

UNESCO Survey: Linguistic Vitality and Diversity

Objectives of the present survey

Our objective is to collect a large and representative sample of comparable data on the world's languages, particularly endangered and indigenous languages, with two specific purposes in mind. First, these data will be used to prepare the third revised print edition of UNESCO's landmark publication, the *Atlas of the World's Languages in Danger of Disappearing* (see: <http://www.unesco.org/culture/en/endangeredlanguages/atlas>), and to create an interactive on-line *Digital Atlas of the World's Languages in Danger of Disappearing*. Second, the data will serve to develop a methodology for an "Indicator on the Status and Trends of Linguistic Diversity and Numbers of Speakers of Indigenous Languages", as requested by the States Parties to the Convention on Biological Diversity (for information please see: <http://www.unesco.org/culture/ich/index.php?pg=00144>). We hope that this questionnaire, if used on an ongoing basis into the future, will offer a basis for verifiable claims about trends in numbers of speakers, language endangerment and linguistic diversity.

The first section of the questionnaire, titled "Language Vitality and Endangerment", is based on a framework that was developed by an international group of linguists in 2002-2003 to assess the degree of endangerment of specific languages (see: <http://www.unesco.org/culture/ich/index.php?pg=00142>). This framework has previously been applied by individual linguists, and – in a few cases – on a larger scale by national authorities, but it has not yet been used for a global-scale data collection exercise, and this is what we are currently attempting to do.

The second section of the questionnaire, titled "Linguistic Diversity Indicators", has been developed very recently and is still very much a work-in-progress. We would highly appreciate it if you could spare a few minutes and supply information under that section as well. Your feedback on the survey design and questions will also be very welcome.

Complete many questionnaires, share blank forms with colleagues

We are interested in gathering as many independent reports covering as many languages as possible, including multiple reports on the same language, which would enhance the reliability of the data and also would allow us to validate the pertinence of the questions we are asking. We are also interested to begin to create time-series data, so if you have had long-term involvement with a given language we encourage you to complete one form reporting the current status of the language and one form reporting its status when you first encountered or began working with the language. The more good data we have, the more reliable will be our generalizations and the more useful they will be for communities, researchers and policy-makers. So, we also encourage you to provide us with information about other people who can be invited to complete a questionnaire for a given language, and we encourage you to pass the survey on to others. In order to help us assess the validity of the survey instrument, it will be more useful if two observers report independently on the same situation than if two observers collaborate on a single report.

Units of analysis: language, dialect and reference community

The primary entity to be reported in this questionnaire is a language as spoken in a given reference community, with particular attention in the second half of the survey to the dialectal situation of that reference community. Better linguists than we have tried and failed to define the difference between “language” and “dialect”, and we do not pretend to have a solution to that problem. We ask you to use common-sense understandings of the two terms, to identify the specific named language variety your report is specific to, and to provide us with sufficient information. This would allow us to link your report on a given language to other reports on the same or related languages or dialects. Our hope is that with the accumulation of fine-grained reports on specific communities, we will be able over time to assemble reliable and generalizable data. Where a gravely endangered language is spoken by only a handful of speakers all living in the same village, language and dialect and reference community are coterminous. However, most reports will be only a snapshot of a specific situation in a particular locality at a certain moment. Where we have only a single report on one community for a language that is known to be spoken over a vaster territory, that report will serve—until others arrive—as representative. Where we accumulate multiple reports on different reference communities speaking the same language or dialect, we will be able to provide both fine-grained detail and more general aggregated statements. Where we receive multiple reports on the same reference community, we will be able to assess the validity of the

questionnaire and, if the reports cover different time periods, to compile diachronic data.

For instance, if you are reporting on the Evenki language in China, you need not worry about reflecting the situation of this language in Russia or Mongolia (unless you fill out separate forms for those communities!). Moreover, if you consider that a group of Evenki-speakers in China forms a distinct cultural-linguistic community due to great differences in lifestyle and/or language vis-à-vis other Evenki communities, please fill in a separate form for this group. Throughout, the important thing will be that you indicate as clearly as possible what the reference community is that serves as the basis for your report, and provide sufficient identifying information about the language so that we can later link reports on the same or related languages or dialects.

Geographic coordinates

We would highly appreciate it if you could provide geographic coordinates for the reference community. This will in particular facilitate the task of mapping the languages in the new edition of the Atlas, especially in its on-line version. We hope to have both fine-grained detail and aggregated data that can allow users to zoom in from larger to smaller units.

Online tools can help you define easily such coordinates placing dots on maps or entering location names. For more information, please consult the following URL:

<http://www.unesco.org/culture/ich/atlas/georef.php>.

Guidelines for filling in the survey

Please provide a rating score for your language of expertise on each of the factors listed below, where possible. Assign those scores that come closest to describing the situation according to your expertise. If your answer falls between two score options, please pick one and then explain in the “Comments” section. Please note that not all choices are mutually exclusive, and, in some cases, it is possible to check more than one box.

An HTML version of this questionnaire will be accessible shortly from here: <http://www.unesco.org/culture/ich/index.php?pg=00144>. In the meantime, the MS Word version can be downloaded from the same URL.

For each assigned score, please also provide a ‘reliability’ score based on the scale below:

Reliability Index - the assigned score is based on:	
3	Evidence from fieldwork and direct observation
2	Evidence from other reliable sources
1	Very little evidence; a 'best guess'
0	No data available [no score provided]

Name of the language being described in this report:	
Alternative names of the language:	
ISO 639 code(s) of the language (can be obtained from here: http://www.sil.org/iso639-3/codes.asp). Please comment if you have any reservations about the ISO code(s):	
Family and branch of the language:	
Country/ies where the reference community whose language is being described is located:	
Province(s) / region(s) where the community is located:	
Reference community (village/town) where the language described is spoken:	
Geographic coordinates of the community whose language is being described (if possible in a decimal format. Multiple entries are welcome. For help, please refer to the paragraph "Geographic coordinates" in the introduction or to: http://www.unesco.org/culture/ich/atlas/georef.php). Please also comment on the accuracy of the geographic coordinates you are providing:	
Year of the data reported in this report:	
Name, address and E-mail address of expert providing report:	
Would you like to have your name associated with this data when it is displayed?	
Name(s) and E-mail address(es) of other linguist(s) who could provide independent information on this language:	

SECTION I: Language Vitality and Endangerment within the reference community

1. Overall vitality / endangerment score:	5	<input type="checkbox"/>	The language is safe	Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/>	Comments
	4	<input type="checkbox"/>	Unsafe/ vulnerable		
	3	<input type="checkbox"/>	Definitely endangered		
	2	<input type="checkbox"/>	Severely endangered		
	1	<input type="checkbox"/>	Critically endangered		
	0	<input type="checkbox"/>	Extinct		

2. Generational language use	Language is used by:			Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments
	5	<input type="checkbox"/>	all generations, including children		
	4	<input type="checkbox"/>	most children		
	3	<input type="checkbox"/>	only some children		
	2	<input type="checkbox"/>	only grandparents and older generations		
	1	<input type="checkbox"/>	only the great grandparental generation		
0	<input type="checkbox"/>	None			

3. Number of speakers	Please provide the number here for:			Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments
	a) Number of speakers in this reference community				
	b) Absolute number of speakers of the language				

4. Proportion of speakers within the reference community	5	<input type="checkbox"/>	Nearly all speak the language (>90%)	Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments (including the size of the reference community, if known)
	4	<input type="checkbox"/>	The great majority speak the language (70-90%)		
	3	<input type="checkbox"/>	A majority speak the language (50-70%)		
	2	<input type="checkbox"/>	A minority speak the language (30-50%)		
	1	<input type="checkbox"/>	Very few speak the language (<30%)		
	0	<input type="checkbox"/>	None speak the language		

5. Domains of language use	5	<input type="checkbox"/>	Universal use: The language is used in all domains and for all functions	Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments
	4	<input type="checkbox"/>	Multilingual parity: Two or more languages may be used in most social domains and for most functions; the use of the language is usually rare in the official domains (e.g., government, business, administration, education, etc) but may still be in the community's public domains (e.g., religious ceremonies, community gatherings, etc.) and informal domains		
	3	<input type="checkbox"/>	Dwindling domains: The dominant language begins to penetrate informal domains, even home.		
	2	<input type="checkbox"/>	Limited domains: The language is used in limited social domains, for limited functions		
	1	<input type="checkbox"/>	Highly limited domains: The language is used only in very restricted domains, for very limited functions		
	0	<input type="checkbox"/>	No longer spoken: The language is not used in any domain at all		

6. New domains, i. e. new media, including broadcast media and the Internet.	4	<input type="checkbox"/>	The language is frequently used in new domains	Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments
	3	<input type="checkbox"/>	The language is sometimes used in new domains		
	2	<input type="checkbox"/>	The language is rarely used in new domains		
	1	<input type="checkbox"/>	The language is never used in new domains		
	0	<input type="checkbox"/>	Not applicable		

7. Domain of traditional knowledge (TK)	For conveying TK, this language is used:			Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments
	5	<input type="checkbox"/>	Frequently		
	4	<input type="checkbox"/>	Sometimes		
	3	<input type="checkbox"/>	Rarely		
	2	<input type="checkbox"/>	Never		
	1	<input type="checkbox"/>	TK is conveyed using another language		
	0	<input type="checkbox"/>	TK is rarely conveyed		

8. Materials for language education and literacy	<p>5 <input type="checkbox"/> There is an established orthography and literacy tradition with fiction and non-fiction and everyday media. The language is used in administration and education</p> <p>4 <input type="checkbox"/> Written materials exist and at school children are developing literacy in the language. The language is not used in written form in the administration.</p> <p>3 <input type="checkbox"/> Written materials exist and children may be exposed to the written form at school. Literacy is not promoted through print media.</p> <p>2 <input type="checkbox"/> Written materials exist but they may be useful only for some members of the community; for others, they may have a symbolic significance. Literacy education in the language is not a part of the school curriculum.</p> <p>1 <input type="checkbox"/> A practical orthography is known to the community and some material is being written.</p> <p>0 <input type="checkbox"/> No orthography is available to the community.</p>	Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments
9. Governmental and institutional language attitudes and policies, including official status and use	<p>5 <input type="checkbox"/> Equal support for all languages, including the target language.</p> <p>4 <input type="checkbox"/> Differentiated support: Non-dominant languages are protected primarily as the language of the private domain. The use of the target language is prestigious.</p> <p>3 <input type="checkbox"/> Passive assimilation: the dominant language prevails in the public domain, and no explicit policy exists for non-dominant languages;.</p> <p>2 <input type="checkbox"/> Active assimilation: Government encourages shift to the dominant language. There is no protection for non-dominant languages, including the target language.</p> <p>1 <input type="checkbox"/> Forced assimilation: The use of non-dominant languages, including the target language, is discouraged; the target language is neither recognized nor protected by the Government.</p> <p>0 <input type="checkbox"/> Prohibition: Non-dominant languages, including the target language are prohibited.</p>	Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments
10. Reference community members' attitudes towards their own language	<p>5 <input type="checkbox"/> All members value the language of their community and wish to see it promoted.</p> <p>4 <input type="checkbox"/> Most members support the continued use of their language.</p> <p>3 <input type="checkbox"/> Many members support language maintenance; many others are indifferent or may even promote shift to the dominant language.</p> <p>2 <input type="checkbox"/> Some members support language maintenance; some are indifferent or may even support language shift.</p> <p>1 <input type="checkbox"/> Only a few members support language maintenance but most are indifferent or may even support shift to the dominant language.</p> <p>0 <input type="checkbox"/> No-one cares if the language disappears; all prefer to use the dominant language.</p>	Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments

11. Type and quality of documentation	5	<input type="checkbox"/>	Superlative: There are comprehensive grammars and dictionaries, extensive texts and a constant flow of language materials. Abundant annotated high-quality audio and video recordings exist.	Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments (Please note whether the material s are specific to this reference community and whether they are available to them)
	4	<input type="checkbox"/>	Good: There is at least one good grammar, a few dictionaries, texts, literature, and everyday media; adequate annotated high-quality audio and video recordings.		
	3	<input type="checkbox"/>	Fair: There may be an adequate grammar, some dictionaries, and texts, but no everyday media; audio and video recordings may exist in varying quality or degree of annotation.		
	2	<input type="checkbox"/>	Fragmentary: There are some grammatical sketches, wordlists, and texts useful for limited linguistic research but with inadequate coverage. Audio and video recordings may exist in varying quality, with or without any annotation.		
	1	<input type="checkbox"/>	Inadequate: Only a few grammatical sketches, short wordlists, and fragmentary texts exist. Audio and video recordings do not exist, are of unusable quality, or are completely un-annotated.		
	0	<input type="checkbox"/>	Undocumented: No material exists.		

12. Status of language programs	5	<input type="checkbox"/>	Successful: A regular and successful program is running involving >5 per cent of the community.	Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments
	4	<input type="checkbox"/>	Good: A program is running with two of the following characteristics: regular; successful; involving >5 per cent of the community.		
	3	<input type="checkbox"/>	Fair: A program is running with one of the following characteristics: regular; successful; involving >5 per cent of the community.		
	2	<input type="checkbox"/>	Basic: A program is running involving <5 per cent of the community, irregularly and with few or no outcomes.		
	1	<input type="checkbox"/>	Aspiring: No language programs but some community members are talking of starting one.		
	0	<input type="checkbox"/>	None: No language program and no interest in starting one.		

SECTION II: Linguistic Diversity

In this section, please describe the reference community as above in Section I. Assign scores for the following factors (where possible and where relevant) to characterize the linguistic situation and experience in the reference community:

(a) External diversity, i.e. linguistic environment:

13. In everyday life, how many languages would a typical member of this community encounter:		Hear	Speak	Read	Write	Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments
	5 or more languages	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	4 languages	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	3 languages	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	2 languages	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	1 language	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
0 language			<input type="checkbox"/>	<input type="checkbox"/>			

14. In how many languages is a typical member of this community fully fluent?¹	<input type="checkbox"/> 5 or more languages	Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments
	<input type="checkbox"/> 4 languages		
	<input type="checkbox"/> 3 languages		
	<input type="checkbox"/> 2 languages		
	<input type="checkbox"/> 1 language		
Which one(s)?			

15. In how many languages is a typical member of this community at least partially fluent?²	<input type="checkbox"/> 5 or more languages	Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments
	<input type="checkbox"/> 4 languages		
	<input type="checkbox"/> 3 languages		
	<input type="checkbox"/> 2 languages		
	<input type="checkbox"/> 1 language		
Which one(s)?			

¹ 'Fully fluent' is here defined as able to comfortably function in the language in everyday interaction and conversation.

² 'Partially fluent' is here defined as able to engage in basic conversation and understand most of what is said.

16. How many languages are represented in the local schools?	Tolerated	Taught as subject	Used for instruction	Reliability Index:	Comments
	5 or more languages	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
4 languages	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	2 <input type="checkbox"/>	
3 languages	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	1 <input type="checkbox"/>	
2 languages	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0 <input type="checkbox"/>	
1 language	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
Which one(s)?					

17. How many languages are represented in the local media?	Television	Radio	Print	Reliability Index:	Comments
	5 or more languages	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
4 languages	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	2 <input type="checkbox"/>	
3 languages	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	1 <input type="checkbox"/>	
2 languages	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	0 <input type="checkbox"/>	
1 language	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
Which one(s)?					

18. How is TV presence (broadcast time) distributed across the various languages?			Reliability Index:	Comments
	5	<input type="checkbox"/>	Each language receives equal amounts of broadcast time	
4	<input type="checkbox"/>	Several languages receive good amounts of time	2 <input type="checkbox"/>	
3	<input type="checkbox"/>	Two or more languages predominate	1 <input type="checkbox"/>	
2	<input type="checkbox"/>	One language predominates, but other language(s) are well-represented	0 <input type="checkbox"/>	
1	<input type="checkbox"/>	Over 90 percent of the TV broadcast time is dominated by only one language		
0	<input type="checkbox"/>	Only one language is represented on TV		
Which one(s)?				

(b) Internal diversity in the language:

19. Would you say this language is characterized by high internal (dialectal) diversity?			Reliability Index:	Comments
	5	<input type="checkbox"/>	Very high internal diversity	
4	<input type="checkbox"/>	High internal diversity	2 <input type="checkbox"/>	
3	<input type="checkbox"/>	Moderate internal diversity	1 <input type="checkbox"/>	
2	<input type="checkbox"/>	A little internal diversity	0 <input type="checkbox"/>	
1	<input type="checkbox"/>	Virtually no internal diversity		

20. In everyday life, how many dialects would a typical member of this reference community encounter?	5 or more dialects	Hear	Speak	Read	Write	Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments
	4 dialects	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	3 dialects	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	2 dialects	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	1 dialect	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>		
	0 dialect			<input type="checkbox"/>	<input type="checkbox"/>		

21. In how many dialects is a typical member of this community fully or partially fluent?	<input type="checkbox"/> More than 2 dialects	Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments
	<input type="checkbox"/> 2 dialects		
	<input type="checkbox"/> 1 dialect		
Which one(s)?			

22. How equal are the dialects in speaker numbers?	5 <input type="checkbox"/> Each dialect has equal numbers	Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments
	4 <input type="checkbox"/> Several dialects have sizable numbers of speakers		
	3 <input type="checkbox"/> Two dialects predominate		
	2 <input type="checkbox"/> One dialect predominates, but other dialect(s) have good numbers of speakers		
	1 <input type="checkbox"/> Over two thirds of speakers use one dialect		
	0 <input type="checkbox"/> One dialect is used by virtually all speakers		
Which one(s)?			

23. How equal are the dialects in symbolic status and prestige?	5 <input type="checkbox"/> Dialects fully equal in status/prestige	Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments
	4 <input type="checkbox"/> Several dialects have parity in status/prestige		
	3 <input type="checkbox"/> Two dialects have higher status/prestige than other dialects		
	2 <input type="checkbox"/> One dialect has higher status/prestige than all other dialects		
	1 <input type="checkbox"/> One dialect has lower status/ prestige than all other dialects		
Which one(s)?			
What is the status and prestige of this dialect(s)?			

24. Would you say this language is characterized by high stylistic diversity, i.e., a variety of different registers and styles are commonly used in interaction?	5	<input type="checkbox"/>	Very high stylistic diversity, frequently encountered	Reliability Index: 3 <input type="checkbox"/> 2 <input type="checkbox"/> 1 <input type="checkbox"/> 0 <input type="checkbox"/>	Comments
	4	<input type="checkbox"/>	High stylistic diversity, often encountered		
	3	<input type="checkbox"/>	Moderate stylistic diversity, often encountered		
	2	<input type="checkbox"/>	Some stylistic diversity, occasionally encountered		
	1	<input type="checkbox"/>	Little stylistic diversity, encountered infrequently		
0	<input type="checkbox"/>	Virtually no stylistic diversity			

Overall comments and suggestions regarding this questionnaire: