

## Digital Cinema Language Codes

This document correlates the language codes used in the [Digital Cinema Naming Convention](#) to valid [RFC 5646](#) tags. SMPTE 429-7 specifies use of xs:language tags. [W3.org](#) specifies that xs:language tags use the definitions in [RFC 3066](#). However, RFC 3066 has been replaced by [RFC 4646](#), which has been replaced by [RFC 5646](#). The “best current practice” of RFC 5646 is reflected in [BCP 47](#). Recommended language tags are listed in the [IANA Subtag Registry](#). Note that BCP 47 2.1.1 states “At all times, language tags and their subtags, including private use and extensions, are to be treated as case insensitive: there exist conventions for the capitalization of some of the subtags, but these MUST NOT be taken to carry meaning.” “Although case distinctions do not carry meaning in language tags, consistent formatting and presentation of language tags will aid users. The format of subtags in the registry is RECOMMENDED as the form to use in language tags.” Note also that IANA has separate tags for use of different scripts for different written representations of the same spoken language. This may be important in portions of the CPL that deal with written text (such as subtitles or captions), but not applicable to spoken language (such as main sound). Another useful resource is Richard Ishida’s [Language Tag Lookup](#).

Language	Digital Cinema Naming Convention (Only used in content title text and annotation text)		RFC 5646	
	Language Code	Notes	xs:language tag	Notes
Albanian	SQ		sq	
Arabic	AR		ar	
Bosnian	BS		bs	
Bulgarian	BG		bg	
Catalan	CA		ca	
Chinese – Cantonese	YUE		yue	
Chinese – Mandarin PRC	CMN	(Audio only)	cmn	
Chinese – Mandarin Simplified	QMS	(Subtitles only)	cmn-Hans	
Chinese – Mandarin Traditional	QMT	(Subtitles only)	cmn-Hant	
Chinese - Taiwanese	NAN	(Audio only)	nan	
Chinese – Taiwanese Mandarin	QTM	(Audio only)	cmn-TW	

Croatian	HR		hr	
Czech	CS		cs	
Danish	DA		da	
Dutch	NL		nl	
English	EN		en	
Estonian	ET		et	
Euskara	EU		eu	
Finnish	FI		fi	
Flemish	VLS		vls	Vlaams
French	FR		fr	
French – Canadian	QFC		fr-CA	
German	DE		de	
German – Swiss	GSW		gsw	
Greek	EL		el	
Hebrew	HE		he	
Hindi	HI		hi	
Hungarian	HU		hu	
Icelandic	IS		is	
Indonesian Bahasa	IND		id	Bahasa means language
Italian	IT		it	
Japanese	JA		ja	
Kazakh	KK		kk	
Korean	KO		ko	
Kyrgyz	KG		ky	
Latvian	LV		lv	
Lithuanian	LT		lt	
Malay Bahasa	MSA		zlm	
Mongolian	MN		mn	
Navajo			nv	
Norwegian	NO		no	
Polish	PL		pl	
Portugese – Brazilian	QBP		pt-BR	
Portugese – European	PT		pt	
Romanian	RO		ro	
Russian	RU		ru	
Serbian	SR		sr	
Slovak	SK		sk	
Slovenian	SL		sl	
Spanish – Argentina	QSA		es-AR	

Spanish - Castilian	ES		es	
Spanish – Latin America	LAS		es-419	Region 419 is Latin America and the Caribbean
Spanish – Mexican	QSM		es-MX	
Swedish	SV		sv	
Tamil	TA		ta	
Telugu	TE		te	
Thai	TH		th	
Turkish	TR		tr	
Ukrainian	UK		uk	
Urdu	UR		ur	
Vietnamese	VI		vi	
Welsh	WEL		cy	

## Region Subtags

RFC 4646 states “Region subtags are used to indicate linguistic variations associated with or appropriate to a specific country, territory, or region. Typically, a region subtag is used to indicate regional dialects or usage, or region-specific spelling conventions. A region subtag can also be used to indicate that content is expressed in a way that is appropriate for use throughout a region, for instance, Spanish content tailored to be useful throughout Latin America.” The language tag list above lists tags for those languages in the [Digital Cinema Naming Convention](#). For example, there are separate tags for Spanish (es) and Latin American Spanish (es-419). Other language tags with region subtags may be used when it is necessary to distinguish between different regional variations of languages (such as en-US and en-GB).

SMPTE 429-16:2014 (Additional Composition Metadata and Guidelines) notes “Matching desired playback language with language tags is not trivial, IETF RFC 4647 specifies practices for such matching.” Playback systems must consider that a show playlist (such as that in a SMPTE 430-11 Auxiliary Content Resource Presentation List) may represent the concatenation of several compositions. For example, a trailer might be have a language attribute of “en,” while the feature might be “en-US.” A playback system must not send these to separate caption streams (or language channels) such that the user would only see captions for the trailer on one channel and the feature on another.

ISDCF recommends that the SYSTEM (server/closed caption rendering) implement a method to direct outputs of a class of languages to a specific channel (e.g., channel 1 is Spanish, channel 2 English, etc.) for each theater. The channel would have a set of rules

for language tag priority (in case there are two region subtags for the same language). The selection would be part of the UI of the system.

SMPTE 429-16:2014 table B.1 provides a list of language tags similar to that above.

## Revision History

- 2011-11-29 – Original version
- 2017-08-07
  - Updated link to digital cinema naming convention
  - Removed reference to Richard Ishida's [Language Tag Lookup](#), since the link no longer works.
  - Added language tags (based on digital cinema naming convention) for Euskara, Kazakh, Kyrgyz, Mongolian, Urdu, and Welsh.
  - Added Region Subtags section
  - Added revision history
- 2017-08-09
  - Added Richard Ishida's Language Tag Lookup back with new URL.
- 2017-08-0a
  - Clarified language priority evaluation on a reel by reel basis.
- 2017-08-29
  - Added Navjo language.
- 2017-10-19
  - Clarified second paragraph of Region Subtags.
  - Added ISDCF recommends paragraph.
  - Added "Only used" text in table header.