
The Linguasphere: kaleidoscope of the world's languages

DAVID DALBY

A description of 'the mantle of communication... extended
around the planet'

[This article is reprinted by kind permission of *The Linguist* (Institute of Linguists, London, 39:4, 2000) and of the author.]

THE NEW perceptions afforded by the era of telecommunications may be compared to those of the Renaissance or the Enlightenment. That the world is a globe, or that blood circulates around the body, became obvious only a few centuries ago... as it now becomes apparent that all spoken and written languages form an organic continuum, a global medium for human communication and the circulation of ideas.

The continuity of communication between communities in contact has always been maintained by those who speak more than one language, by those whose brains have bridged – or attempted to bridge – the differences of viewpoint and expression embodied in the use of differing languages by different communities. It is only in modern times, however, that networks of bilingual links among thousands of speech communities around the world can be seen in global perspective. The ancient continuum of multilingual speech is becoming a system of immediate worldwide communication.

What is the 'linguasphere'?

Just as 'biosphere' denotes the terrestrial mantle of living organisms, so the term 'linguasphere' denotes the mantle of communication gradually extended around the planet by humankind. The onset of the linguasphere may be dated to the times – 50,000 years or more ago – when articulate speech unleashed the

power of the human brain, and when the mutability of speech provided communities with the means to express separate identities, traditions and ideologies.

With the development of writing from around 5,000 years ago, some languages were set into more or less standardised forms, although without arresting the relentless shift in spoken conventions, from generation to generation and from village to village. With the development of printing from over 500 years ago, books have served to protect the imagined 'purity' of certain languages, although the printed word has generally accelerated the flow of words among individual languages.

The distribution of writing systems across the linguasphere has not always corresponded to the landscape of spoken languages, so that some pairs of interintelligible languages, including Serbian and Croatian, or Hindi and Urdu, are written in different alphabets from each other. In contrast, the major Chinese languages (not readily interintelligible when spoken) are the 'same' language in writing, through the use of identical meaning-based characters.

As the multilingual vehicle of human expression, the linguasphere comprises the diversity of all linguistic conventions evolved and modified within and among individual communities since the beginnings of speech. These include:

DAVID DALBY is director of the *Observatoire Linguistique* and emeritus reader at the *London School of Oriental and African Studies*.
Email: speechways@aol.com

- the total lexical repertoire of humankind, made up of the overlapping vocabularies of all languages spoken, written or read at any one time,
- the global distribution of the overlapping and slowly shifting phonological and grammatical patterns which structure those vocabularies,
- and the geographic extension of scripts and other systems employed to transcribe spoken languages.

By providing the key to communication and concerted action by the dominant species on earth, the linguasphere is the single most influential layer of the biosphere. Dependent on the 'recycled' voices of generations of individuals, conveyed by the expansion and multiplication of communities of voices, and reinforced by the written and other derivatives of speech, the linguasphere has come to encompass the entire globe and its surrounding airwaves. Finally, in this generation, it has achieved a global vitality previously undreamed of, an event which deserves urgent reflection and concerted debate around the world.

Languages no longer need to be viewed as countable entities, as though they were apples or nation states. The interwoven vocabularies of French and English, for example, can be studied as a shared resource, alongside their often subtle divergences of meaning. An obsession with numbers and records often leads journalists and others to ask: "How many languages are there?" But what is a 'language'? The distance between varieties of spoken Welsh, in Pembrokeshire and Anglesey for example, are probably as great as between closely related written languages, such as Danish and Swedish.

The number of modern languages spoken in the world may be placed between around 5,000 (using a broad definition of wider or 'outer' languages, including Welsh and Swedish+ Danish) and around 14,000 (using a narrower definition of 'inner' languages, including northern Welsh, southern Welsh, literary Welsh, Swedish, 'book' Norwegian and Danish).

The state of the linguasphere

Rather than trying to quantify languages in the world, it is more profitable to seek a qualitative view of their collective development.

A central issue is the growth of English as a

world language, in its many spoken varieties. One or more languages with an intercontinental spread are clearly required by the current globalisation of economies and the media, but does this necessarily mean the demise of other languages? Some commentators, mainly English-speaking, have warned that up to 90% of all other languages may be extinct before the year 2100. Yet such estimates are almost certainly exaggerations. Languages do not exist only as competitive instruments of communication but also as means of personal expression and communal identity. A substantial proportion of the world's population has always needed to master other languages for communication with the wider world, but without necessarily abandoning their own linguistic heritage. A notable example of the survival of a threatened language is provided by Welsh, which has resisted the invasion of its present territory by speakers of two of the most powerful languages in the history of the world, Latin and English. Or one may cite the creation of written 'country' Norwegian in the 19th century, representing the rural speech of western and northern Norway against the assimilating influence of 'book' Norwegian.

This is not to say that some languages are not now in imminent danger of extinction, especially those spoken by small communities of hunters and gatherers – in such areas as the Amazon basin, the Arctic, southwestern Africa and Australasia. Between the 17th and 20th centuries, the majority of the hunting and gathering communities of two entire continents – North America and Australia – were destroyed by sea-borne invaders from Europe, principally English-speaking. This does not imply, however, that English has a power in itself to destroy the more robust speech communities of other continents. It is much more likely that the world will see a steady growth in bilingualism and multilingualism during the coming decades, as people everywhere seek to participate in an expanding world economy by using English or other major languages. In view of the educational and intellectual advantages acquired in mastering more than one means of communication, monolingual speakers of English may find themselves at a disadvantage in a largely bilingual or multilingual world.

A frequently asked question is: "Which are the world's most spoken languages?" The answer is best formulated in relation to the overall population of the world, and it has

proved useful to define major 'arterial languages' as those spoken each by at least 1% of humankind (in other words, by at least 60 million voices at the turn of the millennium). The recently published *Linguasphere Register* (see below) identifies 29 such arterial languages, of which the four with the greatest number of first and second language speakers are Putonghua (or Mandarin Chinese), English, Hindi+Urdu and Spanish.

The relative importance of each of these four spoken languages currently alternates during the course of each day. When the sun is over the western Pacific, the national language of China is the most in use, but when the sun is over the Atlantic and China sleeps, English takes the lead. The world's second most spoken language also alternates daily, between Hindi+Urdu and Spanish respectively.

No less than 15 of the 28 arterial languages spoken in the world today are derived from or closely related to only five classical languages, the spoken strength of which was amplified in previous millennia by their cultural importance as the written languages of religious and/or political empires. From east to west, these major languages were Classical Chinese, Sanskrit (with Pali), Arabic, Church Slavonic and Latin. All the remaining arterial languages have also been heavily influenced by one or more of these same classical languages. This important fact demonstrates the great power which writing and the propagation of written religions have had over the spoken word – even though the written word derives from the spoken, and though the majority of humankind has been illiterate until recent times.

Since telecommunications have had an even greater and much more immediate impact than the inventions of writing and printing, one may anticipate an even more powerful effect on the future evolution of the world's languages. It should not be expected, however, that this will be to the unique advantage of English and other arterial languages. Even in the 20th century, endangered speech communities in South America and elsewhere had discovered the importance of the tape recorder as a means of preserving the voices and traditions of their languages. Today, speech communities can retain spoken contact even when their members are scattered around the globe. Video recordings can preserve the languages and oral literature of the world's most endangered com-

munities, laying the foundations of an audiovisual record of the linguasphere, of humankind's collective linguistic heritage.

Observation of the linguasphere

The linguasphere is humankind's greatest and most complex creation. The voice of every communicating person participates in its operation and further development, although no one person has ever had knowledge of more than a minute part of its cumulative resources. Now, for the first time, the linguasphere may be observed and studied as a global reality. Individuals around the world can reflect and confer together on its development and management in the era of telecommunications. A global system of documentation, research and related education will be a key to maximising this opportunity.

During the 1980s, a transnational research network – the Observatoire Linguistique – was established in France to promote the study and development of multilingualism. Free from all political, religious and commercial ties, the observatoire has created travelling exhibitions on African and European languages, and has laid the electronic foundations of a virtual institute, Linguasphere.Org. From a research base in west Wales, it has recently completed and published the *Linguasphere Register of the World's Languages and Speech Communities*. The two volumes of this basic work of reference include an introduction to the linguasphere, an annotated inventory of over 21,000 inner languages and dialects, and a classified index of over 70,000 linguistic and ethnic names. Available also in an updatable online version, the Linguasphere Register provides the first comprehensive classification of the world's languages and dialects. Digital codes facilitate the identification and location of all languages and their communities within a referential system of linguistic and geographic zones.

By making the register available worldwide in electronic form, the observatoire has begun to gather additional data from its users online, leading to the preparation of future updated versions. Every interested researcher and academic institution is invited to participate directly in this global programme, which offers major opportunities for the worldwide gathering of data. Financed by voluntary contributions and the sale of its publications, the affairs of the observatoire are the concern of all who partici-

Linguasphere 2000: THE WORLD'S ARTERIAL LANGUAGES each spoken by over 1% of Humankind

Adapted from *The Linguasphere Register of the World's Languages and Speech Communities* (Linguasphere Observatory, Hebron SA34 0XT, Wales: 2000) Language names linked by '+' form a continuum often served by different written standards, e.g. Hindi + Urdu, or Thai + Lao. In contrast, one name may cover wide spoken variations (with one written standard, e.g. Arabiyya or Deutsch). Totals listed are provisional estimates. Populations bilingual in two arterial languages are counted under both. Supplementary data or comments should be sent to <research@linguasphere.org>. This chart may be freely reproduced for educational purposes.

WIDER AFFINITIES or 'language families' 2-digit codes denote Linguasphere zones)		ARTERIAL LANGUAGES estimated at 60 million voices each (= over 1% of humankind, including first and second language speakers)	(1st+2nd language speakers=) in millions, with approx. % of humankind	TOTALS in millions, with approx. % of humankind	
Sinitic/Wider Chinese	[79]	PUTONGHUA (Chinese-Mandarin)	(800+ 200 =)	1000m	16%
		WU (Chinese-Wu)	—	85m	1%
		YUE(Chinese-Cantonese)	—	70m	1%
Indo-European					
Roman/Latin-related	[51]	ESPAÑOL (Spanish)	(400+ 50 =)	450m	7%
		PORTUGUÊS (Portuguese)	(180+ 20 =)	200m	3%
		FRANÇAIS (French)	(90+ 35 =)	125m	2%
		ITALIANO (Italian)	(60+ 10=)	70m	1%
Germanic	[52]	ENGLISH ('Romanised Germanic') DEUTSCH (German)	(400+ 600 =) (100+ 25 =)	1000m 125m	16% 2%
Slavonic	[53]	RUSSKIY (Russian) UKRAINSKA (Ukrainian) + BELARUSSKAYA(Belarusssian)	(170+ 150 =) —	320m 60m	5% 1%
Iranic/Iranian Indic/Sanskrit-related	[58]	FARSI (Persian <i>including</i> Dari) + TAJIKI (Tajik)	(35+ 25 =)	60m	1%
	[59]	HINDI+ URDU	(550+ 350 =)	900m	15%
		PANJABI (Punjabi)	—	85m	1%
		BHOJPURI + MAITHILI	—	60m	1%
		BANGLA (Bengali) + SYLHETI MARATHI	(190+ 60 =) —	250m 80m	4% 1%
Austro-Asiatic: Semitic	[12]	'ARABIYYA (Maghribi/Western Arabic + Mashriqi/Eastern Arabic)	(200+ 50=)	250m	4%
Austronesian Hesperonesic	[31]	MALAYU (Malay) + BAHASA-INDONESIA (Indonesian)	(50+ 159 =)	200m	3%
		JAWA (Javanese)	—	80m	1%
		TAGALOG (<i>including</i> Pilipino/Filipino)	(25+ 35 =)	60m	1%
Turkic	[44]	TÜRKÇE (Turkish) + AZERI+ TURKMEN ('Wider Turkish')	—	80m	1%
(isolated East Asian languages)	[45]	NIHONGO (Japanese) HANKUKMAL (Korean)	(120+ 10 =) —	130m 75m	2% 1%
Mon- Khmer	[46]	VIỆT (Vietnamese)	—	75m	1%
Daic/Tai	[47]	THAI + ISAN + LAO + HUANG + BUYI (Thai+ Lao)	—	90m	1%
Dravidic/Dravidian	[49]	TAMIL	—	70m	1%
		TELUGU	—	70m	1%
Transafrican/Volta-Congo: Bantu	[99]	KISWAHILI (Swahili)	(10+ 50 =)	60m	1%

pate in its collective research, regardless of nationality or academic position.

A Decade of the Linguasphere (Décennie de la Linguasphère) was announced at the recent Expolangues 2000 exhibition in Paris. To be promoted from October 2000, this will provide ten years of reflection on the collective development of the world's languages in an era of expanding telecommunications. It will include the organisation of an ongoing transnational debate on global language strategies, and of links among bilingual schools in different continents. Above all, the decade will present the

opportunity to promote new attitudes towards a multilingual world. Children and adults will be encouraged to view the world's languages and literatures, not as the separate preserves of rival nation-states and communities, but as a dynamic common heritage. ■

Extracts from the *Linguasphere Register* (Linguasphere Press, Hebron: 2000; ISBN 0 95329191 X) may be freely downloaded from <www.linguasphere.org>, and the full work (£95, plus optional £65 for Linguasphere Online) may be ordered through <www.linguasphere.net>.

SNIPPETS 1

The coming silence?

(From Jeffrey Bartholet, 'The Sounds of Silence', *Newsweek* 19 Jun 00; see also pp. 60–62)

Languages have come and gone since the early days of human history, but the wave of destruction today is unprecedented. It's happening wherever technologically advanced societies overwhelm less powerful groups. Australian colonizers helped wipe out more than 150 native languages over the past 222 years, and more than 100 others are on the brink. In South America, Spanish and Portuguese have overwhelmed scores of native Indian languages, and pioneers pursuing "manifest destiny" helped to destroy most of some 300 languages native to North America. Of 100 languages that once were spoken in what is now California, only half remain, and most of those are spoken only by a few tribal elders.

Globalization is probably helping to fuel the destruction. English, in particular, is quickly becoming the indispensable language of successful people from different countries and cultures. That's partly because a disproportionate number of the world's rich speak English, and also because English is the language of the technological revolution.... The obliteration of small languages might seem inevitable and irreversible. But languages, unlike people, can be resurrected. The last fluent speaker of Miami Indian died in the 1960s, but Daryl Baldwin, 37, has nursed Miami back to life.... [H]e taught himself Miami vocabulary, grammar, and pronunciation, and then he brought the language into his home.... Other Miami Indians, among a tribe numbering 7,000..., are joining the effort.