
Chinese, English, Spanish – and the rest

TOM McARTHUR

How do the world's very large languages operate within its 'communicative ecology'?

[An adapted and extended version of 'The Big Three', a paper presented to the Kongres Bahasa Utama Dunia ('Congress of the World's Major Languages'), organized in October 2004 by the Dewan Bahasa dan Pustaka ('Institute of Language and Literature'), Kuala Lumpur, Malaysia. The Dewan, established in 1956, is committed to both linguistic and literary planning with regard to Bahasa Malaysia (Malay as used in Malaysia) and the recognition of Malay as a major Asian and world language.]

THE SECOND edition of the *Oxford Dictionary of English* (2003) defines *ecology* as 'the branch of biology that deals with the relations of organisms to one another and to their physical surroundings'. I would, however, like to extend this definition here to include an 'ecology of communication', covering the nature and evolution of language, the media, and such technologies as radio, telephony, television, and the net/web. This discussion will not, however, deal with an ecology of communication at large, but five particular, related matters: English and the rest of the world's languages; the idea of a 'world language'; the interaction of the world's languages and communicative technologies; the millions of people who use English and at least one other language; and the size, security, and health of our global linguistic inheritance.

The English language complex

English is a paradox. In traditional terms it is a single language, but in recent centuries it has become so large and varied that it has taken on the features of a family of languages: like the

Semitic languages (origin, West Asia) or the *Romance languages* (which might more properly be called the *Latin languages*: cf. Wright 2004). Anyone concerned after the 1980s with *world* or *international* or *global* English also knows (and may use) the form *Englishes*, promoted with marked success by the Indian American linguist Braj B. Kachru. However, although phrases like *the Englishes* and *world Englishes* imply a family, Kachru has not (to my best recollection) explicitly proposed that any of the Englishes are, or might become, distinct languages. One wonders, therefore, whether his 'Englishes' constitute a single complex entity or a family of entities.

Since bringing out *The English Languages* in

TOM McARTHUR has recently been a Distinguished Visiting Professor at the Chinese University of Hong Kong and at Xiamen University, Fujian, China. In addition to editing *ET* he is the creator of 'The Oxford Companion to the English Language' (1992), the 'Oxford Guide to World English' (2002) and 'The Longman Lexicon of Contemporary English' (1981), and the author of 'The English Languages' (Cambridge, 1998) and 'Worlds of Reference' (Cambridge, 1986). In the 1980s, he broadcast regularly for BBC English (World Service), notably in the series 'The Story of English', with David Crystal. He has an M.A. from Glasgow University, and a Diploma, M.Litt., and Ph.D. from Edinburgh University, has been an education officer in the British Army, Head of English at Cathedral School in Bombay, India, an Associate Professor of English at the Université du Québec, Canada, and Deputy Director of the Dictionary Research Centre, the University of Exeter, England.

1998 I have continued to hold the plural view expressed there. However, I have also felt the need for a further phrase: *the English language complex*, which retains the tradition of *English* as a single entity, while allowing for both *the Englishes* and *the English languages* as terms for handling a multiplicity that includes *African English*, *African-American English*, *American English*, *British English*, *China English*, *Indian English*, and *Nigerian English*. Yet wherever one looks there are semantic complications. For example:

- The label *British English* can be ambiguous if not overtly clarified: It may mean upper- and middle-class British usage with an RP accent (as conventionally on offer in ELT textbooks and dictionaries) or all English as used in England alone or all English as used in England, Scotland, Wales, and Northern Ireland.
- The label *Indian English* has a range of subdivisions that includes *Anglo-Indian English*, *Bengali English*, and is at the same time (uncontentiously) part of *Asian English*, yet we do not usually say that *British English* is part of *European English*, although oddly enough (since there are now many such entities as *Dutch English* and *Swedish English*) this is the case. To complicate matters further, people now talk (often disparagingly) of a European Union kind of English, used especially in Brussels, that is often called *Euro-English*, but explicitly *not* what either the British or the Irish use.

Containing by definition these and other (often confusing) *varieties*, the English language complex is, as it were, a system of systems whose membership is unclear, because commentators can always find another English to add to the list: for example, *Afghan English* or *United Nations English*. One cannot predict or apparently exhaust the number of such terms, whether they are geographic, educational, professional, social, or a mix of these, as with: *British English* (as above), *Business English*, *British Business English*, *American Business English*, *Anglo-American Business English* (note: much more likely than *British-American Business English*), *Californian English*, *Californian Media English*, *China English* and *Chinese English* (delicately contrasted in recent issues of *ET*), *Jamaican English Creole*, and indeed that really difficult pair, *X Standard English* and *Standard X English* (where we can replace the X

with whatever nation, region, ethnicity, business, or other activity we wish) *ad infinitum et nauseam*.

We can say, however, that for some time there have been two universally agreed nation-based standards of English, one centred on the UK, the other centered on the US. It has become increasingly clear, however, over the last half-century, that these are no longer alone. Most Australians, Canadians, and New Zealanders now assume that they too have standard varieties, and there are dictionaries to demonstrate the fact. And indeed, the Irish Republic, South Africa, India, and Singapore can arguably be added to this list (with or without national dictionaries).

In a major everyday sense, however, English remains a language among languages. Yet, even so, many would agree that, since at least the mid-twentieth century, it has not been a language in the way that Danish, Hungarian, Korean, Maori, or Nahuatl are languages, for at least six reasons: its scale, its distribution, the number and range of its varieties, its technical and professional applications, and the number of its users and would-be users, whether as a mother tongue or an other tongue. One reason for calling it a *complex* is that many of its varieties are *not*, and never have been, dialects. No one talks of an *Indian dialect of English*, or sub-dialects within Indian English. They do, however, talk about *Punjabi English*, *Bengali English*, and such like, as (sub)varieties of *Indian English*, each of them influenced by at least one major regional language and/or the vagaries of regional life, and *not* by isolation from the mother source, as was the case for example with the usage of Newfoundland and the Falkland Islands.

In traditional terms, dialects are seldom sharply separated, but are areas within continua: that is, Dialects A and B do not usually 'meet' at a boundary line: rather, an area of A shades into an area of B, until A-ness becomes B-ness. In certain politico-cultural instances, however, one area in such a continuum acquires prestige and a script in which records can be kept, orders issued, and books written, becoming because of this the primary and perhaps sole official tool for national record-keeping, news, policy, and officially-received culture, as with *le bon français* or the *King's English*. When a speech form becomes like that, linguists call it an *acrolect* (Greek: 'high dialect') and the rest become *basilects* ('low

dialects'), maybe with an intermediate layer or two, the *mesolects* ('middle dialects'). Such names make commentary more clinical, but *acro* and *basi* are just Greek elements in English that mean 'high' and 'low' – verging (dare I say it?) on a cosmetic exercise. The distance between Standard English and the vernacular Englishes, between Mandarin Chinese [Putonghua] and the Han dialects, or Classical Arabic and the vernacular Arabics remains great however the scale is calibrated.

Standard, dialect, variety, accent, pidgin, creole

By and large, the standard and near-standard forms of British and American English are mutually intelligible (especially in print): in effect 'the same language' while differing in ways which do not invoke questions about which is a dialect of the other. Many British people have of course insisted that US usage is derivative and therefore secondary, but an overwhelming reversal of global roles in the course of the twentieth century has gravely damaged that particular argument. Linguists, however, agree that neither is a lect of the other and each has lects of its own – and that, in any case, the US is vastly larger and more populous, and carries a bigger stick.

After 1945, there was no reason for Americans to feel secondary to anyone in any aspect of their lives, even if some might defer to British niceties. The rest of the twentieth century simply affirmed a state of affairs that seems likely to last well into the twenty-first. But by and large none of this affected other kinds of English elsewhere, including entities far more exotic than traditional dialects and mildly different standard varieties. English-based pidgins and creoles have sometimes been called 'dialects', but usually they have been described not just as 'barbarous' or 'uncouth' (as dialects have often been described) but as 'broken' and 'debased': not real English but bastard offshoots to be regretted, and kept at arm's length.

In such a world climate the term *variety* has been immensely reassuring (but at times a copout) for linguists and language teachers. Although I find the term invaluable, I worry about its catch-all blandness: any distinctive spoken, written, printed, electronic or other aspect of a language, and especially English, can be a 'variety'. And the term has not helped

much in avoiding the original problem: In everyday life, *variety* is not much used for talking about language, and the term *dialect* is far more likely to be applied to mesolects and basilects than to acrolects (which are or become the 'standard' language that is blessed with orthography and print). Meanwhile, the term *pidgin* remains low, often negative, and at best neutral (despite having been adopted by linguists as a technical term), and *creole* is hardly known outside the Caribbean, Louisiana, and linguistics libraries.

There is more, however. In addition to dialects, pidgins, and creoles, English in all its varieties has long been mixing with other languages, producing a range of, as it were, *Anglo-hybrids*. Such entities may seem chaotic to an outsider meeting them for the first time, but blends of this kind are inevitable in locales where two or more avowed languages are in wide daily contact. Such outcomes as *franglais* and *Spanglish* are not, however, chaotic to their users but are pragmatic blends in which the most immediately recalled and relevant material takes pride of place. Many social commentators deplore such hybridization, but deploring it has never put an end to it, any more than it has put an end to dialect.

When people of different backgrounds within 'the same' language come together, the amount of adjustment needed for adequate communication may range from minimal to massive, but by and large exposure to such media as TV and films/movies and news services such as BBC World and CNN appears to have made encounters with variation easier than in the past. Even so, however, remoteness from the world's main travel routes and lack of exposure to the media may make some varieties of a language like English as exotic as the flora and fauna of their locale, as, say, with *Caribbean Creole* and *Tok Pisin* in Papua New Guinea.

Seven levels

English is a vast language, whereas Scottish Gaelic now has about 70,000 speakers left. A thousand years ago Gaelic was the primary language of Scotland, accumulating a strong tradition of orature and literature that only a few can now access. I was born into a family whose last Gaelic speaker died in the 1930s. My children know only a few words (such as *slainte* 'health', and also a toast which few know, and *glen*, which many know because it became an

English word). Both my father and I wanted to acquire it, but life got in the way. However, all three of my children know English, and also French, from living in Quebec; my older daughter is fluent in Japanese, and my son is competent in Italian. Alas, however, none speaks Persian, their mother's language, because she chose not to use it with them. Crucially, however, neither set of grandparents when young could have predicted the languages their grandchildren would know, and this seems increasingly to be the way of the world.

The following is a seven-level model that seeks to represent the world's languages in terms of size and 'clout' (if any). The first five levels 'contain' the largest and safest languages, and here I spend more time on them (alas), while the lowest two levels contain the many languages worldwide suffering diminution and facing extinction (cf. Crystal 2000).

Level 1: English

A vast, unevenly but widely distributed language complex with two globalized national varieties (American and British), each with a standard form, three further major national varieties (Australian, Canadian, New Zealand), and many varieties worldwide. Around a billion people use it, as a first, second, or other language, and its standard varieties are taught to further millions. It can be heard or read almost everywhere, and is the world's primary vehicle for the media, commerce, technology, science, medicine, education, popular and youth culture, travel, trade in armaments, and United Nations peace-keeping. In terms of its widest reach it is often called *World*, *International*, or *Global English*, and there is a multitude of regional and other *Englishes*, known more or less formally as, say, *South African English* and *Japanese English*, or informally and often facetiously as, say, *Japlish* and *Taglish* (English mixed with Tagalog, in the Philippines).

Level 2: Chinese, Spanish, Hindi-Urdu

Three language complexes covering hundreds of millions of people. *Chinese*, the largest, has over a billion users, the vast majority in the People's Republic of China, Hong Kong, and Taiwan. *Spanish* is widely disseminated in the Americas, with a limited role in Europe (its continent of origin, as with English). Despite

the prestige of Castilian usage in Spain, the language has no single strong centre, like Chinese, or centres, like English. Its geopolitical focus, like English, is the western hemisphere, from Argentina to the US. It is also spoken in Morocco and Equatorial Guinea in Africa and the Philippines in Asia. The vast majority of users of *Hindi-Urdu* are in northern India and Pakistan, with diaspora populations elsewhere. Hindi-Urdu is more or less one entity in speech but not in sociocultural terms. Its Hindi component is the official language of India, uses the Devanagari script, and is largely a medium for northern (not southern) Hindus. Its Urdu component uses the Perso-Arabic script and is primarily a medium for Muslims in India and Pakistan. All three complexes are world-regional, with little (though increasing) wider diffusion.

Level 3: Arabic, French, German, Japanese, Malay

The *Arabic* complex is used across North Africa and in much of the Middle East, throughout which it is the primary medium of ethnic Arabs. In addition, from Europe and North Africa to the Philippines it is the historic vehicle of Islam. It has both a Classical form and a range of colloquial, and not necessarily mutually intelligible, forms in Bahrain, Dubai, Egypt, the Emirates, Iraq, Jordan, Kuwait, Lebanon, Morocco, Qatar, Saudi Arabia, Sudan, Syria, and Yemen. *French* has traditionally been in cultural competition with English, German, and Italian, and until the mid-twentieth century was the élite language of diplomacy. After the Second World War, it lost ground to *les anglo-saxons* (the Americans and British viewed as a unity in which the British formerly dominated and the Americans now dominate). The community of *le français mondial* ('World French') is known as *La Francophonie* ('the French-speaking world'). *German*, formerly strong in Europe and some African colonies, and in academic-cum-scientific writing, lost influence as a consequence of the defeat of Germany and Austria in two world wars, but retains strength in business and technology. *Japanese* has never had a world role, despite Japan's global commercial success. *Malay*, however, has begun to develop a higher profile in Asia because of its role in Malaysia, Indonesia, Brunei, Singapore, and southern Thailand.

Level 4: Significant nationally and/or regionally

This level includes such flourishing strong regional languages as Ashanti, Hausa, Ibo, and Yoruba in West Africa; Amharic, Swahili, Somali, and Kikuyu in East Africa; Dutch, Hungarian, Polish, Rumanian, and Swedish in Europe; Portuguese in Brazil, Portugal, Angola, Mozambique, East Timor, and Macau; Turkish and Persian in Western and Central Asia; Bengali, Gujarati, Kannada, Konkani, Marathi, Oriya, Punjabi, Tamil, Telugu, and Malayalam in India: all used in communities with large populations, complex social histories, and thriving cultures. Most of the languages at this level are robustly healthy, and many have diaspora communities.

Level 5: Locally and socially strong

Languages within one or more nations or territories, such as Berber in Morocco, Catalan in Spain and France, Danish in Denmark, Finnish in Finland, Maya in Mexico and Central America, Guarani in Paraguay, Tagalog and Ilocano in the Philippines, Nahuatl in Mexico, Quechua in Peru, and, as an exceptional case, the vibrant small language Icelandic. Some languages at this level are secure in national terms (as with Danish and Finnish, although English is in use for professional, higher-educational, and international purposes), while others occupy a mid-position in their local hierarchy (as with Catalan and Ilocano), and others have large numbers of users but little or no international prominence (as with Berber, Maya, and Quechua).

Level 6: Small and (perhaps) managing

Hundreds of languages used mainly in politically or geographically non-metropolitan areas, or among migrant workers, whose speakers number in the low millions, hundreds of thousands, or less. They may or may not be well-situated nationally, regionally, continentally, socioculturally, or in educational terms, and in many cases have been depleted by social change, including emigration of speakers and immigration of non-speakers. They may, however, be sustaining themselves, sometimes in secure situations, such as Faroese in Denmark's Faroe Islands, sometimes in situations less negative than formerly, as with Maori in New Zealand,

sometimes under politically neutral conditions, such as Welsh in Wales (within the UK), Gallego in Spain, and more secure Amerindian languages such as Navajo.

Level 7: Extremely small and endangered

Some thousands of languages in the Americas, Asia, Africa, Australasia, Europe, and the Pacific whose speakers number in the thousands, hundreds, or tens and less, often in communities that are shrinking for various reasons (including migration and intermarriage) or disrupted and sometimes demoralized, within and ranging across the boundaries of nation-states, and therefore semi-assimilated into more socially and economically powerful societies. Some manage to sustain themselves, as with Lapp in northern Scandinavia, while others have for many years been in dire straits, particularly the remaining Aboriginal languages of Australia, the 'heritage languages' of Canada's 'Native Peoples', such as Ojibwa and Inuit; and comparable 'American Indian' languages in the US, such as Cherokee and Mohawk. Some may be dichotomous, where for example Irish Gaelic (or Irish) is somewhat sustained by government effort and a national education policy in the Irish Republic, but Scottish Gaelic has no official role or protection, and dwindles despite efforts to promote it (including through internet lessons). Some groups are stable and fairly secure but many are socially disrupted, the numbers of active speakers dwindling, and many probably will not survive the next half-century.

Discussion of the languages in Categories 1 and 2 of this list is difficult without bringing in the languages of 3 to 7, because in many instances multilingual individuals and communities use both large languages alongside their own local-cum-ethnic languages, or have given up a local tongue, or are using it less, or see the next generation giving it up in favour of one or more languages offering better prospects. Significantly, for our purposes here, if users of a language at the sixth or seventh level are not only learning and using a language of wider local distribution but also learning English at school or using it in their work, then *two* kinds of pressure are being exerted on small languages that may not yet be endangered, but could become so in, say, fifty years' time.

Millions of busy people worldwide *can* and *do*

become successful trilinguals. In India this is often the norm: mother tongue, other Indian tongue, and English ('the window on the world'). Many native-speakers of English, however, do not use any other language beyond, say, school-days French. It is often therefore difficult for them to imagine or take seriously competition among languages *inside* a nation-state. I don't have Gaelic because local work and migration patterns meant that Highlanders and Islanders moved to the Lowland cities, where they acquired Lowland Scots as the language of work and their children learned Scottish English in school while using versions of Scots vernacular in the street and playground. Nothing simple there, nor anything simple elsewhere. Language ecologies are like bio-ecologies: they normally function well enough, but can be thrown off balance. I don't speak Gaelic, but I *do* know Scots and English natively, plus French and Persian non-natively. Maybe two out of three isn't too bad. And probably Lowland Scots will survive, partly because many people don't accept that it's a language. It may therefore manage to sneak around for quite some time, heavily disguised as bad English.

Meanwhile, at the big end of the scale

The spread since 1945 of 'world' or 'international' or 'global English' has been remarkable, which is why it has Category 1 to itself. Yet even if I'd put Levels 1 and 2 of 7 levels together as 1 of 6 levels, it would have made no difference to language reality. I didn't do this because the distributions are so different: Spanish, Chinese, and Hindi-Urdu massively concentrated in particular regions, with limited impact elsewhere, and Hindi-Urdu not well known as a single entity, even in India. As a result, this vast complex removes itself from the world discussion, despite its scale and significance. The following three factors apply to it, but in global terms relate more to Chinese, Spanish, and English:

- 1 Far more users of Chinese and Spanish are learning English than there are users of English learning Spanish and Chinese.
- 2 Far more users of Chinese and Spanish are learning English than are learning one another's language.
- 3 Far more users of other languages are learning English than are learning Chinese and Spanish.

However, we can note that Spanish occupies a unique role as the only language currently making inroads into the English-speaking world, and in its most powerful locale, the United States. The demographic and linguistic advance of Latino migrants into the US (from Mexico and further south) has been so marked in recent times that alarmed US linguistic conservatives campaign for English to be made the official federal language (as if this would make much difference). Ironically, in the most powerful fortress of English, enough people have been nervous enough to band together to protect the one language on the planet that is least in need of protection.

The Chinese complex is vast, but the majority of its speakers are Han (ethnic Chinese) living in the People's Republic of China. The vast majority of the world's non-native learners of Chinese are non-Han citizens of the PRC, with however a small but growing number of learners in neighbouring East Asia, their eyes on trade. Like Hindi-Urdu and Spanish, therefore, Chinese is (despite its size) a world-regional language, whereas English (with all its built-in spelling and other hurdles) is the closest we have come to universality in the use of a language, and so, even if English has nearly destroyed Gaelic and long ago upstaged Scots (which was my mother's tongue), I have learned to live with it, and, like my revered English teacher Miss Frances Anderson, at Woodside Senior Secondary School in Glasgow, love it as much as I love Scots. Not to forget French, Latin, Greek, and Persian.

In *The Future of English?* (1997), David Gradol suggests three options for English as a lingua franca in Asia: that it might keep this role indefinitely; that it might be supplanted by Mandarin; or that there might not in future be any Asian lingua franca. A few years further on, in 2005, it would seem that English will sustain its transnational role in Asia, as illustrated by recent developments in the Association of South East Asian Nations (ASEAN), which since its formation in 1967 has used English as its common medium. One must suppose that, when ASEAN members meet such prospective regional trading partners as China and India, English will continue in its current role. In that role, however, it will not be operating as an Atlantic language, but as an Asian lingua franca. And when in the nearish future ASEAN will be dealing with both India and China, English will be the broker language, working

together with Hindi-Urdu, Chinese, and Malay as key languages in South and East Asia.

One might call this *English Plus*. In the Americas it will be English plus Spanish (plus Portuguese for Brazil, and French for Quebec). In North and West Africa, it will be English plus Arabic (plus French). In West Asia, it will be Arabic plus English. In South Asia, it will be English plus Hindi-Urdu (South Indians preferring English to Hindi; Pakistanis comfortable with Urdu and English; Sri Lankans opting for English). In parts of Eastern Europe and Central Asia, it will be English plus Russian (with French in the wings in Romania, and German in the Balkans). In western Europe, however, it looks as though English will be the *lingua franca* (even for the French). That is, English will be everywhere, but often in partnership with other large languages. Not much help (alas) for the languages in Level 7, but less monolithic than some may fear. And maybe an ecology in which people from traditionally

multilingual backgrounds can work towards reducing the downward drift. ■

References and related reading

- Crystal, David. 1997. *English as a Global Language*. Cambridge: University Press.
— . 2nd edition, 2003.
— . 2000. *Language Death*. Cambridge: University Press.
- Graddol, David. 1997. *The Future of English?* London: The British Council.
- Kachru, Braj B. 1982. *The Other Tongue: English across Cultures*. Urbana: University of Illinois Press.
— . 1992. *The Other Tongue: English across Cultures*. [new collection]. Urbana & Chicago: University of Illinois Press.
— . 2005. *Asian Englishes beyond the Canon*. Hong Kong: University Press.
- McArthur, Tom. 1998. *The English Languages*. Cambridge: University Press.
— . 2002. *Oxford Guide to World English*. Oxford: University Press.
- Wright, Roger. 2004. 'Latin and English as world languages.' In *English Today* 80 (20:4).