

**Språkkoder –**

Del 2: Treställiga bokstavskoder  
(ISO 639-2:1998, IDT)

**Codes for the representation of names of  
languages –**

Part 2: Alpha-3 code (ISO 639-2:1998, IDT)

Den internationella standarden ISO 639-2:1998 gäller som svensk standard. Detta dokument innehåller den officiella engelska versionen av ISO 639-2:1998.

The International Standard ISO 639-2:1998 has the status of a Swedish Standard. This document contains the official English version of ISO 639-2:1998.

---

Upplýsingar om **sakinnehållet** i standarden lämnas av SIS, Swedish Standards Institute, telefon 08 - 555 520 00.

Standarder kan beställas hos SIS Förlag AB som även lämnar **allmänna upplýsingar** om svensk och utländsk standard.

*Postadress:* SIS Förlag AB, 118 80 STOCKHOLM  
*Telefon:* 08 - 555 523 10. *Telefax:* 08 - 555 523 11  
*E-post:* [sis.sales@sis.se](mailto:sis.sales@sis.se). *Internet:* [www.sis.se](http://www.sis.se)

## Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

International Standard ISO 639-2 was prepared jointly by Technical Committees ISO/TC 37, *Terminology (principles and coordination)*, Subcommittee SC 2, *Layout of vocabularies* and ISO/TC 46, *Information and documentation*, Subcommittee SC 4, *Computer applications in information and documentation*.

ISO 639 consists of the following parts, under the general title *Codes for the representation of names of languages*:

- *Part 1: Alpha-2 code (revision of ISO 639:1988)*
- *Part 2: Alpha-3 code*

Annex A forms an integral part of this part of ISO 639.

## **SS-ISO 639-2:2005**

### **Introduction**

ISO 639 provides two sets of language codes, one as a two-letter code set (ISO 639-1) and another as a three-letter code set (this part of ISO 639) for the representation of names of languages. ISO 639-1 was devised primarily for use in terminology, lexicography and linguistics. This part of ISO 639 represents all languages contained in ISO 639-1 and in addition any other language, as well as language groups, as they may be coded for special purposes when more specificity in coding is needed. The languages listed in ISO 639-1 are a subset of the languages listed in this part of ISO 639; every language code in the two-letter code set has a corresponding language code in the alpha-3 list, but not necessarily vice versa.

Both code lists are to be considered as open lists. The codes were devised for use in terminology, lexicography, information and documentation (i.e. for libraries, information services, and publishers) and linguistics. This part of ISO 639 also includes guidelines for the creation of language codes and their use in some applications.

## Codes for the representation of names of languages —

### Part 2: Alpha-3 code

#### 1 Scope

This part of ISO 639 provides two sets of three-letter alphabetic codes for the representation of names of languages, one for terminology applications and the other for bibliographic applications. The code sets are the same except for twenty-five languages that have variant language codes because of the criteria used for formulating them (see 4.1). The language codes were devised originally for use by libraries, information services, and publishers to indicate language in the exchange of information, especially in computerized systems. These codes have been widely used in the library community and may be adopted for any application requiring the expression of language in coded form by terminologists and lexicographers. The alpha-2 code set was devised for practical use for most of the major languages of the world that are most frequently represented in the total body of the world's literature. Additional language codes are created when it becomes apparent that a significant body of literature in a particular language exists. Languages designed exclusively for machine use, such as computer programming languages, are not included in this code.

## Codes pour la représentation des noms de langue —

### Partie 2: Code alpha-3

#### 1 Domaine d'application

La présente partie de l'ISO 639 fournit deux jeux de codets à trois caractères pour la représentation des noms de langue, l'un destiné à des applications en terminologie, l'autre destiné à des applications bibliographiques. Les deux jeux de codets sont les mêmes, excepté pour vingt-cinq langues qui ont des codets différents à cause du critère utilisé pour les former (voir 4.1). À l'origine, des codets de langues ont été créés à l'usage des bibliothèques, des services documentaires et des éditeurs, pour indiquer la langue dans l'échange d'information, particulièrement entre des systèmes informatisés. Ces codets ont été utilisés largement à l'intérieur de la communauté des bibliothèques, et ils peuvent être adoptés pour toute application qui requiert l'expression de la langue sous forme codée, en terminologie ou en lexicographie. Le code alpha-2 a été créé pour un usage pratique pour la plupart des grandes langues du monde, langues qui sont les plus souvent représentées dans le corpus de la littérature mondiale. De nouveaux codets sont créés quand il devient évident qu'il existe un corpus significatif de littérature dans une langue particulière. Les langages créés pour être utilisés exclusivement par des machines, comme les langages de programmation, ne sont pas inclus dans le présent code.

## SS-ISO 639-2:2005

### 2 Normative reference

The following standard contains provisions which, through reference in this text, constitute provisions of this part of ISO 639. At the time of publication, the edition indicated was valid. All standards are subject to revision, and parties to agreements based on this part of ISO 639 are encouraged to investigate the possibility of applying the most recent edition of the standard indicated below. Members of IEC and ISO maintain registers of currently valid International Standards.

ISO 3166-1:1997, *Codes for the representation of names of countries and their subdivisions — Part 1: Country codes*.

### 3 Definitions

For the purposes of this part of ISO 639, the following definitions apply.

#### 3.1 code

data representation in different forms according to a pre-established set of rules

#### 3.2 language code

combination of characters used to represent a language or languages

#### 3.3 collective language code

language code used to represent a group of languages

### 4 Language codes

#### 4.1 Form of the language codes

The language codes consist of three Latin-alphabet characters in lowercase. No diacritical marks or modified characters are used. Implementors should be aware that these codes are not intended to be an abbreviation for the language, but to serve as a device to identify a given language or group of languages. The language codes are derived from the language name.

Two code sets are provided, one for bibliographic applications (ISO 639-2/B), and one for terminology applications (ISO 639-2/T). Criteria for selecting the form of a language code for code set B were

— preference of the countries using the language,

### 2 Référence normative

La norme suivante contient des dispositions qui, par suite de la référence qui en est faite, constituent des dispositions valables pour la présente partie de l'ISO 639. Au moment de la publication, l'édition indiquée était en vigueur. Toute norme est sujette à révision, et les parties prenantes des accords fondés sur la présente partie de l'ISO 639 sont invités à rechercher la possibilité d'appliquer l'édition la plus récente de la norme indiquée ci-après. Les membres de la CEI et de l'ISO possèdent le registre des Normes internationales en vigueur à un moment donné.

ISO 3166-1:1997, *Codes pour la représentation des noms de pays et leurs subdivisions — Partie 1: Codes pays*.

### 3 Définitions

Pour les besoins de la présente partie de l'ISO 639, les définitions suivantes s'appliquent.

#### 3.1 code

représentation de données sous différentes formes, selon un jeu de règles préétablies

#### 3.2 codet de langue

combinaison de caractères utilisée pour représenter une langue ou des langues

#### 3.3 codet de langue collectif

codet de langue utilisé pour représenter un groupe de langue

### 4 Codets de langue

#### 4.1 Forme des codets de langue

Les codets de langue sont écrits en minuscules et sont composés de trois lettres de l'alphabet latin, à l'exclusion de tout signe diacritique et de tout caractère modifié. Il est recommandé aux utilisateurs d'être attentifs à ce que ces codes n'ont pas été conçus pour être une abréviation du nom de la langue, mais pour servir de symbole pour identifier une langue donnée ou un groupe de langue. Les codets de langue sont dérivés du nom de la langue.

Deux codes sont fournis, l'un pour des applications bibliographiques (ISO 639-2/B), l'autre pour des applications en terminologie (ISO 639-2/T). Les critères

- established usage of codes in national and international bibliographic databases, and
- the vernacular or English form of the language.

Code set T was based on

- the vernacular form of the language, or
- preference of the countries using the language.

There are twenty-three language names that have variant codes assigned depending on the code set chosen.

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 bibliographic or terminology code set must be used in its entirety, and the choice of the set used must be made clear by exchanging partners prior to information interchange. Users shall refer to ISO 639-2/B for the code set for bibliographic applications and ISO 639-2/T for the code set for terminology applications.

To ensure continuity and stability, codes shall only be changed for compelling reasons. After a change in codes, the previous code shall not be reassigned for at least five years. To accommodate large applications that build continuously, the codes in ISO 639-2/B shall not be changed if a language name or its abbreviation are changed.

When adapting this part of ISO 639 to languages using other alphabets (e.g. Cyrillic), codes shall be formed according to the principles of this part of ISO 639.

#### 4.1.1 Collective language codes

Collective language codes are provided for languages where a relatively small number of documents exist or are expected to be written, recorded or created. The words *languages* or *(other)* as part of a language name in the following tables 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.

pour choisir la forme du codet de langue dans le code B sont

- la préférence des pays utilisant la langue,
- l'utilisation avérée de ce code dans des banques de données bibliographiques nationales et internationales, et
- la forme du nom de la langue en anglais ou dans la langue vernaculaire.

Le code T est fondé sur

- la forme vernaculaire du nom de la langue, ou
- la préférence des pays qui utilisent cette langue.

Il existe vingt-trois noms de langue qui ont des codets différents selon le code que l'on a choisi.

La mise au point de nouveaux codets de langue sera fondée, chaque fois que possible, sur la forme vernaculaire de la langue, à moins que le ou les pays utilisant la langue ne choisissent un codet de langue différent.

Les codes terminologique et bibliographique doivent être utilisés intégralement, et le choix du code que l'on utilise doit être précisé par les partenaires qui échangent de l'information avant qu'ait lieu l'échange. Les utilisateurs doivent faire référence à l'ISO 639-2/B pour le code destiné à des applications bibliographiques, et à l'ISO 639-2/T pour le code destiné à des applications en terminologie.

Afin d'assurer la continuité et la stabilité du code, les codets ne doivent être modifiés que pour des raisons valables. Après qu'un codet a été modifié, le codet précédent ne doit pas être réutilisé avant une période d'au moins cinq ans. Afin de se conformer aux besoins des grandes applications qui se développent continûment, un codet de l'ISO 639-2/B ne doit pas être modifié si le nom de la langue ou son abréviation se modifie.

Lorsqu'on adapte la présente partie de l'ISO 639 à des langues qui utilisent un autre alphabet (par exemple le cyrillique), les codets doivent être créés selon les principes de la présente partie de l'ISO 639.

#### 4.1.1 Codets de langue collectifs

Les codets de langue collectifs sont destinés aux langues pour lesquelles n'existent, ne devraient être écrits, enregistrés ou créés, qu'un nombre relativement faible de documents. Les termes *langues* ou *autres*, lorsqu'ils apparaissent dans le nom d'une langue dans les tableaux suivants, doivent être utilisés lorsqu'un codet de langue est un codet de langue collectif. Un codet de langue collectif ne doit pas être utilisé lorsqu'il existe un codet de langue individuel ou un autre codet de langue collectif plus spécifique.

## SS-ISO 639-2:2005

### 4.1.2 Special situations

The language code mul (for multiple languages) should be applied when several languages are used and it is not practical to specify all the appropriate language codes.

The language code und (for undetermined) is provided for those situations in which a language or languages must be indicated but the language cannot be identified.

### 4.1.3 Scripts and dialects

A single language code is normally provided for a language even though the language is written in more than one script. A separate standard may be developed for the purpose of designating information concerning the script or writing system of a language.

The dialect of a language is usually represented by the same language code as that used for the language. If the language is assigned to a collective language code, the dialect is assigned to the same collective language code. If the language has an individual language code, the dialect is also included in that code rather than the code for the group to which both belong. In a few instances, however, both the language and a dialect of that language have their own individual language codes.

### 4.1.4 Local Codes

Codes qaa through qtz are reserved for local use, including for local treatment of dialects. These codes may only be used locally, and may not be exchanged internationally.

### 4.1.5 Ancient languages

Ancient languages that are not given individual language codes are assigned the code for the major language collective to which each belongs, rather than the code for the modern language which evolved from the ancient language. For example, Old Frisian is assigned the language code gem for the language group Germanic (Other) instead of the language code fry for the modern language Frisian.

In cases of doubt, assistance may be sought from the Registration Authority.

### 4.1.2 Cas particuliers

Il est recommandé d'utiliser le codet de langue mul (pour langues multiples) lorsque plusieurs langues sont utilisées et qu'il est impossible en pratique de donner tous les codets de langues qui conviendraient.

Le codet de langue und (pour indéterminée) a été prévu pour les cas où l'on doit signaler une langue ou des langues mais où il n'est pas possible de les identifier.

### 4.1.3 Écritures et dialectes

Le code prévoit un codet de langue unique par langue, même si cette langue s'écrit en plusieurs écritures. Une norme particulière devrait être mise au point afin de qualifier l'information concernant l'écriture ou le système d'écriture des langues.

Le dialecte d'une langue est habituellement représenté par le même codet de langue que celui utilisé pour la langue. Si l'on donne à une langue un codet de langue collectif, on donne au dialecte le même codet de langue collectif. Si une langue a un codet de langue particulier, son dialecte est compris dans ce codet, plutôt que dans le codet du groupe linguistique auquel tous deux appartiennent. Toutefois, il existe des cas où la langue et le dialecte de cette langue ont tous les deux un codet de langue particulier.

### 4.1.4 Codets locaux

Les codets qaa à qtz sont réservés à une utilisation locale qui comprend aussi la façon de traiter les dialectes. Ces codes doivent être utilisés uniquement localement, et ne peuvent pas être échangés au niveau international.

### 4.1.5 Langues anciennes

On donne aux langues anciennes, pour lesquelles il n'a pas été attribué de codet particulier, le codet du groupe de langue le plus important auquel elles appartiennent, plutôt que le codet correspondant à la langue moderne qui a dérivé de cette langue ancienne. Par exemple, on attribue le codet gem, qui indique le groupe linguistique germaniques, autres langues au vieux frison, plutôt que le codet fry qui indique la langue moderne, le frison.

En cas de doute, on peut demander l'assistance de l'Agence d'enregistrement.



## 4.2 Registration of new language codes

The Registration Authority for this part of ISO 639 shall be the Library of Congress, Washington, D.C. 20540-4102 USA (c/o Network Development and MARC Standards Office)<sup>1)</sup>. The Registration Authority for ISO 639-1 is Infoterm, Simmeringer Hauptstrasse 24, A-1110 Vienna, Austria.

## 4.3 Application of language codes

Language codes can be used in the following specific instances. Examples of how language codes may be used follow each.

**4.3.1** To indicate the languages in which documents are or have been written or recorded.

### EXAMPLE

UNIMARC Format<sup>2)</sup>, Field 101 Language of the item  
A pamphlet is issued in Danish (language code dan used for Danish)  
Field 101: 0b\$adan

**4.3.2** To indicate the languages in which document-handling records (order records, bibliographic records, and the like) have been created.

### EXAMPLE

UNIMARC Format<sup>2)</sup>, Field 100 Language of cataloguing; positions 22-24  
A book is written in Dutch, but the catalog record is in English (language code eng used for English)  
Field 100 positions 22-24: eng

**4.3.3** To indicate the language-speaking capabilities of delegates to a meeting. (Alternatively, another International Standard list may be used, for example ISO 639-1.)

EXAMPLE — In a list of delegates issued at an ISO meeting, the codes eng, fre and rus indicate whether the delegates spoke in (eng) English, (fre) French, or (rus) Russian.

1) Subject to approval of ISO Council.

2) The UNiversal MACHine Readable Cataloging format is used for exchange of bibliographic data.

## 4.2 Enregistrement de nouveaux codets de langue

L'Agence d'enregistrement pour la présente partie de l'ISO 639-2 sera la Bibliothèque du Congrès, Washington, D.C. 20540-4102 USA (c/o Network Development and MARC Standards Office)<sup>1)</sup>. L'Agence d'enregistrement pour l'ISO 639-1 est Infoterm, Simmeringer Hauptstrasse 24, A-1110 Vienna, Austria.

## 4.3 Utilisation des codets de langue

Les codets de langue peuvent être utilisés dans les cas particuliers indiqués ci-après. Chaque cas est suivi d'exemples donnant la façon d'utiliser les codets.

**4.3.1** Ils servent à indiquer la langue dans laquelle un document est écrit ou a été écrit ou a été enregistré.

### EXEMPLE

Format UNIMARC<sup>2)</sup>, zone 101 Langue du document  
Une brochure est publiée en danois (le codet de langue dan est utilisé pour le danois)  
Zone 101: 0b\$adan

**4.3.2** Ils servent à indiquer la langue dans laquelle la notice représentant le document a été créée (notice d'acquisition, notice bibliographique et assimilée).

### EXEMPLE

Format UNIMARC<sup>2)</sup>, zone 100, positions 22-24 Langue de catalogage  
Un livre est écrit en néerlandais, mais sa notice de catalogage est en anglais (le codet de langue eng est utilisé pour l'anglais)  
Zone 100 positions 22-24: eng

**4.3.3** Ils servent à indiquer la compétence linguistique des délégués lors d'une réunion. (Alternative-ment, une autre Norme internationale peut être employée, par exemple l'ISO 639-1.)

EXEMPLE — Dans une liste des délégués produite lors d'une réunion de l'ISO, les codes eng, fre et rus indiquent que les délégués parlent anglais (eng), français (fre), russe (rus).

1) Doit être approuvé par le Conseil de l'ISO.

2) Le format UNIMARC (UNiversal MACHine Readable Cataloging) est employé pour l'échange de données bibliographiques.

## SS-ISO 639-2:2005

**4.3.4** To indicate the original language of a document.

EXAMPLE — Document of the United Nations:  
ST/DCS/1/Rev. 2 eng fre  
[Document is bilingual.]

### 4.4 Application of the country code

Country codes from ISO 3166 may be combined with language codes to denote the area in which a term, phrase, or language is used.

#### EXAMPLES

— a spool of thread (eng US)  
— a bobbin of cotton (eng GB)

## 5 Structure of the list of language codes

The list of language codes is presented in four tables:

Table 1: Alpha-3 codes arranged alphabetically by English name of language

Table 2: Alpha-3 codes arranged alphabetically by French name of language

Table 3: Alpha-3 codes arranged alphabetically by bibliographic code (ISO 639-2/B)

Table 4: Alpha-3 codes arranged alphabetically by terminology code (ISO 639-2/T)

**4.3.4** Ils servent à indiquer la langue originale d'un document.

EXEMPLE — Document des Nations Unies:  
ST/DCS/1/Rev. 2 eng fre  
[Le document est bilingue.]

### 4.4 Utilisation des codets de pays

Les codets de pays issus de l'ISO 3166 peuvent être combinés avec des codets de langue pour indiquer dans quelle zone on utilise un mot, une phrase ou une langue.

#### EXEMPLES

— a spool of thread (eng US)  
— a bobbin of cotton (eng GB)

## 5 Structure du code de langue

Les codes de langues sont présentés en quatre tableaux.

Tableau 1: Code alpha-3 classé alphabétiquement par le nom de la langue en anglais

Tableau 2: Code alpha-3 classé alphabétiquement par le nom de la langue en français

Tableau 3: Code alpha-3 classé alphabétiquement par codets bibliographiques (ISO 639-2/B)

Tableau 4: Code alpha-3 classé alphabétiquement par codets en terminologie (ISO 639-2/T)

**Table 1 — Alpha-3 codes arranged alphabetically by English name of language**  
**Tableau 1 — Code alpha-3 classé alphabétiquement par le nom de la langue en anglais**

<b>English name</b> <b>Nom anglais</b>	<b>French name</b> <b>Nom français</b>	<b>ISO 639-2/B</b> <b>Bibliographic code</b> <b>Codet bibliographique</b>	<b>ISO 639-2/T</b> <b>Terminology code</b> <b>Codet terminologique</b>
Abkhazian	abkhaze	abk	abk
Achinese	aceh	ace	ace
Acoli	acoli	ach	ach
Adangme	adangme	ada	ada
Afar	afar	aar	aar
Afrihili	afrihili	afh	afh
Afrikaans	afrikaans	afr	afr
Afro-Asiatic (Other)	afro-asiatiques, autres langues	afa	afa
Akan	akan	aka	aka
Akkadian	akkadien	akk	akk
Albanian <sup>3)</sup>	albanais <sup>3)</sup>	alb	sqi
Aleut	aléoute	ale	ale
Algonquian languages	algonquines, langues	alg	alg
Altaic (Other)	altaïques, autres langues	tut	tut
Amharic	amharique	amh	amh
Apache languages	apache	apa	apa
Arabic	arabe	ara	ara
Aramaic	araméen	arc	arc
Arapaho	arapaho	arp	arp
Araucanian	araucan	arn	arn
Arawak	arawak	arw	arw
Armenian <sup>3)</sup>	arménien <sup>3)</sup>	arm	hye
Artificial (Other)	artificielles, autres langues	art	art
Assamese	assamais	asm	asm
Athapascan languages	athapascanes, langues	ath	ath
Australian languages	australiennes, langues	aus	aus
Austronesian (Other)	malayo-polynésiennes, autres langues	map	map
Avaric	avar	ava	ava
Avestan	avestique	ave	ave
Awadhi	awadhi	awa	awa
Aymara	aymara	aym	aym
Azerbaijani	azéri	aze	aze
Balinese	balinais	ban	ban

3) For explanation of variant codes see 4.1.

3) Voir en 4.1 l'explication de la divergence des deux codets.