Codes for the Representation of Languages for Information Interchange

Status: For Ballot: February 9, 2001 - March 23, 2001

Abstract: A standardized 3-character code to indicate language in the exchange of information is defined. Codes are given for languages, contemporary and historical.

A Draft American National Standard
Developed by the
National Information Standards Organization

DRAFT STANDARD SUBJECT TO CHANGE

DRAFT STANDARD SUBJECT TO CHANGE

Contents

DRAFT STANDARD SUBJECT TO CHANGE

Foreword

This standard was originally prepared by Standards Committee C, Language Codes, which was organized in 1979. Charged with "providing a standard code for indicating languages for information interchange purposes," the committee produced a standard based on the list of MARC language codes developed by the Library of Congress in cooperation with the National Agricultural Library and the National Library of Medicine. This code list is now published as the MARC Code List for Languages.

Practical application of the MARC language codes has shown that in order to serve as an appropriate retrieval device for information, a standard list of language codes must reflect the linguistic content of the universal collection to which it is applied, with language codes assigned as needed to distinguish information in a given language or group of languages. The MARC language codes constitute such a list.

The committee's decision to base the standard on the existing MARC list took into account these contributing factors: (a) several years' successful application of the MARC language codes resulting in many millions of bibliographic records containing the accepted MARC codes, (b) the mnemonic relationship of the MARC codes to the English language names of the languages with English being the operational language of most American libraries, information services, and publishers, and (c) the flexibility inherent in a three-character code. The MARC list may be consulted for references from alternative forms of language names, as well as for the assignments to collective codes of languages for which individual codes have not been established.

This revised edition reflects a thorough review of the document and includes changes which are a result of requests and demonstrated need from users and implementors. In addition, it includes numerous changes necessary for compatibility with bibliographic language codes in ISO 639-2 (*Codes for the representation of names of languages: alpha-3 code).* The MARC code list is kept consistent with both ANSI/NISO Z39.53 and ISO 639-2/B.

NISO acknowledges with thanks the contributions of Rebecca Guenther, Library of Congress Network Development and MARC Standards Office, and Millicent Werwerka, Cataloging Policy and Support Office, Library of Congress, in revising this standard.

This standard was processed and approved for submittal to ANSI by the National Information Standards Organization. Suggestions for improving this standard are welcome. They should be sent to the National Information Standards Organization, 4733 Bethesda Avenue, Suite 300, Bethesda, MD 20814. NISO approval of this standard does not necessarily imply that all Voting Members voted for its approval. At the time it approved this standard, NISO had the following members:

NISO Voting Members

3M
Jerry Karel
Susan Boettcher (Alt)

Academic Press
Anthony Ross
Bradford Terry (Alt)

ALSi
Martin Sach
Tony O'Brien (Alt)

American Association of Law Libraries
Robert L. Oakley
Mary Alice Baish (Alt)

(continued)

NISO Voting Members *(continued)*

American Chemical Society
Robert S. Tannehill, Jr.

American Library Association
Paul J. Weiss

American Society for Information
  Science and Technology
Mark H. Needleman

American Society of Indexers
Charlotte Skuster
Marie Kascus (Alt)

American Theological Library
  Association
Myron Chace

ARMA International
Diane Carlisle

Art Libraries Society of North America
David L. Austin

Association for Information and Image
  Management
Betsy A. Fanning

Association of Information and
  Dissemination Centers
Bruce H. Kiesel

Association of Jewish Libraries
Carloine R. Miller
Elizabeth Vernon (Alt)

Association of Research Libraries
Duane E. Webster
Julia Blixrud (Alt)

Baker & Taylor
Robert H. Doran

Bell & Howell Information and Learning
Todd Fegan
James Brei (Alt)

Book Industry Communication
Brian Green

Broadcast Music Inc.
Edward Oshanani
Robert Barone (Alt)

Cambridge Scientific Abstracts
Matthew Dunie
Anthea Gotto (Alt)

CARL Corporation
Ward Shaw

College Center for Library Automation
J. Richard Madaus
Ann Armbrister (Alt)

Congressional Information Service, Inc.
Robert Lester

Data Research Associates, Inc.
Michael J. Mellinger
Mark H. Needleman (Alt)

Elsevier Science Inc.
John Mancia

Endeavor Information Systems, Inc.
Verne Coppi
Cindy Miller (Alt)

epixtech, inc.
John Bodfish
Ricc Ferante (Alt)

Ex Libris
James Steenbergen
Carl Grant (Alt)

Follett Corp.
D. Jeffrey Blumenthal
Don Rose (Alt)

Fretwell-Downing Informatics
Robin Murray

Gale Group
Katherine Gruber
Justine Carson (Alt)

Gaylord Information Systems
James English
William Schickling (Alt)

GCA Research Institute
Jane Harnad

Geac Computers, Inc.

H.W. Wilson Company
Ann Case

Indiana Cooperative Library Services Authority
Millard Johnson
Dea Szatkowski (Alt)

Infotrieve
Jan Peterson

Innovative Interfaces, Inc.
Gerald M. Kline
Sandy Westall (Alt)

Institute for Scientific Information
Helen Atkins
Richard Newman (Alt)

The International DOI Foundation
Norman Paskin

Kluwer Academic Publishers
Mike Casey

Library Binding Institute
Donald Dunham

The Library Corporation
Mark Wilson
Nancy Capps (Alt)

Library of Congress
Winston Tabb
Sally H. McCallum (Alt)

Los Alamos National Laboratory
Richard E. Luce

Lucent Technologies
M.E. Brennan

(continued)

DRAFT STANDARD SUBJECT TO CHANGE

NISO Voting Members *(continued)*

Medical Library Association
Nadine P. Ellero
Carla J. Funk (Alt)

MINITEX
Cecelia Boone
William DeJohn (Alt)

Modern Language Association
Terence Ford
Daniel Bokser (Alt)

Motion Picture Association of America
William M. Baker
Axel aus der Muhlen (Alt)

Music Library Association
Lenore Coral
Mark McKnight (Alt)

National Agricultural Library
Pamela Q. J. Andre
Gary K. McCone (Alt)

National Archives and Records
  Administration
Mary Ann Hadyka

National Federation of Abstracting and
  Information Services
Marion Harrell

National Library of Medicine
Betsy L. Humphreys

Nylink
Mary-Alice Lynch
Jane Neale (Alt)

OASIS
Laura Walker

OCLC, Inc.
Donald J. Muccino

Openly Informatics
Eric Hellman (Alt)

R.R. Bowker
Albert Simmonds

Recording Industry Assn. of America
Linda R. Bocchi
Michael Williams (Alt)

Research Libraries Group
Lennie Stovel
Joan Aliprand (Alt)

RoweCom, Inc.
Robert Boissy
Marilyn Geller (Alt)

SIRS Mandarin, Inc.
Leonardo Lazo
Harry Kaplanian (Alt)

SIRSI Corporation
Greg Hathorn
Slavko Manojlovich (Alt)

Society for Technical Communication
Annette Reilly
Kevin Burns (Alt)

Society of American Archivists
Lisa Weber

Special Libraries Association
Marjorie Hlava

Sun Microsystems Inc.
Cheryl Wright
John C. Fowler (Alt)

Triangle Research Libraries Network
Jordan M. Scepanski
Mona C. Couts (Alt)

U.S. Department of Commerce, National
  Institute of Standards and Technology,
  Office of Information Services

U.S. Department of Defense, Defense Technical
  Information Center
Gopalakrishnan Nair
Jane Cohen (Alt)

U.S. National Commission on Libraries and
  Information Science
Denise Davis

VTLS, Inc.
Vinod Chachra

(continued)

At the time NISO approved this standard, the following individuals served on its Board of Directors:

Codes for the Representation of Languages for Information Interchange

1. Purpose and Scope

1.1 Purpose

The purpose of this standard is to provide libraries, information services, and publishers a standardized code to indicate language in the exchange of information. This standard for language codes is not a prescriptive device for the definition of language and dialects but rather a list reflecting the need to distinguish recorded information by language.

1.2 Scope

1.2.1 Usage

Language codes are designed to be used:

1. To designate the languages in which documents are or have been written or recorded;
2. To designate the languages in which document handling records (order records, bibliographic records, and the like) have been created.

1.2.2 Limitation

Language codes are not designed to be used:

1. To designate machine programming languages (FORTRAN, BASIC, and the like);
2. To distinguish languages from dialects. The dialect of a language is usually represented by the same language code as that used for the language.

2. General Principles of the List

2.1 Structure of the Language Codes

Each code comprises three roman alphabet characters. Codes generally were created using three characters usually based on an English form of the language name or, in some cases, a vernacular form of the corresponding language name. Future development of language codes will be based, whenever possible, on the vernacular form of the language, unless another language code is requested by the country or countries using the language. The codes are varied where necessary to resolve conflicts.

2.2 Specificity of the Language Codes

Language codes are assigned either to individual languages or to related groups of languages. The level of specificity of the language code assigned is determined in each case to be the level necessary to maintain the utility of the standard based on the volume of documents or document handling records that have been or are expected to be written, recorded, or created. Levels of specificity represented by the language codes include:

1. Language codes for individual languages;

2. Collective language codes for linguistically or otherwise related groups of languages;

3. Collective language codes for linguistically or otherwise related groups of languages having individual language codes for some but not all languages so related.

This standard does not indicate which level of specificity is represented by each code. The word "languages" or "other" as part of a descriptor may be taken to indicate that a language code is a collective language code. A collective language code is not intended to be used when an individual language code or another more specific collective language code is available.

2.3 Future Development

To ensure continuity and stability, language codes shall rarely be changed and only for compelling reasons. To accommodate large applications that build continuously, the codes in Z39.53 shall not be changed if a language name or its abbreviation are changed.

2.4. Script

A single language code is normally assigned for a language even though the language is written in more than one script.

2.5. Special Language Code

2.5.1 Language Code MIS (Miscellaneous)

The language code MIS for Miscellaneous is assigned to languages that do not have specific individual or collective codes assigned; Etruscan is an example of a language to which the code MIS has been assigned.

2.5.2 Language Code ART (Artificial [Other])

The collective language code ART for Artificial (Other) is assigned for those languages not serving as native tongues but developed for international communication or some

DRAFT STANDARD SUBJECT TO CHANGE

more specific purpose. Among the languages assigned this code are Babm and Loglan. Individual artificial languages in which a sufficiently significant number of documents or document handling records are or have been written, recorded, or created are assigned individual language codes (e.g., EPO for Esperanto).

2.5.3 Language Code MUL (Multiple Languages)

The language code MUL for Multiple Languages is assigned for those situations in which several language codes would be needed to reflect the linguistic content of a document or document handling record, and it is not practical to use all the appropriate codes.

2.5.4 Language Code UND (Undetermined)

The language code UND for Undetermined is assigned for those situations in which the agency applying the standard is unable to determine the linguistic content of a document or document handling record sufficiently to apply the appropriate individual or collective code(s).

3. Structure of the List of Language Codes

3.1 Descriptors

Each language code is accompanied by a descriptor. Descriptors generally are based on the form of language name found in Library of Congress Subject Headings. When the descriptor is changed, the code is not also changed, but remains the same.

3.2 Arrangement of the List of Language Codes

The list of language codes is presented in two versions:

1. In alphabetical order by language code (see Table 1);

2. In alphabetical order by descriptor (see Table 2).

A list of formerly valid but now obsolete codes are included in Table 3.

4. Maintenance Agency

The maintenance agency designated in the Appendix will be responsible for reviewing and acting upon suggestions for new language codes and for maintaining a current listing of all language codes and their definitions. Questions concerning the implementation of this standard and requests for information should be sent to the maintenance agency.

DRAFT STANDARD SUBJECT TO CHANGE

Table 1: List of Language Codes in Alphabetical Order by Code

Code	Descriptor	Code	Descriptor
aar	Afar	bla	Siksika
abk	Abkhaz	bnt	Bantu (Other)
ace	Achinese	bos	Bosnian
ach	Acoli	bra	Braj
ada	Adangme	bre	Breton
afa	Afroasiatic (Other)	btk	Batak
afh	Afrihili (Artificial language)	bua	Buriat
afr	Afrikaans	bug	Bugis
aka	Akan	bul	Bulgarian
akk	Akkadian	bur	Burmese
alb	Albanian		
ale	Aleut	cad	Caddo
alg	Algonquian (Other)	cai	Central American Indian (Other)
amh	Amharic	car	Carib
ang	English, Old (ca. 450-1100)	cat	Catalan
apa	Apache languages	cau	Caucasian (Other)
ara	Arabic	ceb	Cebuano
arc	Aramaic	cel	Celtic (Other)
arm	Armenian	cha	Chamorro
arn	Mapuche	chb	Chibcha
arp	Arapaho	che	Chechen
art	Artificial (Other)	chg	Chagatai
arw	Arawak	chi	Chinese
asm	Assamese	chk	Truk
ath	Athapascan (Other)	chm	Mari
aus	Australian languages	chn	Chinook jargon
ava	Avaric	cho	Choctaw
ave	Avestan	chp	Chipewyan
awa	Awadhi	chr	Cherokee
aym	Aymara	chu	Church Slavic
aze	Azerbaijani	chv	Chuvash
		chy	Cheyenne
bad	Banda	cmc	Chamic languages
bai	Bamileke languages	cop	Coptic
bak	Bashkir	cor	Cornish
bal	Baluchi	cos	Corsican
bam	Bambara	cpe	Creoles and Pidgins, English-based (Other)
ban	Balinese		
baq	Basque	cpf	Creoles and Pidgins, French-based (Other)
bas	Basa		
bat	Baltic (Other)	cpp	Creoles and Pidgins, Portuguese-based (Other)
bej	Beja		
bel	Belarusian	cre	Cree
bem	Bemba	crp	Creoles and Pidgins (Other)
ben	Bengali	cus	Cushitic (Other)
ber	Berber (Other)	cze	Czech
bho	Bhojpuri		
bih	Bihari	dak	Dakota
bik	Bikol	dan	Danish
bin	Bini	day	Dayak
bis	Bislama	del	Delaware

Code	Descriptor	Code	Descriptor
den	Slave	goh	German, Old High (ca.750-1050)
dgr	Dogrib	gon	Gondi
din	Dinka	gor	Gorontalo
div	Divehi	got	Gothic
doi	Dogri	grb	Grebo
dra	Dravidian (Other)	grc	Greek, Ancient (to 1453)
dua	Duala	gre	Greek, Modern (1453-)
dum	Dutch, Middle (ca. 1050-1350)	grn	Guarani
dut	Dutch	guj	Gujarati
dyu	Dyula	gwi	Gwich'in
dzo	Dzongkha		
		hai	Haida
efi	Efik	hau	Hausa
egy	Egyptian	haw	Hawaiian
eka	Ekajuk	heb	Hebrew
elx	Elamite	her	Herero
eng	English	hil	Hiligaynon
enm	English, Middle (ca. 1100-1500)	him	Himachali
epo	Esperanto	hin	Hindi
est	Estonian	hit	Hittite
ewe	Ewe	hmn	Hmong
ewo	Ewondo	hmo	Hiri Motu
		hun	Hungarian
fan	Fang	hup	Hupa
fao	Faroese	iba	Iban
fat	Fanti	ibo	Igbo
fij	Fijian	ice	Icelandic
fin	Finnish	ijo	Ijo
fiu	Finno-Ugrian (Other)	iku	Inuktitut
fon	Fon	ile	Interlingue
fre	French	ilo	Iloko
frm	French, Middle (ca. 1400-1600)	ina	Interlingua (International Auxiliary Language Association)
fro	French, Old (842-ca. 1400)	inc	Indic (Other)
fry	Frisian	ind	Indonesian
ful	Fula	ine	Indo-European (Other)
fur	Friulian	ipk	Inupiaq
		ira	Iranian (Other)
gaa	Gã	iro	Iroquoian (Other)
gay	Gayo	ita	Italian
gba	Gbaya		
gem	Germanic (Other)	jav	Javanese
geo	Georgian	jpn	Japanese
ger	German	jpr	Judeo-Persian
gez	Ethiopic	jrb	Judeo-Arabic
gil	Gilbertese		
gla	Scottish Gaelic	kaa	Kara-Kalpak
gle	Irish	kab	Kabyle
glg	Galician	kac	Kachin
glv	Manx	kal	Kalâtdlisut
gmh	German, Middle High (ca. 1050-1500)	kam	Kamba
		kan	Kannada

DRAFT STANDARD SUBJECT TO CHANGE

Code	Descriptor	Code	Descriptor
kar	Karen	man	Mandingo
kas	Kashmiri	mao	Maori
kau	Kanuri	map	Austronesian (Other)
kaw	Kawi	mar	Marathi
kaz	Kazakh	mas	Masai
kha	Khasi	may	Malay
khi	Khoisan (Other)	mdr	Mandar
khm	Khmer	men	Mende
kho	Khotanese	mga	Irish, Middle (ca. 1100-1550)
kik	Kikuyu	mic	Micmac
kin	Kinyarwanda	min	Minangkabau
kir	Kyrgyz	mis	Miscellaneous languages
kmb	Kimbundu	mkh	Mon-Khmer (Other)
kok	Konkani	mlg	Malagasy
kom	Komi	mlt	Maltese
kon	Kongo	mnc	Manchu
kor	Korean	mni	Manipuri
kos	Kusaie	mno	Manobo languages
kpe	Kpelle	moh	Mohawk
kro	Kru (Other)	mol	Moldavian
kru	Kurukh	mon	Mongolian
kua	Kuanyama	mos	Mooré
kum	Kumyk	mul	Multiple languages
kur	Kurdish	mun	Munda (Other)
kut	Kutenai	mus	Creek
		mwr	Marwari
lad	Ladino	myn	Mayan languages
lah	Lahndā		
lam	Lamba	nah	Nahuatl
lao	Lao	nai	North American Indian (Other)
lat	Latin	nau	Nauru
lav	Latvian	nav	Navajo
lez	Lezgian	nbl	Ndebele (South Africa)
lin	Lingala	nde	Ndebele (Zimbabwe)
lit	Lithuanian	ndo	Ndonga
lol	Mongo-Nkundu	nds	Low German
loz	Lozi	nep	Nepali
ltz	Letzeburgesch	new	Newari
lua	Luba-Lulua	nia	Nias
lub	Luba-Katanga	nic	Niger-Kordofanian (Other)
lug	Ganda	niu	Niuean
lui	Luiseño	nno	Nynorsk
lun	Lunda	nob	Bokmål
luo	Luo (Kenya and Tanzania)	non	Old Norse
lus	Lushai	nor	Norwegian
		nso	Northern Sotho
mac	Macedonian	nub	Nubian languages
mad	Madurese	nya	Nyanja
mag	Magahi	nym	Nyamwezi
mah	Marshall	nyn	Nyankole
mai	Maithili	nyo	Nyoro
mak	Makasar	nzi	Nzima
mal	Malayalam		

Code	Descriptor	Code	Descriptor
oci	Occitan (post-1500)	sem	Semitic (Other)
oji	Ojibwa	sga	Irish, Old (to 1100)
ori	Oriya	sgn	Sign languages
orm	Oromo	shn	Shan
osa	Osage	sid	Sidamo
oss	Ossetic	sin	Sinhalese
ota	Turkish, Ottoman	sio	Siouan (Other)
oto	Otomian languages	sit	Sino-Tibetan (Other)
		sla	Slavic (Other)
paa	Papuan (Other)	slo	Slovak
pag	Pangasinan	slv	Slovenian
pal	Pahlavi	sme	Northern Sami
pam	Pampanga	smi	Sami
pan	Panjabi	smo	Samoan
pap	Papiamento	sna	Shona
pau	Palauan	snd	Sindhi
peo	Old Persian (ca. 600-400 B.C.)	snk	Soninke
per	Persian	sog	Sogdian
phi	Philippine (Other)	som	Somali
phn	Phoenician	son	Songhai
pli	Pali	sot	Sotho
pol	Polish	spa	Spanish
pon	Ponape	srd	Sardinian
por	Portuguese	srr	Serer
pra	Prakrit languages	ssa	Nilo-Saharan (Other)
pro	Provençal (to 1500)	ssw	Swazi
pus	Pushto	suk	Sukuma
		sun	Sundanese
que	Quechua	sus	Susu
		sux	Sumerian
raj	Rajasthani	swa	Swahili
rap	Rapanui	swe	Swedish
rar	Rarotongan	syr	Syriac
roa	Romance (Other)		
roh	Raeto-Romance	tah	Tahitian
rom	Romany	tai	Tai (Other)
rum	Romanian	tam	Tamil
run	Rundi	tat	Tatar
rus	Russian	tel	Telugu
		tem	Temne
sad	Sandawe	ter	Terena
sag	Sango	tet	Tetum
sah	Yakut	tgk	Tajik
sai	South American Indian (Other)	tgl	Tagalog
sal	Salishan languages	tha	Thai
sam	Samaritan Aramaic	tib	Tibetan
san	Sanskrit	tig	Tigré
sas	Sasak	tir	Tigrinya
sat	Santali	tiv	Tiv
scc	Serbian	tkl	Tokelauan
sco	Scots	tli	Tlingit
scr	Croatian	tmh	Tamashek
sel	Selkup	tog	Tonga (Nyasa)

DRAFT STANDARD SUBJECT TO CHANGE

Code	Descriptor	Code	Descriptor
ton	Tongan	wak	Wakashan languages
tpi	Tok Pisin	wal	Walamo
tsi	Tsimshian	war	Waray
tsn	Tswana	was	Washo
tso	Tsonga	wel	Welsh
tuk	Turkmen	wen	Sorbian languages
tum	Tumbuka	wol	Wolof
tur	Turkish		
tut	Altaic (Other)	xho	Xhosa
tvl	Tuvaluan		
twi	Twi	yao	Yao
tyv	Tuvinian	yap	Yapese
		yid	Yiddish
uga	Ugaritic	yor	Yoruba
uig	Uighur	ypk	Yupik languages
ukr	Ukrainian		
umb	Umbundu	zap	Zapotec
und	Undetermined	zen	Zenaga
urd	Urdu	zha	Zhuang
uzb	Uzbek	znd	Zande
		zul	Zulu
vai	Vai	zun	Zuni
ven	Venda		
vie	Vietnamese		
vol	Volapük		
vot	Votic		

Table 2: List of Language Codes in Alphabetical Order by Descriptor

Descriptor	Code
Abkhaz	abk
Achinese	ace
Acoli	ach
Adangme	ada
Afar	aar
Afrihili (Artificial language)	afh
Afrikaans	afr
Afroasiatic (Other)	afa
Akan	aka
Akkadian	akk
Albanian	alb
Aleut	ale
Algonquian (Other)	alg
Altaic (Other)	tut
Amharic	amh
Apache languages	apa
Arabic	ara
Aramaic	arc
Arapaho	arp
Arawak	arw
Armenian	arm
Artificial (Other)	art
Assamese	asm
Athapascan (Other)	ath
Australian languages	aus
Austronesian (Other)	map
Avaric	ava
Avestan	ave
Awadhi	awa
Aymara	aym
Azerbaijani	aze
Balinese	ban
Baltic (Other)	bat
Baluchi	bal
Bambara	bam
Bamileke languages	bai
Banda	bad
Bantu (Other)	bnt
Basa	bas
Bashkir	bak
Basque	baq
Batak	btk
Beja	bej
Belarusian	bel
Bemba	bem
Bengali	ben
Berber (Other)	ber
Bhojpuri	bho

Descriptor	Code
Bihari	bih
Bikol	bik
Bini	bin
Bislama	bis
Bokmål	nob
Bosnian	bos
Braj	bra
Breton	bre
Bugis	bug
Bulgarian	bul
Buriat	bua
Burmese	bur
Caddo	cad
Carib	car
Catalan	cat
Caucasian (Other)	cau
Cebuano	ceb
Celtic (Other)	cel
Central American Indian (Other)	cai
Chagatai	chg
Chamic languages	cmc
Chamorro	cha
Chechen	che
Cherokee	chr
Cheyenne	chy
Chibcha	chb
Chinese	chi
Chinook jargon	chn
Chipewyan	chp
Choctaw	cho
Church Slavic	chu
Chuvash	chv
Coptic	cop
Cornish	cor
Corsican	cos
Cree	cre
Creek	mus
Creoles and Pidgins (Other)	crp
Creoles and Pidgins, English-based (Other)	cpe
Creoles and Pidgins, French-based (Other)	cpf
Creoles and Pidgins, Portuguese-based (Other)	cpp
Croatian	scr
Cushitic (Other)	cus
Czech	cze

DRAFT STANDARD SUBJECT TO CHANGE

Descriptor	Code	Descriptor	Code
Dakota	dak	German, Middle High	
Danish	dan	(ca. 1050-1500)	gmh
Dayak	day	German, Old High (ca. 750-1050)	goh
Delaware	del	Germanic (Other)	gem
Dinka	din	Gilbertese	gil
Divehi	div	Gondi	gon
Dogri	doi	Gorontalo	gor
		Gothic	got
Dogrib	dgr	Grebo	grb
Dravidian (Other)	dra	Greek, Ancient (to 1453)	grc
Duala	dua	Greek, Modern (1453-)	gre
Dutch	dut	Guarani	grn
Dutch, Middle		Gujarati	guj
(ca. 1050-1350)	dum	Gwich'in	gwi
Dyula	dyu		
Dzongkha	dzo	Haida	hai
		Hausa	hau
Efik	efi	Hawaiian	haw
Egyptian	egy	Hebrew	heb
Ekajuk	eka	Herero	her
Elamite	elx	Hiligaynon	hil
English	eng	Himachali	him
English, Middle		Hindi	hin
(ca. 1100-1500)	enm	Hiri Motu	hmo
English, Old (ca. 450-1100)	ang	Hittite	hit
Esperanto	epo	Hmong	hmn
Estonian	est	Hungarian	hun
Ethiopic	gez	Hupa	hup
Ewe	ewe		
Ewondo	ewo	Iban	iba
		Icelandic	ice
Fang	fan	Igbo	ibo
Fanti	fat	Ijo	ijo
Faroese	fao	Iloko	ilo
Fijian	fij	Indic (Other)	inc
Finnish	fin	Indo-European (Other)	ine
Finno-Ugrian (Other)	fiu	Indonesian	ind
Fon	fon	Interlingua (International	ina
French	fre	Auxiliary Language Association)	
French, Middle		Interlingue	ile
(ca. 1400-1600)	frm	Inuktitut	iku
French, Old (842-ca.1400)	fro	Inupiaq	ipk
Frisian	fry	Iranian (Other)	ira
Friulian	fur	Irish	gle
Fula	ful	Irish, Middle (ca. 1100-1550)	mga
		Irish, Old (to 1100)	sga
Gã	gaa	Iroquoian (Other)	iro
Galician	glg	Italian	ita
Ganda	lug		
Gayo	gay	Japanese	jpn
Gbaya	gba	Javanese	jav
Georgian	geo	Judeo-Arabic	jrb
German	ger	Judeo-Persian	jpr

Descriptor	Code	Descriptor	Code
Kabyle	kab	Magahi	mag
Kachin	kac	Maithili	mai
Kalâtdlisut	kal	Makasar	mak
Kamba	kam	Malagasy	mlg
Kannada	kan	Malay	may
Kanuri	kau	Malayalam	mal
Kara-Kalpak	kaa	Maltese	mlt
Karen	kar	Manchu	mnc
Kashmiri	kas	Mandar	mdr
Kawi	kaw	Mandingo	man
Kazakh	kaz	Manipuri	mni
Khasi	kha	Manobo languages	mno
Khmer	khm	Manx	glv
Khoisan (Other)	khi	Maori	mao
Khotanese	kho	Mapuche	arn
Kikuyu	kik	Marathi	mar
Kimbundu	kmb	Mari	chm
Kinyarwanda	kin	Marshall	mah
Komi	kom	Marwari	mwr
Kongo	kon	Masai	mas
Konkani	kok	Mayan languages	myn
Korean	kor	Mende	men
Kpelle	kpe	Micmac	mic
Kru (Other)	kro	Minangkabau	min
Kuanyama	kua	Miscellaneous languages	mis
Kumyk	kum	Mohawk	moh
Kurdish	kur	Moldavian	mol
Kurukh	kru	Mon-Khmer (Other)	mkh
Kusaie	kos	Mongo-Nkundu	lol
Kutenai	kut	Mongolian	mon
Kyrgyz	kir	Mooré	mos
		Multiple languages	mul
Ladino	lad	Munda (Other)	mun
Lahndā	lah		
Lamba	lam	Nahuatl	nah
Lao	lao	Nauru	nau
Latin	lat	Navajo	nav
Latvian	lav	Ndebele (South Africa)	nbl
Letzeburgesch	ltz	Ndebele (Zimbabwe)	nde
Lezgian	lez	Ndonga	ndo
Lingala	lin	Nepali	nep
Lithuanian	lit	Newari	new
Low German	nds	Nias	nia
Lozi	loz	Niger-Kordofanian (Other)	nic
Luba-Katanga	lub	Nilo-Saharan (Other)	ssa
Luba-Lulua	lua	Niuean	niu
Luiseño	lui	North American Indian (Other)	nai
Lunda	lun	Northern Sami	sme
Luo (Kenya and Tanzania)	luo	Northern Sotho	nso
Lushai	lus	Norwegian	nor
		Nubian languages	nub
Macedonian	mac	Nyamwezi	nym
Madurese	mad	Nyanja	nya

DRAFT STANDARD SUBJECT TO CHANGE

Descriptor	Code	Descriptor	Code
Nyankole	nyn	Sardinian	srd
Nynorsk	nno	Sasak	sas
Nyoro	nyo	Scots	sco
Nzima	nzi	Scottish Gaelic	gla
		Selkup	sel
Occitan (post-1500)	oci	Semitic (Other)	sem
Ojibwa	oji	Serbian	scc
Old Norse	non	Serer	srr
Old Persian (ca. 600-400 B.C.)	peo	Shan	shn
Oriya	ori	Shona	sna
Oromo	orm	Sidamo	sid
Osage	osa	Sign languages	sgn
Ossetic	oss	Siksika	bla
Otomian languages	oto	Sindhi	snd
		Sinhalese	sin
Pahlavi	pal	Sino-Tibetan (Other)	sit
Palauan	pau	Siouan (Other)	sio
Pali	pli	Slave	den
Pampanga	pam	Slavic (Other)	sla
Pangasinan	pag	Slovak	slo
Panjabi	pan	Slovenian	slv
Papiamento	pap	Sogdian	sog
Papuan (Other)	paa	Somali	som
Persian	per	Songhai	son
Philippine (Other)	phi	Soninke	snk
Phoenician	phn	Sorbian languages	wen
Polish	pol	Sotho	sot
Ponape	pon	South American Indian (Other)	sai
Portuguese	por	Spanish	spa
Prakrit languages	pra	Sukuma	suk
Provençal (to 1500)	pro	Sumerian	sux
Pushto	pus	Sundanese	sun
		Susu	sus
Quechua	que	Swahili	swa
		Swazi	ssw
Raeto-Romance	roh	Swedish	swe
Rajasthani	raj	Syriac	syr
Rapanui	rap		
Rarotongan	rar	Tagalog	tgl
Romance (Other)	roa	Tahitian	tah
Romanian	rum	Tai (Other)	tai
Romany	rom	Tajik	tgk
Rundi	run	Tamashek	tmh
Russian	rus	Tamil	tam
		Tatar	tat
Salishan languages	sal	Telugu	tel
Samaritan Aramaic	sam	Temne	tem
Sami	smi	Terena	ter
Samoan	smo	Tetum	tet
Sandawe	sad	Thai	tha
Sango	sag	Tibetan	tib
Sanskrit	san	Tigré	tig
Santali	sat	Tigrinya	tir

DRAFT STANDARD SUBJECT TO CHANGE

Descriptor	Code	Descriptor	Code
Tiv	tiv	Vai	vai
Tlingit	tli	Venda	ven
Tok Pisin	tpi	Vietnamese	vie
Tokelauan	tkl	Volapük	vol
Tonga (Nyasa)	tog	Votic	vot
Tongan	ton	Wakashan languages	wak
Truk	chk	Walamo	wal
Tsimshian	tsi	Waray	war
Tsonga	tso	Washo	was
Tswana	tsn	Welsh	wel
Tumbuka	tum	Wolof	wol
Turkish	tur		
Turkish, Ottoman	ota	Xhosa	xho
Turkmen	tuk		
Tuvaluan	tvl	Yakut	sah
Tuvinian	tyv	Yao	yao
Twi	twi	Yapese	yap
		Yiddish	yid
		Yoruba	yor
Ugaritic	uga		
Uighur	uig	Zande	znd
Ukrainian	ukr	Zapotec	zap
Umbundu	umb	Zenaga	zen
Undetermined	und	Zhuang	zha
Urdu	urd	Zulu	zul
Uzbek	uzb	Zuni	zun

DRAFT STANDARD SUBJECT TO CHANGE

Table 3: List of Obsolete Language Codes

Old code	Descriptor	New code
cam	Khmer	khm
esk	Eskimo languages	(deleted)
esp	Esperanto	epo
eth	Ethiopic	gez
far	Faroese	fao
fri	Frisian	fry
gae	Scottish Gaelic	gla
gag	Galician	glg
gal	Oromo	orm
gua	Guarani	grn
int	Interlingua (International Auxiliary Language Association)	ina
iri	Irish	gle
kus	Kusaie	kos
lan	Occitan (post-1500)	oci
lap	Sami	smi
max	Manx	glv
mla	Malagasy	mlg
sao	Samoan	smo
sho	Shona	sna
snh	Sinhalese	sin
sso	Sotho	sot
swz	Swazi	ssw
tag	Tagalog	tgl
taj	Tajik	tgk
tar	Tatar	tat
tru	Truk	chk
tsw	Tswana	tsn

APPENDIX
Designation of Maintenance Agency
(This Appendix is not part of American National Standard Z39.53-200x Codes for the
Representation of Languages for Information Interchange. It is included for information only.)

The functions assigned to the maintenance agency as specified in Section 4 will be
administered by the Library of Congress. Questions concerning the implementation of
this standard and requests for information should be sent to the Cataloging Policy and
Support Office, Library of Congress, Washington, D.C. 20540.

DRAFT STANDARD SUBJECT TO CHANGE