

***Ethnologue* as a global sourcebook on linguistic diversity and language vitality**

Gary F. Simons
SIL International

1. Linguistic diversity by world regions

Data from M. Paul Lewis, ed. 2009. *Ethnologue: Languages of the world, 16th edition*. Dallas, TX: SIL International. Online edition: <http://www.ethnologue.com>

Metrics

<i>GDI</i>	Greenberg's Diversity Index	The probability that two persons selected at random have the same mother tongue
<i>Rich</i>	Richness	The total number of distinct languages
<i>Even</i>	Evenness	The number of equal-sized languages that would give the same probability of having different mother tongues as the GDI; equals $1 / (1 - GDI)$
<i>Dom</i>	Dominance	The proportional size of the largest language

Table 1: Language diversity metrics for world and five major areas

	<i>GDI</i>	<i>Rich</i>	<i>Even</i>	<i>Dom</i>	<i>Largest Language</i>
World	0.967	6,909	30.30	14.2%	Mandarin Chinese
Africa	0.986	2,162	71.43	7.2%	Egyptian Spoken Arabic
Europe	0.934	376	15.15	16.6%	Russian
Asia	0.933	2,355	14.93	23.4%	Mandarin Chinese
Americas	0.723	1,187	3.61	38.0%	Spanish
Pacific	0.464	1,310	1.87	73.1%	English

Table 2: Language diversity by UN regions (highest to lowest)

<i>Region</i>	<i>GDI</i>	<i>Rich</i>	<i>Even</i>	<i>Dom</i>	<i>Largest Language</i>
Melanesia	0.986	1,055	71.43	7.1%	Fiji Hindi
Eastern Africa	0.979	413	47.62	7.3%	Amharic
Middle Africa	0.977	720	43.48	8.0%	Luba-Kasai
Western Africa	0.959	861	24.39	12.8%	Hausa
Southern Asia	0.951	665	20.41	12.6%	Bengali

South-Eastern Asia	0.941	1,269	16.95	15.7%	Javanese
Western Asia	0.891	145	9.17	27.8%	Turkish
Micronesia	0.891	33	9.17	18.1%	Chamorro
Southern Africa	0.890	78	9.09	20.0%	Zulu
Southern Europe	0.862	79	7.25	30.6%	Italian
Northern Africa	0.846	167	6.49	31.1%	Egyptian Spoken Arabic
Central Asia	0.800	52	5.00	36.3%	Northern Uzbek
Eastern Europe	0.768	152	4.31	43.7%	Russian
Western Europe	0.765	98	4.26	38.2%	German
Polynesia	0.741	25	3.86	43.3%	Samoan
Eastern Asia	0.630	320	2.70	59.4%	Mandarin Chinese
South America	0.592	417	2.45	48.7%	Portuguese
Northern Europe	0.589	89	2.43	62.8%	English
Caribbean	0.559	28	2.27	62.1%	Spanish
Northern America	0.360	379	1.56	79.3%	English
Central America	0.186	362	1.23	90.2%	Spanish
Australia and NZ	0.121	191	1.14	93.7%	English

2. Genetic diversity at risk

From: D. H. Whalen and Gary F. Simons, "Endangered language families" Proceedings of the 1st International Conference on Language Documentation and Conservation, University of Hawaii, 12-14 March 2009.
Online: <http://scholarspace.manoa.hawaii.edu/handle/10125/5017>

We identified 342 distinct linguistic stocks, i.e., 240 groupings that are reconstructable to a common protolanguage and 102 language isolates.

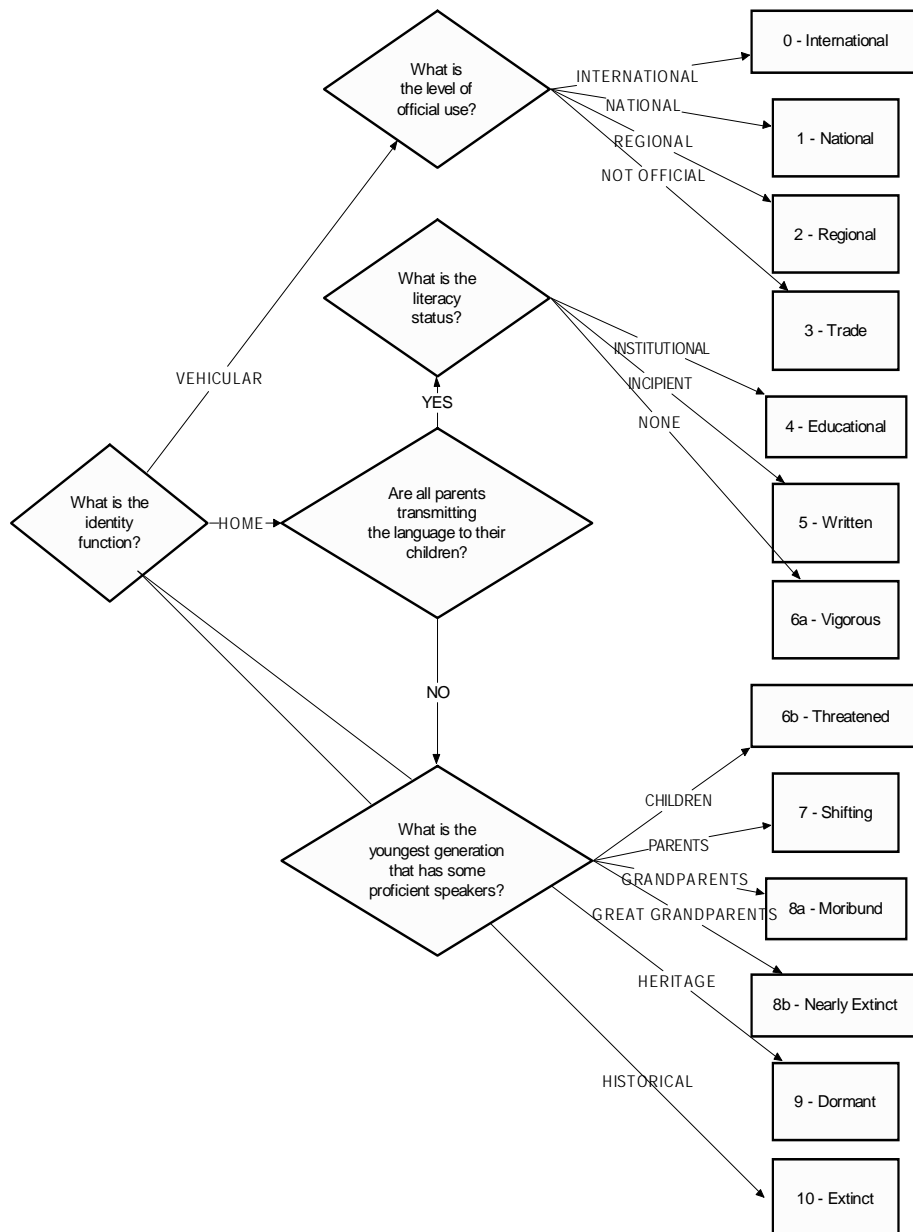
Used population of largest (*i.e.*, safest) language as a proxy indicator: safe if in top 10% of languages (> 330,000), endangered if in bottom 50% (< 7,500), and potentially endangered otherwise.

Table 3: Summary of linguistic stock endangerment by world area

	<i>Endangered</i>	<i>Potentially Endangered</i>	<i>Safe</i>	<i>Total</i>
Africa	14	21	20	55 (16%)
Europe	4	6	3	13 (4%)
Asia	47	21	13	81 (24%)
Americas	72	32	6	110 (32%)
Pacific	57	26	0	83 (24%)
<i>Total</i>	194 (57%)	106 (31%)	42 (12%)	342

3. Expanded Graded Intergenerational Disruption Scale

(Adapted from Joshua Fishman, 1991. *Reversing Language Shift*. Clevedon, UK: Multilingual Matters Ltd.)



Level	Label	Description	UNESCO
0	International	The language is used internationally for a broad range of functions.	Safe
1	National	The language is used in education, work, mass media, government at the nationwide level.	Safe
2	Regional	The language is used for local and regional mass media and governmental services.	Safe
3	Trade	The language is used for local and regional work by both insiders and outsiders.	Safe
4	Educational	Literacy in the language is being transmitted through a system of public education.	Safe
5	Written	The language is used orally by all generations and is effectively used in written form in parts of the community.	Safe
6a	Vigorous	The language is used orally by all generations and is being learned by children as their first language.	Safe
6b	Threatened	The language is used orally by all generations but only some of the child-bearing generation are transmitting it to their children.	Vulnerable
7	Shifting	The child-bearing generation knows the language well enough to use it among themselves but none are transmitting it to their children.	Definitely Endangered
8a	Moribund	The only remaining active speakers of the language are members of the grandparent generation.	Severely Endangered
8b	Nearly Extinct	The only remaining speakers of the language are members of the grandparent generation or older who have little opportunity to use the language.	Critically Endangered
9	Dormant	The language serves as a reminder of heritage identity for an ethnic community. No one has more than symbolic proficiency.	Extinct
10	Extinct	No one retains a sense of ethnic identity associated with the language, even for symbolic purposes.	Extinct

From: M. Paul Lewis and Gary F. Simons., "Assessing Endangerment: Expanding Fishman's GIDS" To appear in *Revue Roumaine de Linguistique*, vol. 55, no. 2, 2010. Preprint : <http://www.sil.org/~simonsg/preprint/EGIDS.pdf>