, y:

middle: ə, ʌ:, a, ʊ, ʊ:

back: ɑ, ɔ, ɔ:, o:, u, u:, ʊ

/e:/ is realized as [e(:)ə] or [ɪə]; /o:/ as [o(:)ə] or [ʊə]. /a, ɛ, ɔ/ are nasalized when followed by *n* + fricative: e.g. *mense* [mɛ̃:ɪsə], 'people'; *aangesig* [ā:xəsɪx], 'face'.

DIPHTHONGS

Simple (short) or lengthened; all are falling.

short: əy, ʊʊ, ɛy (= Dutch /ʌy/), ai, ɔi, ui

long: a:i, o:i/ɔ:i, e:u

Stress

Stress is free, associated with pitch. The main stress is normally on the root. In separable verbs, however, the stress is transferred to the prefix.

MORPHOLOGY AND SYNTAX

The division into common and neuter nouns, retained in Dutch, has been lost in Afrikaans. A single definite article – *die* – applies to all nouns, singular and plural: e.g. *die vader* 'the father' – *die moeder* 'the mother' – *die kinders* 'the children'. 'n = [ə] is used as singular indefinite article: e.g. 'n *vliegtuig* 'a plane'.

NUMBER

-e is a frequent plural marker, with phonetic adjustment, where necessary, at juncture; -s and -ers are also used: e.g. *wolf* – *wolwe*; *skip* 'ship' – *skepe*; *dag* 'day' – *dae*; *oom* 'uncle' – *ooms*; *kind* 'child' – *kinders*.

CASE

There is no inflection for case in Afrikaans; the genitive relationship is expressed periphrastically with the particle *se*: e.g. *sy vader se huis* 'his father's house'; *Pretoria se koerante* 'Pretoria's newspapers, the Pretoria papers'.

The indirect objective case of a personal noun is marked by the preposition *vir* 'for, to': *gee ... vir die arme kind* 'give the poor child ...'; *Sê vir oom, hy moet tuis kom* 'Tell uncle to come home.' The use of *vir* with a direct personal object seems to be possible. Mironov (1969) gives the example: *Jan slaan vir Piet* 'Jan strikes Piet.'

Adjective

As attribute, adjective precedes noun. After either article, -e is normally added to all polysyllabic and some monosyllabic adjectives: e.g. *die Nederlandse taal*

'the Dutch language'. The addition of *-e* may induce change in stem final: e.g. *Die kind is goed*; but, *Dit is 'n goeie kind* 'This/he/she is a good child.' Many adjectives remain uninflected: e.g. *Dit is 'n donker nag* 'It's a dark night.'

COMPARATIVE

Made with *-er*: phonetic change at junctures: e.g. *donker* 'dark' – *donkerder*; *doof* 'deaf' – *dower*.

Pronoun

Some vestiges of common Germanic inflection survive in the singular forms: the third person singular forms are marked for gender:

Singular				Plural	
	Subject	Object	Possessive adjective	Subject/object	Possessive adjective
1	ek	my	my	ons	onse
2	jy	jou	jou	julle	julle se
	U	U	U	U	U
3	masc. hy	hom	sy		
	fem. sy	haar	haar	hulle	hulle se
	nt. dit	dit	—		

The predicative possessive forms are *myne*, *joune*, *U sy'n/s'n*, *syne*, *haar se*; pl. *ons sy'n*, *julle sy'n*, *hulle sy'n*: e.g. *my boek – die boek is myne – die boek is julle s'n*. *U* is the polite address form.

DEMONSTRATIVE PRONOUN/ADJECTIVE

e.g. *hierdie* 'this', *daardie* 'that'.

INTERROGATIVE PRONOUN

wie 'who?'; *wat* 'what?': e.g. *In wie se naam?* 'In whose name?'

RELATIVE PRONOUN

wat is used for all referents: e.g. *die boek wat daar lê* 'the book that is lying there'; *die man/die vrou wat daar woon* 'the man/woman who lives there'. 'Whose' is rendered as *wie se*: *die man wie se huis verkoop sal word* 'the man whose house is to be sold'.

Numerals

As in Dutch with some changes in spelling: e.g. D *vijf* 'five' = Afr. *vyf*.

Verb

Apart from the *ge-* prefix on most past participles, all verbal inflection has been lost. The verbal base is identical with the present tense of the indicative mood: e.g. *val* 'to fall', *ek/hulle val* 'I/they fall'. Auxiliaries such as *het*, *is*, *sal*, *word* are

used to make composite tenses. The typical Germanic past-tense forms, whether strong or weak (*see German*), have disappeared, leaving the composite form with *het* as the sole past tense in Afrikaans, apart from *was*, the past tense of *is* (cf. *Yiddish*): e.g. *ek het geval* 'I fell/have fallen'. If the verb begins with an inseparable prefix, e.g. *be-*, *er-*, *her-*, *ver-*, etc., the *ge-* of the past participle is dropped; i.e. here, the past participle coincides with the stem and the present tense.

SEPARABLE VERBS

The order of components in the tense structure of verbs with separable prefixes may be illustrated as follows, for *saambring* 'to bring along with one':

present: *ek bring my broeder saam* 'I'm bringing my brother along with me'
 past: *ek het my broeder saam.ge.bring* 'I brought/I've brought ...'
 future: *ek sal my broeder saam.bring* 'I shall bring ...'
 split infinitive: *om saam te bring* '(so as) to bring along with one'

PASSIVE

The model auxiliary *word* is used in the present and the future:

present: *die huis word gebou* 'the house is being built'
 past: *die huis is gebou* 'the house was built'
 past anterior: *die huis was gebou* 'the house had been built'
 future: *die huis sal gebou word* 'the house will be built'

IMPERATIVE MOOD

A simple imperative form is once again identical with stem and present tense; and, as in the present tense, a separable prefix follows: e.g. *oppas* 'to take care', *Pas op!* 'Take care!' A polite hortative form can be made with the auxiliary *moet*: e.g. *jy moet weggaan* 'you should/must go'. *Moet + nie* → *moenie* provides a negative imperative: e.g. *Moenie hier staan nie!* 'Don't stand here!'

NEGATIVE

The negative particle *nie* is recapitulated in composite verbal forms: e.g. *hy het nie gekom nie* 'he didn't come'; *hy het niks gedoen nie* 'he didn't do anything'.

NON-FINITE FORMS

The infinitive form is heavily eroded; e.g. *hê* (< *heb* < *hebben*) 'to have'; *sê* (< *seg* < *seggen*) 'to say'.

The present participle retains its original Germanic form: e.g. *lesende* 'reading'; *vallende* 'falling'; *Al pratende het hulle uit die kamer gestap* 'Talking, they left the room.'

Past participle: *ge-* prefix, except in verbs with an inseparable prefix; cf.

Hy het die boek gelees 'He read/has read the book'
Hy het die boek vertaal 'He translated/has translated the book'

Prepositions

Prepositions require the objective form of personal pronoun: e.g. *met haar* 'with

her'; *Ek hou nie van hom nie* 'I don't like him' (lit. 'hold not with him'); *Ek wil môre graag by jou kom kuier* 'I'd like to come and visit you tomorrow' (*graag* = Gm. *gerne*)

Word order

SVO in principal clause; in subordinate clause SOV; if the subordinate clause precedes the principal clause, the word order in the latter is VSO.

Subordinate clause, past tense: OSV aux.: e.g. *die motorkar wat ek gister gekoop het* 'the car I bought yesterday'.

Subordinate clause, future tense: OS aux. V: e.g. *die brief wat ek môre sal ontvang* 'the letter which I'll get tomorrow'.

If a modal auxiliary is present, the order is SO aux. modal aux. V: e.g. (*Hy het gesê dat*) *hy ons môre sou kan help* ('He said that) he would be able to help us tomorrow.'

- 1 IN die begin was die Woord, en die Woord was by God, en die Woord was God.
- 2 Hy was in die begin by God.
- 3 Alle dinge het deur Hom ontstaan, en sonder Hom het nie een ding ontstaan wat ontstaan het nie.
- 4 In Hom was lewe, en die lewe was die lig van die mense.
- 5 En die lig skyn in die duisternis, en die duisternis het dit nie oorweldig nie.
- 6 Daar was 'n man van God gestuur, wie se naam Johannes was.
- 7 Hy het tot 'n getuienis gekom om van die lig te getuig, sodat almal deur hom sou glo.
- 8 Hy was nie die lig nie, maar hy moes van die lig getuig.

AKAN or TWI

INTRODUCTION

This language is spoken in two major dialects, Ashanti and Fante, by up to 8 million people in Ghana. Akan belongs to the Kwa group of the Niger-Congo family. The form described here is Ashanti, the dialect spoken by the great majority of the Akan-speaking population.

The Basler Evangelische Missionsgesellschaft undertook the task of creating a Twi literary language in the 1840s, and the next 40 years saw the appearance of a Twi Bible, a definitive grammar, and a dictionary, mostly the work of J.G. Christaller. This early literary activity was based on the minor but politically and geographically accessible Akuapem dialect. Through the first half of the twentieth century, however, Ashanti began to be recognized as the proper base for a national literary language. Though certain scholars (e.g. Danquah) continued to write in Akuapem, others like Nketia and Tabi switched to Ashanti. In 1968 A.C. Denteh launched the important literary periodical *Odawuru*.

SCRIPT

Roman plus *ɛ*, *ɔ*. In 1961 the Bureau of Ghana Languages devised a common standardized script for all forms of Akan. Tone is not normally marked, nor is nasalization. The correspondence between the script and the actual sounds of Akan is rather weak.

PHONOLOGY

Consonants

stops: p, b, t, d, k, g, ɔ
affricates: tʃ, dʒ
fricatives: f, ɸ, s, ʃ (ɕ), h
nasals: m, n, ɲ, ŋ
trill: r

Note: all consonants tend to be palatalized before front vowels, and stops tend to affricates.

There is a labialized series: /k^o, g^o, h^o, n^o, ŋ^o/; /v/ and /l/ occur in foreign words. /d/ is notated as *d̥*; /ɸ/ as *w*.

Vowels

i, ɪ, e, ɛ, æ, ǣ, ɔ, o, ʊ, u, y

/e/ and /ǣ/ have no nasalised counterparts; all other vowels take nasalization.

VOWEL HARMONY

lax series: ɪ, ɛ, ɑ, o, ʊ[˘]

tense series: i, e, a, o, u

This division is important for vowel harmony: briefly, a *lax* vowel followed by /i/, /a/, or /u/ is promoted to the next-highest vowel in the *tense* series, e.g. the sequence /ɪ...ʊ/ → /i...u/.

Vowel harmony determines the vocalic structure of the possessive and subject pronouns.

Tones

Three tones, which are phonemic: high, middle, and low. The tones in Akan are characteristically ‘terraced’, i.e. successive highs begin on a slightly lower level. This in turn affects successive middle tones, but low pitch is not affected.

MORPHOLOGY AND SYNTAX

Noun

No grammatical gender. Nouns are uninflected, though tonal patterns vary, depending on whether the noun is isolate, in construct (e.g. preceded by possessive pronoun), or in a compound.

There are several ways of making a singular noun plural. *-nóm* is a pluralizing suffix: *ɔyére* ‘wife’, pl. *ɔyérenóm*.

Pluralizing prefixes:

m-: *ɔba* ‘child’, pl. *mma*;

a-: *ɔkérāmāñ* ‘dog’, pl. *akérāmāñ*;

n-: *edá* ‘day’, pl. *nná*.

Some plurals are suppletive: *osáni* ‘warrior’, pl. *asáfoɔ*

POSSESSION

Marked by order: X’s Y = XY – *Ghánà mǎñ* ‘the country of Ghana’; or, a possessive pronoun may be inserted: *abofára nǒ nhómā* ‘the child’s (his) book’.

In compounds *ab*, *a* may retain its isolate tone, *b* takes possessive tonal pattern: *Asante.héne.fíé* ‘the king of Ashanti’s palace’ (*héne* ‘king’, *ofíé* ‘house’).

LOCATIVE

There is a locative marker in *-beá*, and the noun expressing agency adds *-fo*; thus from the root *yaré* ‘to be ill’, we get: *ayaresábèa* ‘hospital’; *ɔyaresáfò* [*ɔyar-esáfòɔ*], ‘doctor’.

Adjective

As attribute, adjective follows noun, and is often reduplicated: *mmára fófòro pii* ‘many new laws’ (*mmará* ‘law’, *fófòro* ‘new’); *búùku kétéwaa tunturh* ‘a small black book’ (*tunturh* ‘dark in colour’).

Exceptionally, some adjectives make a plural form.

Comparative is made with *sēn*: *eyé duru* ‘it is heavy’, *eyé duru sēn búùku nó* ‘It is heavier than the book.’

Pronoun

Emphatic personal independent + subject markers (here + copula)

	Singular		Plural	
	Independent	Subject marker	Independent	Subject marker
1	mē	mēye	yéŋ	yeyé
2	wó	wóyè	mó	móyè
3	ɔnó	ɔye	wóŋ	wɔye
impersonal	enó	eye	enó	eye

The objective forms: *mē*, *wo*, *nō* are reduced as verbal suffixes to *-m̃*, *-w̃*, *-ñ*.

DEMONSTRATIVE PRONOUN

(*ε*)há ‘this/here’; (*ε*)hó ‘that/there’. *Eyí* is a demonstrative adjective: *nhómā yí* ‘this book’.

INTERROGATIVE PRONOUN

chēna ‘who?’

RELATIVE PRONOUN

nea, *áà*; *ɔyaresáfóɔ áà ɔɔkɔ* ‘the doctor who is going’.

Numerals

1–10: *baakó*, *mmienú*, *mmiensá*, *ennán*, *enúm*, *ensíá*, *ensón*, *ɛŋwɔtwé*, *ɛŋkorón*, *edú*. 11, *dúbàakó*; 12, *dúmienú* ... 20, *aduonū*; 21, *aduonúbàakó* ... 30, *aduasá*, 40, *aduanáñ*. 100, *ɔhá*, 101, *ɔhá né baakó* ...

Verb

Subject markers prefixed to stem: *hu(nu)* ‘to see’, present tense:

singular	1	mīhu	plural	yehū
	2	wúhū		múhū
	3	ohū		wohū

Negated by low-tone nasal prefixed to stem: *mēte Twū* ‘I speak Twi’; *mēnté Twū* ‘I don’t speak Twi’.