

世界の「言語天文台」

高橋 智恵*・中平 勝子**・三上 喜貴***

Language Observatories in the World

Tomoe TAKAHASHI*, Katsuko T. NAKAHIRA**, Yoshiki MIKAMI***

Key words : Language Observatory, Digital-Divide, Internet, UNESCO

1. はじめに

独立行政法人・科学技術振興機構（JST）の「社会技術研究プログラム（RISTEX）」の公募研究テーマのひとつとしての選定を受け、筆者らは2003年10月から「言語間デジタルディバイドの解消を目指した言語天文台の創設」に関する研究プロジェクトを開始した。

世界では6,000を超える言語が話されているといわれるが、そのうちデジタルネットワーク上で不自由なく利用できる言語は限られており、デジタルネットワーク上に存在する知識へのアクセス、コミュニケーション手段へのアクセスには言語間に大変な格差が存在する。筆者らはかねてよりこの問題についての調査研究を進めてきたところであるが¹、言語間デジタルディバイドの実態を客観的かつ詳細に明らかにすることのできる「観測手段」としての言語天文台を構築し、あわせてその克服策を技術面、政策面から明らかにしようというのが本研究プロジェクトの目的である。

「言語天文台」なる概念、用語は本研究プロジェクトが初めて用いたものであり、上述の研究目的からも明らかとなっており、デジタルネットワーク上での言語活動を観測する、バーチャルな世界における天文台として構築されたものである。しかし、世界を見渡すと、言語天文台に相当する名称や類似の活動を行

原稿受付：平成16年5月20日

*科学技術振興機構、社会技術研究システム

**長岡技術科学大学、eラーニング研究実践センター

***長岡技術科学大学、経営情報系

1 三上喜貴、言語観デジタル・ディバイド、長岡技術科学大学言語人文科学論集、Vol.14、pp.83-94、2000

っている組織やグループはいくつか存在する。本稿ではこうした世界の「言語天文台」関連活動に関する我々の調査結果を報告するとともに、我々の研究活動の持つ意義について考察する。

1.1 言語間デジタルディバイドの実態

本研究プロジェクトの開始にあたり、2004年2月20日・21日に第1回言語天文台ワークショップ（FLOW2004: The First Language Observatory Workshop 2004）が、インド、タイ、マレーシア、シンガポール、ハンガリー等の研究者を招いて開催された。その際、ハンガリーから出席した Ms. Robin Lee Nagano² は、ネットワーク上での情報アクセスに関して生じている言語間の格差を端的に指摘する報告を行った。それは、英語、ドイツ語、ハンガリー語、日本語の各言語で共通のキーワードについて Google 検索を行いネット上で得られる情報量の言語間格差について調査したものである。以下にその一例を示す。

Table 1 ある映画についての言語別ヒット件数

<i>Search term (s)</i>	<i>language</i>	<i>#of hits (approx)</i>
Amelie (title) + Jeunet (director's name)	English	11,800
Die fabelhafte Welt der Amélie (title)	German	49,800
Amelie csodálatos élete (title)	Hungarian	540
アメリ (title)	Japanese	64,100

注：2001年フランスで作成されたラブストーリー映画。原題は 'Le Febuleux destin d'Amelie Poluain'、監督は Jean-Pierre Jeunet（ジャン・ピエール・ジュネ）

Table 2 キャベツの調理法についての言語別ヒット件数

<i>Search term (s)</i>	<i>language</i>	<i>#of hits (approx)</i>
cabbage + (recipe or cook)	any	1,510,000
cabbage + (recipe or cook)	English	511,000
Kohl + (Rezept or kochen)	German	101,000
Kraut + (Rezept of kochen)	German	71,700
Káposzta + (recept or főz)	Hungarian	1,190
Kelkáposzta + (recept or főz)	Hungarian	438
キャベツ + (調理法 or 作る)	Japanese	829

2 the Language Teaching Centre, University of Miskolc, Hungary

Table 3 風力エネルギーについての言語別ヒット件数

<i>Search term (s)</i>	<i>language</i>	<i>#of hits (approx)</i>
wind power	any	982,000
wind energy	any	718,000
wind power	English	937,000
wind energy	English	674,000
wind power	German	2,390
wind energy	German	5,860
Windenergie	German	136,000
wind power	Hungarian	26
wind energy	Hungarian	65
szélenergia	Hungarian	1,270
wind power	Japanese	1,310
wind energy	Japanese	892
風力エネルギー	Japanese	6,620

Table 4 ある疾患名についての言語別ヒット件数

<i>Search term (s)</i>	<i>language</i>	<i>#of hits (approx)</i>
FSGS(+focal or segmental or glomerulosclerosis)	any	3,160
FSGS(+focal or segmental or glomerulosclerosis)	English	2,920
FSGS(+focal or segmental or glomerulosclerosis)	German	23
FSGS(+FoKal or segmentale or Glomerulosklerose)	German	57
FSGS(+focal or segmental or glomerulosclerosis)	Hungarian	7
FSGS(+focalis or segmentalis or glomerulosclerosis)	Hungarian	9
FSGS(+focal or segmental or glomerulosclerosis)	Japanese	39
単状（分節状）系球体硬化症	Japanese	764

注：単一の疾患ではなく、多彩な原疾患の経過中に起こる共通の臨床症状を示す疾患群である。症状は高度のタンパク尿、低タンパク血症を必須条件として、その結果生ずる高脂血症を特徴とする。

Robin はある話題映画、キャベツの調理法、風力エネルギー、FSGS という特殊疾患名という4つの検索単語による検索結果を比較し、映画や調理法などの日常生活に関連するトピックであれば、言語間の大きな格差はなく十分情報収集が可能であるが、より専門的な分野では英語での情報発信が他の言語よりも群を抜いて多いという事実を強調し、この傾向は今後も変わることなく、英語

は専門家たちの第一の情報交換手段として君臨し続けるだろうと予想している。

こうした言語間の情報量格差は、言語別の利用者数の相違を反映したものである。インターネット上のマーケティング関連調査を行っているある調査機関の推計によれば、2004年3月時点における世界のインターネット利用者数は約7.3億人で、そのうち英語による利用者が2.87億人（35.8%）、英語以外のヨーロッパ言語による利用者が2.76億人（37.9%）であり、英語を含めヨーロッパ言語だけで世界の利用者の70%以上を占めている（Fig. 1）。この調査機関が行っている言語別のウェブページ数についての推計によれば、世界中のウェブページの68.4%は英語によって書かれたものである。

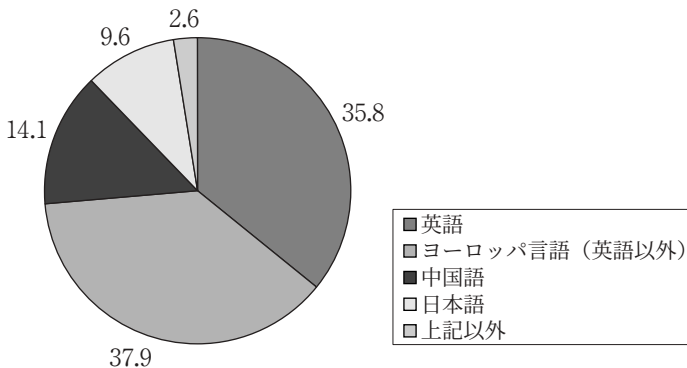


Fig. 1 言語別インターネット利用者 (%)

出典：Global Internet Statistics

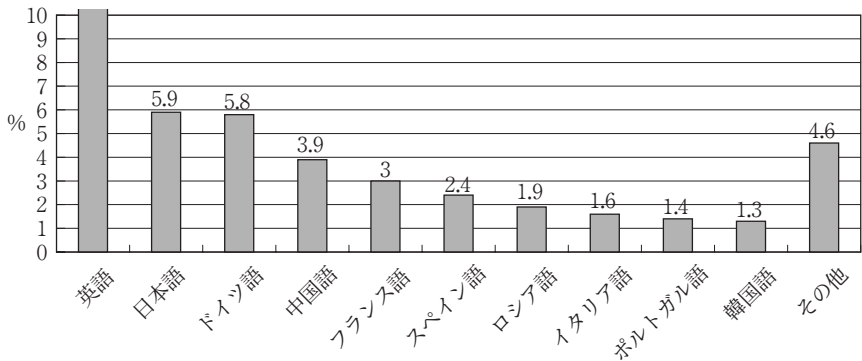


Fig. 2 言語別インターネットページ数割合

注：英語の割合は68.4%である。また全体のWeb page数は313Bとされている。

出典：前掲

しかしながら、こうした指摘は言語間デジタルディバイドの実像の一部を示すものでしかない。冒頭に述べたように、地球上には6,000以上の言語の利用者があり、その中には、デジタルネットワーク上で母国語を表記することすらできない利用者が依然として数多く存在している。一例として、国連の「世界人権宣言」(Universal Declaration of Human Rights)を取り上げてみよう。同宣言は地球上のすべての人々に等しく読まれることを意図して宣言されたものであり、国連人権高等弁務官のウェブサイトには、同宣言が327言語に翻訳され、掲載されている³。しかしながら、各言語に翻訳されたページを実際に閲覧してみると、すべての言語への翻訳文がデジタルな符号化文字列として表現されているとは限らない。特に、非ラテン文字を利用する72言語への翻訳ページ中、符号化文字列として表記されているのはロシア語、中国語、日本語、韓国語などわずか15言語にすぎず、大多数は単に画像として貼り付けられているもの(52言語)および文字化けして判読不能なもの(5言語)であった。これらの言語・文字の利用者は、仮にネットワーク環境やパソコンの利用に問題がなくても、依然としてネットワーク上で母国語を利用することができないのである⁴。この事実は、筆者らが言語天文台の活動を通じて探求しようと考えている言語間デジタルディバイドの実態を端的に示すものであり、本稿末に各言語に翻訳された人権宣言の表記状況一覧をまとめて示した。

2. 言語天文台に関連する研究チームについての調査

次に、主としてインターネット上で検索可能な既存の研究結果・研究グループを調査対象とし、一部現地調査の結果も加えながら、言語天文台と類似する活動を志向しているもの、または活動内容は異なるが language observatory と同様な名称を持つ組織やグループについて調査した結果を紹介する。

2.1 フランス語圏における類似研究

フランス語は1999年の統計によれば母国語としての話者数が世界全体で7,200万人とされており、国連公用語のひとつとして、英語に次ぐ国際語の地位を得ている。フランス語利用者人口の多い諸国家の国際組織として "L'Organisation internationale de la Francophonie" がある⁵。この国際組織には世界48カ国が参加し

3 <http://www.unhchr.ch/udhr/>

4 筆者らは、この問題が広範に存在しているインドの言語情報処理技術関係者に対してこの問題を指摘した。

5 <http://www.francophonie.org/> 参照

ているが、これは国の数という指標で見て英語圏、アラビア語圏に次ぐ規模である。

しかし、こうした国際語としての地位、特にインターネット上での地位は、今日、英語によって大きく脅かされつつあり、こうした危機感を反映して、言語天文台に相当する活動が最も顕著に見られる言語である。フランス語で言語天文台を表す "observatoire linguistique" をキーワードとして Google 検索を行うと、240件のヒットがあった⁶。その大半は以下で示す二つの研究グループに関するものである。

2.1.1 Research group Observatoire de linguistique Sens-Texte (OLST)

1998年にカナダ・モントリオール大学内に創設されたこの組織の主な活動目標は、言語学における意味-文字データ理論 (the Meaning-Text Theory) に基づいた調査、研究を促進させることである。フランス語のコンピュータ言語の構築や、形式化された専門用語の研究等に携わっているが、特にフランス語に限った調査、研究というわけではない。テキストの語形成や分析を自動的に行えるような Meaning-Text モデルの開発も進められており、コンピュータサイエンスに関わる分野での研究にも携わっている。2004年3月にはワークショップも開催しており、他の言語活動関連組織とも連携を図っているようある。

2.1.2 Réseau international des observatoires francophones de l'inforoute et du traitement informatique des langues (RIOFIL)

この組織RIOFILは新しい情報通信技術を活用するフランス語利用者を支援するためのデータ加工ツールの作成、普及を目指して設立されたものであり、アフリカ、フランス、カナダ、スイス、ベルギーと世界にまたがる6つのグループが存在する。フランス語利用者が新しい情報通信技術を迅速に利用できる環境を整備することは、世界の言語的、文化的多様性を維持発展させていく上で極めて重要であるとの認識のもと、情報通信技術のフランス語環境へのローカライゼーションや、フランス語圏における高速ネットワークの導入を積極的に推進してきたが、RIOFIL の公式ホームページが最後に更新されたのは1999年1月22日であり、それ以降現在に至るまで RIOFIL 単独の活動の形跡は見られない。

しかし、2000年には、RIOFIL と同様な活動を行っていた RINT⁷ と呼ばれる組

6 2004年3月7日現在

7 International Network of Neology and Terminology

織を統合し、RIFAL (Réseau international francophone d'aménagement linguistique) を新たに発足させた。この組織はAIF (L'Organisation internationale de la Francophonie) の一組織として言語活動に関わる政策をサポートする役割をはたしており、引き続き新しい情報通信技術の普及に力を注いでいる。

RIFAL の特徴として特筆しておくべき点は、南半球の国々において、高速ネットワークを導入し言語コミュニケーションや情報技術に関わるデータ加工処理システムの構築を目指しているということである。社会言語学やその分野に関連する記事を電子上で配信している Noves SL, Journal on Sociolinguistics でアフリカ言語を取り上げた記事の中でも、RIFAL の活動内容について言及されている⁸。RIOFIL と RINT の統合以前にも、1996年にフランス語圏の国際協力機関ACCT (Agence de coopération culturelle et technique) の要請を受け、両組織が協力し3年計画で南半球におけるフランス語の運用能力高上を目指すプロジェクトを実施された。

新組織 RIFAL の活動範囲はフランス、カナダ、ベルギーといった先進国地域のフランス語圏のみに止まらず、ベナン共和国、コンゴ共和国、モロッコといったアフリカ諸国を含め世界21都市で調査活動等を行っている。組織運営に関わる機関は現在、フランス、ケベック、ベルギー、ハイチ、モロッコ、そして中央アフリカと西アフリカの7箇所に置かれ、フランス語のさらなる地位向上を目指した活動に取り組んでいる。

2.2 カタロニア語圏における類似研究



カタロニア語はスペイン北東部のカタロニア地方で話されている言語であり、話者数は約6.5百万人とされているが、ユネスコの調査では絶滅危惧言語とされている。カタロニア地方では、カタロニア語文化の保護のために、外国映画は可能な限りカタロニア語に吹き替える、各大学においてカタロニア語の授業を実施する、などの努力がなされているものの、例えば、昨今では学校教育におけるカタロニア語の履修率は低下してきている等の問題が指摘されている⁹。

8 <http://cultura.gencat.net/llengcat/noves/hm01setiu/internacional/>

9 private discussion with Universitat de Barcelona Staff and Miquel Strubell

2.2.1 L'Observatori de la llengua catalana

スペイン語で言語天文台を表す "L'observatori de la llengua" を検索語とした Google 検索の結果は300件のヒットを得た¹⁰。これらの記事の多くはカタロニア言語天文台 (L'Observatori de la llengua catalana) の活動に関するものであった。この組織はカタロニア文化圏において指導者的役割を果たす文化組織として自らを位置づけている。主な活動としては、社会でのカタロニア語使用に関する調査、またこの言語の文法的な変遷についての継続的な観測の2点が挙げられる。その調査結果を報告書としてまとめ、国際社会に問題提起しており、最近の報告では、2004年1月20日に開催された欧州理事会へ地方言語、少数言語に関する調査結果を提示し、ヨーロッパ憲章の適用を求め、スペインにおけるカタロニア語迫害を公然と批判した。また、2004年2月24日にはジュネーブの人種差別撤廃委員会へカタロニア地方の人々に対する人種差別の実態と、スペインにおけるカタロニア語の現状を報告した。

この組織の主たる目的はカタロニア語圏の人々が少数派として受けている社会的迫害の排除であり、同様な立場にあるヨーロッパの少数言語も調査対象として捉えているようである。しかし、カタロニア語の社会的地位の拡大こそが最重要目標であり、彼らの活動を一民族として自分たちの言語、民族を守る戦いと捉えることも可能ではないだろうか。

2.2.2 WICCAC (Webmasters Independents en Català de Cultura i d'Àmbits Cívics,)

この組織の名称を英訳すれば Independent Webmasters in Catalan of Culture and Civic Scopes となり、文字通り、カタロニア語サイトを運営する Web Master 達の連合組織である。その目的は、メンバーである Web Master 達が円滑な意見交換を行える環境を整え、Web Page の改善を行い、インターネットにおけるカタロニア言語の円滑な利用を促進することにある。

その活動の一環として、カタロニア語圏の Web サイトにおけるカタロニア語で記載されたサイトの調査を行っており、彼らのサイト¹¹にて、各ジャンル別の Web サイト、サイト内において利用されている言語の調査結果を毎週公表している。2004年8月中旬に筆者の一人が現地調査した結果では、大学機関や劇場、ラジオ・テレビ局などではほぼカタロニア語 Web サイトを用意しているが (90%

¹⁰ 2004年3月7日現在

¹¹ <http://wiccac.org/resum.html>

前後)、ビジネス関係ではあまりカタロニア語 Web サイトが用意されていない(50%未満)という報告が挙げられている。

2.2.3 SoftCatala¹²

元来はテクノロジー関係の団体であり、特にフリーウェアのカタロニア語化につとめている。その一方で、インターネット上でのカタロニア語利用状況についての調査なども行っており、最近では、カタロニア語で書かれた Web ページ数は世界で第23位を占めている(約300万ページ)、カタロニア語で書かれた Web ページ数は対カタロニア語(を話す)人口比0.45であり、これは世界で第20位である、などと報告している¹²。しかし、これらは AllTheWebなどを一次情報源とするものであり、独自の調査活動を行っているわけではない。

2.3 バスク語圏における類似研究

同様にスペイン、フランスにまたがるバスク地方で使用されているバスク語を対象とした組織として Hizkuntz Eskubideen Behatokia (英: The Observatory of Linguistic Rights, Behatokia)がある。バスク語を話す人々の言語権の保護が目的で、その活動範囲は公的なものに止まらず、個人レベルまでも保護対象としている。

バスク語使用の際に個人的な基本権利を迫害された場合も、この組織を通じて訴えることが出来るようだ。

スペインでは、バスク語(Euskara)は公用語として認められておらず、よってバスク語を話す人たちの言語権は常に迫害されているといえる。そこでこの組織ではまず保健、教育、裁判、情報収集等の基本的な公的サービスをバスク語で受けられる制度の確立を目指している。

バスク語の場合もカタロニア語と同様に、基本的な生活レベルで言語的迫害を受けていることが大きな問題であり、われわれ言語天文台が取り扱う問題よりは、人々にとってより差し迫ったものであるかもしれない。しかし、Web 上においては Windows の言語オプションにもカタロニア語、バスク語ともに含まれている。この事実からも、ここで取り上げたカタロニア語、バスク語を調査対象としている2つの組織は、言語的には少数派であるが、文字コードの面から捉えればヨーロッパ言語の恩恵を大いに受け、自分たち民族が直面している言語問題の現状を全世界にむけて情報として発信できているといえるのではな

¹² <http://www.softcatala.org/articles/article26.htm>

いだろうか。

2.4 英語圏における類似研究

次に英語圏での現状をまとめる。検索エンジン Google を利用し language-observatory で検索したところ63件のヒットを得た¹³。しかし、このうちの大半が筆者らの言語天文台プロジェクトに関するものであった。ここまで取り上げてきた他の言語とは違い、社会的迫害とはまったく無縁である英語は、自ら何らかの行動を起こさなくとも、言語使用においては世界中で不自由なく通用する。それは Web 上でも同様である。英語圏から筆者らのプロジェクトに賛同を表明する研究機関、団体等が現れないのも十分納得のいく結果である。しかし、グローバルな視点で言語を観測している組織が英語圏にも存在している。その中でもここで特筆すべき組織を次に紹介する。

2.4.1 The Linguasphere Observatory

1983年に政治や宗教との関連を一切持たず、多言語主義の下でグローバルな調査を行う非営利組織として Observatoire Linguistique がフランスに発足した。発足当時から世界中の言語、方言に関する情報収集に努め multilingualism の世界的な普及を目標としており、この組織がウェールズに移転し The Linguasphere Observatory として現在も活動を続けている。世界言語の多様性の擁護を目的とする UNESCO の Linguapax プロジェクトにも協力しており、自らを多国間に及ぶ言語環境の観察者と位置づけ、サイバースペース上の観測プラットフォームと呼んでいる。この組織による最初の報告書が2000年に発表されたが、世界の言語および方言の前代未聞の詳しいカタログとしてグローバルな言語の記録を世界に提供した。

2.5 中南米における類似研究

ドミニカ共和国に母体を置く国際 NGO 組織の FUNREDES¹⁴ は発展途上国への情報技術の普及を目指した研究、調査活動に取り組んでいる。この組織内に、1996年から2001年にかけて Observatory of the languages and the cultures という名のもとで、特にインターネット上の言語と文化に焦点をあてた活動に携わるグループが編成された。約5年間の活動中に、言語に関する研究調査結果を5回、

13 2004年4月26日現在

14 ASSOCIATION RÉSEAUX ET DÉVELOPPEMENT

文化に関しては3回公表している。調査対象は主に英語、フランス語、スペイン語と、その文化圏であるが、その中でも特に Web 上でのスペイン語の普及に関する調査に力を入れている。

しかしながら、このグループの5年間の調査は、Web 上での英語の使用に対して、フランス語やスペイン語の割合はどのように変化していくのかという結果に止まっている。同様な他の調査と比べるとより細部にまでわたり比較されているものの、それは調査段階で終結してしまったようにも感じられる。ただ、FUNREDES としては、特定の言語に限った情報収集の手段や、個人の楽しみとしてのインターネットの普及を目指しているのではなく、多くの個人、そしてその個人を構成員とする地域が、言語や文化の多様性をより認識し受容するために情報技術の普及が必要だと強調している。今現在では、Observatory of the languages and the cultures に続く活動が表立ったところでは発見できていないが、先の5年間の調査結果を基礎とし、今後さらにデジタルネットワーク上で言語の多様性に関わる研究に携わっていくことを期待する。

2.6 その他の地域での類似研究

まずはアジア・アフリカ地域の類似研究であるが、顕著な活動が見受けられるとはいえないものの、Regional Observatory of The Information Society in Asia Pasific (ROISAP) が存在している。アジア・アフリカと広大な地域を一括りで捉えてしまうことは、文化的背景からも、もちろん言語の面から考えても適切であるとはいえないが、この組織以外で類似研究といえるものはアジア・アフリカ地域では発見できなかった。アフリカ地域から、われわれのプロジェクトに関して多くの関心を寄せられている事実からも、いまだアジア・アフリカ地域にいたっては、デジタルネットワーク上での言語、特に自らが直面しているデジタルディバイドに関する認識は低いと推測できる。

次にロシア語圏では、обсерватория Информационного Общества という組織の存在を認めることができた。しかし実際のところ彼らのホームページからは活動内容等について情報を得ることはできなかった。あくまで予想でしかないが、組織の設立は最近であり、ロシア語圏ではネットワーク上で情報発信手段としての役割を果たす言語に関する認識がようやく高まってきたのではないだろうか。

3. まとめ

3.1 既存の言語天文台活動

以上のように、“Language Observatory”という名のもとで活動する組織は地球上にいくつか存在しているが、その活動内容は必ずしも筆者らの構想する言語天文台の活動と一致するものではない。まず、対象言語については、関連言語を視野に入れてはいるものの特定の関心言語に重きが置かれているものがほとんどであり、また、活動内容も、バスク語に関する Hizkuntz Eskubideen Behatokia のように当該言語圏の基本的な人権の保護—デジタルネットワークに限らず市民生活における当該言語の利用範囲を広げるためのあらゆる活動を対象として言語権の確立—を目指そうとするもの、フランス語に関する OLST のように言語に関する学術的研究そのものを目的とするもの、など多様な広がりを持つ。

しかし、フランス語に関する RIFAL の活動、カタロニア語圏における L'Observatori de la llengua catalana や WICCAC の活動、中南米地域における FUNREDES の活動などは、サイバースペース上における当該言語の利用促進を目標とするという点において言語天文台の活動目標ときわめて近いものがあり、部分的には、われわれの目指す言語別ウェブページ調査（Cyber Language Census）のような調査活動をすでに小規模ながら実施していることも確認された。FUNREDES のように、すでに筆者らに対してコンタクトを求めてきた組織もあるが、目的を共有しえる組織との連携を推進していくことが必要である。

3.2 UNESCO「サイバースペース勧告」と「言語天文台」

言語間デジタルディバイド問題への対応は国際機関にとっても重要な課題であるとの認識が次第に広がりつつある。2003年秋に開催された第32回 UNESCO 総会では、「多言語主義の促進及び使用並びにサイバースペースへの普遍的アクセスに関する勧告」（Recommendation concerning the promotion and use of multilingualism and universal access to cyberspace）が採択され、「世界規模の情報ネットワークにおける言語の多様性及びサイバースペースにおける情報への普遍的アクセスは現代の議論の核心であり、かつ、知識を基盤とした社会の発展における決定要因となり得る」との認識を明らかにした¹⁵。

同勧告は、問題に対処するための具体的な課題のひとつとして、「ユネスコは、他の国際機関と協調し言語情報処理における技術革新を含めた多言語主義、多言語リソース、アプリケーションに関する既存の製作、規則、技術的勧告及び

15 ほぼ同様の内容が、第31回 UNESCO 総会の決議として採択されている。

最善の実施事例について共同オンライン観測所（collaborative online laboratory）を設立すべきである」と述べている¹⁶。ここでは、言語天文台という表現こそ使用されていないものの、その趣旨からすれば、同勧告の提言する「オンライン観測所」の概念は、筆者らの構想する言語天文台にきわめて近いものといえる。実際、第1回言語天文台ワークショップに UNESCO 関係者の出席を求めたところ、共通の問題意識を有するものであることが相互に確認され、今後、本プロジェクトは UNESCO と密接に連携をとりながらプロジェクトを進めることが合意されたところであり、今後、前項で述べた世界各地の関係組織を含めた共同研究ネットワークを構築していくことが求められている。

16 勧告第5パラグラフ。UNESCO, in cooperation with other international organizations, should establish a collaborative online observatory on existing policies, regulations, technical recommendations, and best practices relating to multilingualism and multilingual resources and applications, including innovations in language computerization.

ISO 639-2	ISO name	Script	UDHR name	Usage Area	Form
abk	Abkhazian	Cyrillic	Abkhaz	Georgia, Ukraine	code
ace	Achinese	Latin	Achehnese	Sumatra	code
sai	South American Indian (Other)	Latin	Achuar-Shiwiari	Peru, Ecuador	code
nic	North American Indian	Modified Latin	Adja	Benin, Togo	code
afr	Afrikaans	Latin	Afrikaans	South Africa, Namibia	code
sai	South American Indian (Other)	Latin	Aguaruna	Peru	code
aka	Akan	Modified Latin	Akuapem Twi	Ghana, Ivory Coast	code
alb	Albanian	Latin	Albanian	Albania	code
sai	South American Indian (Other)	Latin	Amahuaca	Peru, Brazil	code
sai	South American Indian (Other)	Latin	Amarakaeri	Peru	code
amh	Amharic	Amhar	Amharic	Ethiopia	PDF
		Latin	Amuesha-Yanesha	Peru	code
sai	South American Indian (Other)	Latin	Arabela	Peru	code
ara	Arabic	Arabic	Arabic	Africa and Asia	code
arm	Armenian	Armen	Armenian	Armenia	PDF
		Modified Latin	Asante	Ghana	code
sai	South American Indian (Other)	Latin	Ashaninca	Peru	code
sai	South American Indian (Other)	Latin	Asheninca	Peru, Brazil	code
asm	Assamese	Assamese	Assamese	Assam	image
sem	Semitic (Other)	Assyrian	Assyrian	Iraq, Syria, Turkey	image
ast	Asturian; Bable	Latin	Asturian	Spain	code
aym	Aymara	Latin	Aymara	Bolivia, Peru	code
aze	Azerbaijani	Cyrillic	Azeri	Azerbaijan	PDF
		Latin	Azeri		PDF
ban	Balinese	Latin	Balinese	Bali	code
bal	Baluchi	Balochi	Balochi	Pakistan, Iran	image
bam	Bambara	Modified Latin	Bambara	Mali, Ivory Coast	code
nic	Niger-Kordofanian (Other)	Modified Latin	Baoule	Ivory Coast	code
bsq	Basque	Latin	Basque	The Basque Provinces	code
nic	Niger-Kordofanian (Other)	Modified Latin	Batonu	Benin, Nigeria	code
bel	Belarusian	Cyrillic	Belorus	Belarus	code
bem	Bemba	Latin	Bemba	Zambia, Tanzania	code
ben	Bengali	Bengal	Bengali	Bangladesh	image
bnt	Bantu (Other)	Latin	Beti	Cameroon	code
bho	Bhojpuri		Bhojpuri	India, Nepal	illegal
bis	Bislama	Latin	Bichelamar	Vanuatu	code
bik	Bikol	Latin	Bikol/Bicolano	Luzon	code
sai	South American Indian (Other)	Latin	Bora	Peru, Colombia	code
bos	Bosnian	Cyrillic	Bosnian	Bosnia-	code
		Latin	Bosnian	Herzegovina	code
bre	Breton	Latin	Breton	France	code
bug	Buginese	Latin	Buginese	Sulawesi	code
bul	Bulgarian	Cyrillic	Bulgarian	Bulgaria	code
bur	Burmese	Burmese	Burmese/Myanmar	Myanmar	PDF
myn	Mayan languages	Latin	Cakchiquel	Guatemala	code
sai	South American Indian (Other)	Latin	Campa pajonalino	Peru	code
sai	South American Indian (Other)	Latin	Candoshi-Shapra	Peru	code
sai	South American Indian (Other)	Latin	Caquinte	Peru	code

世界の「言語天文台」

ISO 639-2	ISO name	Script	UDHR name	Usage Area	Form
sai	South American Indian (Other)	Latin	Cashibo-Cacataibo	Peru	code
sai	South American Indian (Other)	Latin	Cashinahua	Peru, Brazil	code
cat	Catalan	Latin	Catalan	Andorra, Catalonia	code
ceb	Cebuano	Latin	Cebuano	The Philippines	code
cha	Chamorro	Latin	Chamorro	Micronesia, Guam	code
sai	South American Indian (Other)	Latin	Chayahuita	Peru	code
nya	Chichewa; Nyanja	Latin	Chechewa	Malawi	code
nai	North American Indian	Latin	Chickasaw	USA	code
cai	Central American Indian	Modified Latin	Chinanteco	Mexico	code
cai	Central American Indian	Latin	Chinanteco, Ajitlan	Mexico	code
chi	Chinese	Chinese	Chinese	China, Singapore	code
bnt	Bantu (Other)	Latin	Cokwe	Angola, Congo	code
cos	Corsican	Latin	Corsican	France	code
bos	Bosnian	Latin	Croatian	Croatia	code
cze	Czech	Latin	Czech	Czech Republic	code
nic	Niger-Kordofanian (Other)	Modified Latin	Dagaare	Ghana	code
nic	Niger-Kordofanian (Other)	Modified Latin	Dagbani	Northeast Ghana	code
ada	Adangme	Modified Latin	Dangme	Togo	code
dan	Danish	Latin	Danish	Denmark	code
ira	Iranian (Other)	Arabic	Dari	Afghanistan	image
ssa	Nilo-Saharan	Modified Latin	Dendi	Benin, Niger	code
nic	Niger-Kordofanian (Other)	Modified Latin	Ditammari	Benin	code
dut	Dutch	Latin	Dutch	The Netherlands	code
doz	Dzongkha	Tibetan	Dzongkha, Bhutanese	Nepal, Bhutan	image
bin	Bini	Latin	Edo	Bendel State, Nigeria	code
eng	English	Latin	English	Europe, America	code
epo	Esperanto	Latin	Esperanto	international	code
est	Estonian	Latin	Estonian	Estonia	code
tut	Altaic (other)	Cyrillic	Even	Yakutia	image
ewe	Ewe	Modified Latin	Ewe/Eve	Ghana Togo	code
		Modified Latin	Fante	Ghana	code
fao	Faroese	Latin	Faroese	the Faroese islands	code
per	Persian	Arabic	Farsi/Persian	Iran, Afghanistan	image
fij	Fijian	Latin	Fijian	Fiji	code
tgl	Tagalog	Latin	Filipino (Tagalog)	Philippines	code
fin	Finnish	Latin	Finnish (Suomi)	Finland	code
fon	Fon	Modified Latin	Fon	Benin, Togo	code
fre	French	Latin	French	Europe, Africa	code
fry	Frisian	Latin	Frisian	The Netherlands	code
fur	Friulian	Latin	Friulian (Friulano)	Italy, Austria	code
gaa	Ga	Modified Latin	Ga	Ghana, Togo	code
tut	Altaic (other)	Latin	Gagauz (Gagauzi)	Moldova, Bulgaria	code
glg	Gallegan	Latin	Galician (Galego)	Galicia	code
sai	South American Indian (Other)	Latin	Garifuna	Honduras, Guatemala	code
geo	Georgian	Georgian	Georgian	Georgia	PDF
ger	German	Latin	German (Deutsch)	Austria, Germany	code
nic	Niger-Kordofanian (Other)	Modified Latin	Gonja	Ghana	code
gre	Greek	Greek	Greek (Ellinika)	Cyprus, Greece	code

ISO 639-2	ISO name	Script	UDHR name	Usage Area	Form
kal	Kalaallisut	Latin	Greenlandic(Inuktitut)	Greenland	code
sai	South American Indian (Other)	Latin	Guarani	Paraguay	code
nis	North American Indian	Modified Latin	Guen (Mina)	Benin, Togo	code
guj	Gujarati	Gujarati	Gujarati	Pakistan, Bangladesh	image
cpf	Creoles and pidgins, French-based (Other)	Latin	Haitian Creole (Kreyol)	Haiti	code
			Haitian Creole (popular)		code
sit	Sino-Tibetan (Other)	Latin	Hani	China	code
hau	Hausa	Latin	Hausa/Haoussa	Cameroon, Chad	code
haw	Hawaiian	Latin	Hawaiian	United States	code
heb	Hebrew	Hebrew	Hebrew (Ivrit)	Israel	PDF
hil	Hiligaynon	Latin	Hiligaynon	Panay	code
hin	Hindi	Devanagari	Hindi	India	image
hmn	Hmong	Latin	Hmong(Miao)/Northern East-Guizhou	China	code
hmn	Hmong	Latin	Hmong (Miao) Sichuan-Guizhou-Yunnan	China	code
hmn	Hmong	Latin	Hmong(Miao)/Southern East-Guizhou	China	code
myn	Mayan languages	Latin	Huasteco	Mexico	code
sai	South American Indian (Other)	Latin	Huitoto Murui	Peru, Colombia	code
hun	Hungarian	Latin	Hungarian	Hungary	code
efi	Efik	Latin	Ibibio (Efik)	Nigeria	code
ice	Icelandic	Latin	Icelandic	Iceland	code
		Latin	Ido	Constructed Language	code
ibo	Igbo	Modified Latin	Igbo	South-eastern Nigeria	code
ilo	Iloko	Latin	Iloko/Ilocan	Luzon	code
ind	Indonesian	Latin	Indonesian	Indonesia	code
ina	Interlingua (International Auxiliary Language Association)	Latin	Interlingua	international	code
iku	Inuktitut	Canadian Syllabics	Inuktitut	West of Hudson Bay	PDF
gle	Irish	Latin	Irish Gaelic (Gaeilge)	Eire	code
ita	Italian	Latin	Italian (Italiano)	Italy, Switzerland	code
jpn	Japanese	Japanese	Japanese (Nihongo)	Japan	code
jav	Javanese	Latin	Javanese	Indonesia, Malaysia	code
nic	Niger-Kordofanian (Other)	Latin	Jola-Fogny (Diola)	Senegal	code
nic	Niger-Kordofanian (Other)	Modified Latin	Kabye	Togo	code
kan	Kannada	Kannada	Kannada	Karnataka	image
kau	Kanuri	Latin	Kanuri Yerwa	Borno State, Nigeria	code
bnt	Bantu (Other)	Latin	Kaonde	Congo	code
pam	Pampanga	Latin	Kapampangan	Philippines	code
kar	Karen	Burmese	Karen (Pwo)	Myanmar, Thailand	PDF
kar	Karen	Burmese	Karen (S'gaw)	Myanmar, Thailand	PDF
nic	Niger-Kordofanian (Other)	Modified Latin	Kasem	Ghana	code
kas	Kashmiri	Arabic	Kashmiri	India	image
kaz	Kazakh	Cyrillic	Kazakh	Kazakhstan	code
khm	Khmer	Khmer	Khmer	Cambodia	PDF
myn	Mayan languages	Latin	Kiche (Quiche)	Guatemala	code
kon	Kongo	Latin	Kicongo	Congo, Angola	code
kmb	Kimbundu	Latin	Kimbundu (Mbundu)	Angola	code
nym	Nyamwezi	Latin	Kinyamwezi (Nyamwezi)	Tanzania	code
kin	Kinyarwanda	Latin	Kinyarwanda	Rwanda	code
crp	Creoles and pidgins (Other)	Latin	Kituba	Congo	code

世界の「言語天文台」

ISO 639-2	ISO name	Script	UDHR name	Usage Area	Form
kor	Korean	Korean	Korean (Hangukeo)	Korea	code
kpe	Kpelle	Modified Latin	Kpelewo	Guinea, Liberia	code
cpe	Creoles and pidgins, English based (Other)	Modified Latin	Krio	Sierra Leone, Gambia	code
kur	Kurdish	Latin	Kurdish	Kurdistan	code
kir	Kirghiz	Cyrillic	Kyrgyz	Kyrgystan	PDF
bnt	Bantu (Other)	Latin	Lamso	Cameroon, Nigeria	code
lao	Lao	Lao	Lao	Laos	image
lat	Latin	Latin	Latin (Latina)	the Vatican City	code
lav	Latvian	Latin	Latvian	Latvia	code
nic	Niger-Kordofanian (Other)	Modified Latin	Limba	Sierra Leone	code
lin	Lingala	Latin	Lingala	Congo	code
lit	Lithuanian	Latin	Lithuanian (Lietuviskai)	Lithuania	code
loz	Lozi	Latin	Lozi	Zambia	code
lua	Luba-Lulua	Latin	Luba-Kasai (Tshiluba)	Congo	code
lug	Ganda	Latin	Luganda/Ganda	Uganda	code
lun	Lunda	Latin	Lunda/Chokwe-lunda	Angola, Congo	code
bnt	Bantu (Other)	Latin	Luvala	Angoka, Zambia	code
ltz	Letzeburgesch	Latin	Luxembourgish	Luxembourg	code
mac	Macedonian	Cyrillic	Macedonian	Macedonia	code
mad	Madurese	Latin	Madurese	Madura, Java	code
mag	Magahi		Magahi	Bihar (India)	illegal
bnt	Bantu (Other)	Latin	Makonde	Tanzania, Mozambique	code
mlg	Malagasy	Latin	Malagasy	Madagascar	code
may	Malay	Latin	Malay	Malaysia, Singapore	code
mal	Malayalam	Malayalam	Malayalam	Karala (India)	image
mlt	Maltese	Latin	Maltese	Malta	code
myn	Mayan languages	Latin	Mam	Guatemala	code
nic	Niger-Kordofanian (Other)	Modified Latin	Maninka	Guinea, Mali, Senegal	code
mao	Maori	Latin	Maori	New Zealand	code
rar	Rarotongan	Latin	Maori Cook Islands	Cook Islands	code
arn	Araucanian	Latin	Mapudungun	Chile	code
mar	Marathi	Devanagari	Marathi	Maharashtra	image
mah	Marshallese	Latin	Marshallese	Marshall Islands	code
mwr	Marwari	Devanagari	Marwari	India	PDF
sai	South American Indian (Other)	Latin	Matses	Peru, Brazil	code
myn	Mayan languages	Latin	Mayan (Yucateco)	Mexico, Belize	code
oto	Otomian languages	Latin	Mazahua	Mexico	code
cai	Central American Indian (Other)	Latin	Mazateco	Mexico	code
men	Mende	Modified Latin	Mende	Sierra Leone, Liberia	code
mic	Micmac	Latin	Mikmaq (Micmac)	Maritime Province	code
min	Minangkabau	Latin	Minangkabau	Sumatra	code
cai	Central American Indian (Other)	Latin	Miskito	Nicaragua, Honduras	code
cai	Central American Indian (Other)	Latin	Mixteco	Mexico	code
mon	Mongolian	Mongolian	Mongolian	Mongolia	PDF
mos	Mossi	Modified Latin	Moore	Burkina Faso, Mali	code
bnt	Bantu (Other)	Latin	Nagangela (Nyemba)	Angola	code
nah	Nahuatl	Latin	Nahuatl	Mexico	code
nav	Navajo	Latin	Navaho (Dine)	USA	code

ISO 639-2	ISO name	Script	UDHR name	Usage Area	Form
nde	Ndebele, North	Latin	Ndebele	Zimbabwe	code
nep	Nepali	Devanagari	Nepali	Nepal, Bhutan, Assam	image
cpe	Creoles and pidgins, English based (Other)	Latin	Niger pidgin English	Nigeria	code
sai	South American Indian (Other)	Latin	Nomatsiguenga	Peru	code
nso	Sotho, Northern	Latin	Northern Sotho	South Africa	code
nob	Norwegian Bokmål	Latin	Norwegian Bokmal	Norway	code
nno	Norwegian Nynorsk	Latin	Norwegian Nynorsk	Norway	code
nya	Chichewa; Nyanja	Latin	Nyanja/Chinyanja	Malawi	code
nzi	Nzima	Modified Latin	Nzema	Ghana, Ivory Coast	code
oci	Occitan (post 1500); Provençal	Latin	Occitan Auvergnat	France	code
oci	Occitan (post 1500); Provençal	Latin	Occitan Languedocien	France	code
oji	Ojibwa	cree	Ojibway	Canada, USA	PDF
ori	Oriya	Oriya	Oriya	Orissa (India)	image
orm	Oromo	Latin	Oromiffa	Ethiopia	code
oss	Ossetian; Ossetic	Cyrillic	Osetin	Georgia, Russia	code
ndo	Ndonga	Latin	Oshiwambo (Ndonga)	Namibia, Angola	code
oto	Otomian languages	Latin	Otomi	Mexico	code
sai	South American Indian (Other)	Latin	paez	Colombia	code
pau	Palauan	Latin	Palauan	Palau	code
pus	Pushto	Perso-Arabic	Pashto/Pakhto	Afghanistan	image
ful	Fulah	Latin	Peuhl	Gongola State	code
roa	Romance (Other)	Latin	Picard	Belgium, France	code
cai	Central American Indian (Other)	Latin	Pipil	El Salvador, Honduras	code
pol	Polish	Latin	Polish	Poland	code
pon	Pohnpeian	Latin	Ponapean	Micronesia	code
por	Portuguese	Latin	Portuguese	Portugal, Angola	code
ful	Fulah	Modified Latin	Pulaar	Gongola State	code
pan	Panjabi	Gurmukhi	Punjabi/Panjabi	Punjab (India)	PDF
		Latin	Purhepecha	Mexico	code
myn	Mayan languages	Latin	Q'echi/Kekchi	Guatemala, Belize	code
que	Quechua	Latin	Quechua	Bolivia, Peru, Ecuador	code
que	Quechua	Latin	Quechua de Ambo-Pasco	Peru	code
que	Quechua	Latin	Quechua de Ayacucho	Peru	code
que	Quechua	Latin	Quechua de Cajamarca	Peru	code
que	Quechua	Latin	Quechua de Cotahuasi (Arequipa)	Peru	code
que	Quechua	Latin	Quechua de Huamalistes (Huanuco)	Peru	code
que	Quechua	Latin	Quechua de Margos (Sur de Dios de Mayo, Huanuco)	Peru	code
que	Quechua	Latin	Quechua de Pomabamba (Ancash)	Peru	code
que	Quechua	Latin	Quechua del Callejon de Huaylas	Peru	code
que	Quechua	Latin	Quechua del Cusco	Peru	code
que	Quechua	Latin	Quechua del Norte de Junin	Peru	code
que	Quechua	Latin	Quichua	Argentina, Bolivia	code
roh	Raeto-Romance	Latin	Rhaeto-Romance (Rumantsch)	Grisons, Switzerland	code
rom	Romany	Latin	Romani	Hungary, Bulgaria	code
mol	Moldavian	Latin	Romanian	Romania, Moldavia	code
bnt	Bantu (Other)	Latin	Rukonzo (Konjo)	Uganda	code
run	Rundi	Latin	Rundi/Kirundi	Burundi	code
nyn	Nyankole	Latin	Runyankore-rukiga/Nkore-kiga	Uganda	code

世界の「言語天文台」

ISO 639-2	ISO name	Script	UDHR name	Usage Area	Form
rus	Russian	Cyrillic	Russian (Russky)	Russian	code
smi	Sami languages (Other)	Latin	Sami/Lappish	Norway, Sweden	code
roa	Romance (Other)	Latin	Sammarinese	San Marino, Italy	code
smo	Samoan	Latin	Samoan	Samoa	code
sag	Sango	Latin	Sango	Central Africa	code
san	Sanskrit	Devanagari	Sanskrit	India	illegal
inc	Indic (Other)	Arabic	Saraiki	Pakistan, India	PDF
srđ	Sardinian	Latin	Sardian	Sardinia (Italy)	code
glā	Gaelic	Latin	Scottish Gaelic (Gaidhlig Albanach)	Outer Hebrides	code
srr	Serer	Latin	Seereer	Senegal	code
bos	Bosnian	Cyrillic	Serbian (Srpski)	Serbia	code
		Latin	Serbian (Srpski)		code
shn	Shan	Burmese	Shan	Myanmar	PDF
sai	South American Indian (Other)	Latin	Sharanahua	Peru, Brazil	code
sai	South American Indian (Other)	Latin	Shipibo-Conibo	Peru	code
sna	Shona	Latin	Shona	Zimbabwe, Mozambique	code
snd	Sindhi	Arabic	Sindhi	Sindh Province	image
sin	Sinhalese	Sinhalese	Sinhala	Sri Lanka	image
ssw	Swati	Latin	Siswati	Swaziland	code
slo	Slovak	Latin	Slovak (Slovenscina)	Slovak Republic	code
slv	Slovenian	Latin	Slovenian (Slovenscina)	Slovenia	code
cpe	Creoles and pidgins, English based (Other)	Latin	Solomons Pidgin (Pijin)	Solomon Island	code
som	Somali	Latin	Somali	Somalia	code
snk	Soninke	Latin	Soninke (Soninkanxaane)		code
wen	Sorbian languages	Latin	Sorbian	Lusatia	code
sot	Sotho, Southern	Latin	Southern Sotho/Sotho/Sesotho/Sutu/Sesutu	Lesotho	code
spa	Spanish	Latin	Spanish (Espanol)	Spain, Canary Islands	code
suk	Sukuma	Latin	Sukuma	Tanzania	code
sun	Sundanese	Latin	Sundanese	Indonesia	code
sus	Susu	Latin	Sussu/Soussou/Sosso/Soso/Susu	Guinea, Guinea Bissau	code
swa	Swahili	Latin	Swahili/Kiswahili	Tanzania	code
cre	Cree	Cree	Swampy Cree	Canada	PDF
swe	Swedish	Latin	Swedish	Sweden, Finland	code
tah	Tahitian	Latin	Tahitian	French Polynesia	code
tgk	Tajik	Arabic	Tajiki	Tajikistan	PDF
sit	Sino-Tibetan (Other)	Tibetan	Tamang (Tam)	Nepal	PDF
ber	Berber (Other)	Latin	Tamazight (Beraber)	Algeria, Libya	code
tam	Tamil	Tamil	Tamil	Tamil, Nadu	image
tat	Tatar	Cyrillic	Tatar	Tatarstan	image
tel	Telugu	Telugu	Telugu	Andhra Pradesh	image
myn	Mayan languages	Latin	Tenek (Huasteco)	Mexico	code
crp	Creoles and pidgins (Other)	Latin	Tetum	East Timor, Indonesia	code
tha	Thai	Thai	Thai	Thailand	image
tem	Timne	Modified Latin	Themne (Temne)	Sierra Leone	code
tib	Tibetan	Tibetan	Tibetan	Tibet	PDF
mis	Miscellaneous languages	Latin	Ticuna	Peru, Brazil	image
tir	Tigrinya	Amharic	Tigrinya (Tigrigna)	Eritrea, Ethiopia	illegal
tiv	Tiv	Latin	Tiv	Nigeria	code

ISO 639-2	ISO name	Script	UDHR name	Usage Area	Form
sai	South American Indian (Other)	Latin	Toba	Argentina	code
myn	Mayan languages	Latin	Tojol-a b 'al	Mexico, Cuba	code
tpi	Tok Pisin	Latin	Tok Pisin	Papua New Guinea	code
bnt	Bantu (Other)	Latin	Tonga	Zambia, Zimbabwe	code
ton	Tonga (Tonga Islands)	Latin	Tongan (Tonga)		code
cai	Central American Indian (Other)	Latin	Totonaco	Mexico	code
chk	Chuukese	Latin	Trukese (Chuuk)	Micronesia	code
tur	Turkish	Latin	Turkish	Turkey, Cyprus	code
tuk	Turkmen	Cyrillic	Turkmen	Turkmenistan	PDF
myn	Mayan languages	Latin	Tzeltal	Mexico, Cuba	code
myn	Mayan languages	Latin	TzotZil	Mexico, Cuba	code
uig	Uighur	Arabic	Uighur (Uyghur)	China	image
ukr	Ukrainian	Cyrillic	Ukrainian (Ukrayins ꞑa)	Ukraine	code
umb	Umbundu	Latin	Umbundu	Angola	code
mis	Miscellaneous languages	Latin	Urarina	Peru	code
urd	Urdu	Arabic	Urdu	Pakistan, Kashmir	image
uzb	Uzbek	Cyrillic Latin	Uzbek Uzbek	Uzbekistan	PDF code
vie	Vietnamese	Modified Latin	Vietnamese	Vietnam	PDF
roa	Romance (Other)	Latin	Vlach	Macedonia, Romania	code
		Latin	Walloon/Wallon	Belgium	code
nic	Niger-Kordofanian (Other)	Modified Latin	Wama	Benin	code
war	Waray	Latin	Waray	Philippines	code
sai	South American Indian (Other)	Latin	wayuu	Colombia, Venezuela	code
wel	Welsh	Latin	Welsh (Cymraeg)	Wales	code
tsn	Tswana	Latin	Western Sotho/Tswana/Setswana	Botswana, South Africa	code
wol	Wolof	Latin	Wolof	Senegal, Mauritania	code
xho	Xhosa	Latin	Xhosa	South Africa	code
sai	South American Indian (Other)	Latin	Yagua	Peru, Brazil	code
yao	Yao	Latin	Yao	Malawi, Tanzania	code
yap	Yapese	Latin	Yapese	Micronesia, Palau	code
sit	Sino-Tibetan (Other)	Yi	Yi	China	image
yid	Yiddish	Hebrew	Yiddish	USA, Europe, Israel	image
yor	Yoruba	Modified Latin	Yoruba (Yorùbá)	Nigeria	code
mis	Miscellaneous languages	Cyrillic	Yukagir	Russian Federation	image
zap	Zapotec	Latin	Zapoteco	Mexico	code
zap	Zapotec	Latin	Zapoteco, San Lucas Quiavini	Mexico	code
zha	Zhuang	Latin	Zhuang	China	code
zul	Zulu	Latin	Zulu	South Africa, Lesotho	code