

国際連合および関連機関が作成する文書の最初のページには、A, C, E, F, R, S というラテン文字がつけられている。いうまでもなく Arabic, Chinese, English, French, Russian, Spanish の略称で、国際連合設立時に公文書に使用を認めた 6 言語識別子 (language identifier) である。

世界中で使われている言語は 5,000 とも 6,000 ともいわれる。これらの言語で書かれた語、用語、文書が当該言語圏を越えて流通するにつれて、それらにも言語識別子を付与する必要が高まった。2 桁のラテン小文字からなる言語識別子 140 種を含むコード*が ISO 639:1988 として制定され、その改訂版である ISO 639-1:2002 に含まれる言語識別子は 181 種に増加した。それでも収容しきれない言語を含めるため、3 桁のラテン小文字からなる言語識別子 438 種を含むコードが ISO 639-2:1998 として制定された。さらに言語関係者の強い要望にこたえ、現在のコードは下記の 6 部門への展開となった。このうち、ISO 639-1 および ISO 639-2 については、対応する日本工業規格 (JIS X 0412-1:2004 および JIS X 0412-2:2004) が制定されている。

ISO 639 Codes for the representation of names of languages

ISO 639-1:2002	Part 1: Alpha-2 code
ISO 639-2:1998	Part 2: Alpha-3 code
ISO 639-3:2007	Part 3: Alpha-3 code for comprehensive coverage of languages
ISO 639-4:2010	Part 4: Implementation guidelines and general principles for languages coding
ISO 639-5:2008	Part 5: Alpha-3 code for language families and groups
ISO 639-6:2010	Part 6: Alpha-4 code for comprehensive coverage of language variants

代表的なものを以下に示す。

	2 文字識別子		3 文字識別子	
			用語学用	書誌用
イタリア語	it		ita	ita
英語	en		eng	eng
韓国語/朝鮮語	ko		kor	kor
スペイン語	es		spa	spa
中国語 (漢語)	zh		zho	chi
ドイツ語	de		deu	ger
日本語	ja		jpn	jpn
フランス語	fr		fra	fre
ロシア語	ru		rus	rus

3 文字識別子には、用語学用 (T) と書誌用 (B) とが

ある。ほとんどの言語については同一で、異なるものは 24 言語だけだが、この中にはフランス語 (fra (T), fre (B)), ドイツ語 (deu (T), ger (B)), 中国語 (漢語) (zho (T), chi (B)) などが含まれる。どのように使い分けるかについての説明はなく、使用者の立場にまかしているようだ。ただし、混用を避けねばならない。

識別子は原則として当該言語での呼称に基づいて作成される。日本語については例外で、決定に当たって日本側が「ni」でなく「ja」でよいと回答したためである。3 文字識別子に書誌用 (T) があるのも、英語圏の強い要請があったためとみられる。

言語名の新規登録、言語識別子の変更についての審議機関として、2 文字については Austrian Standards Institute が、3 文字については Library of Congress が担当している。日本はアイヌ語を追加申請し、「ain」が追補版に収載された。

われわれの周辺では 2 文字識別子でほとんどの用が足りるが、使用範囲の狭い言語および古語については 3 文字識別子を使うことになる。ハワイ語 (haw)、中期英語 (1100-1500) (enm)、古高ドイツ語 (ca.750-1050) (goh)、中高ドイツ語 (ca.1050-1500) (gmh) などである。地域によって使い方が異なる場合には、地域識別子を付与する。

実例をいくつか挙げておこう。

- (1) *en* ion
fr ion
de Ion
it ione
ja イオン
- (2) elevator (en US)
lift (en GB)
ascenceur (fr)
エレベータ (ja)
- (3) corn (en GB) = wheat (en)
corn (en US) = maize (en)

言語識別子を文中で使用する場合には、括弧でくくるか、または斜体で表記することをすすめる。

*ISO 639-1 の定義によれば、コード (code) は表であり、それぞれの文字列は識別子 (identifier) と呼ばれる。国名のコードでは、後者を「コード要素」と称している。ISO でも JIS でもコードの定義は複数存在し、使い方に注意する必要がある。

(太田泰弘)