

Web English —the future?

A.I.C. Monaghan

National Centre for Language Technology
Dublin City University

September 1997

Abstract

Electronic communication, and particularly the World Wide Web, is becoming increasingly indispensable in our daily lives. The vast majority of the information currently exchanged electronically is in English, and it might be assumed that this will promote the use of English. Rarely is the contrary view presented, that the adoption of English as a general-purpose medium for global communication will change the English language and perhaps even lead to the creation of a «Web English» which replaces native varieties.

Key words: English, World Wide Web, Future.

Table of Contents

- | | |
|-----------------------------------|-------------------|
| 1. Introduction | 4. The other view |
| 2. English on the Web | Bibliography |
| 3. Futures for English on the Web | |

1. Introduction

With roughly 350 million native speakers worldwide, English is second in the world language league. Chinese is first, with more than 700 million native speakers. Nobody knows how many people use Chinese as a second language, but it seems probable that the number of non-native speakers of English far outweighs the total for Chinese. If we count native and non-native English, then, English is probably the most widely-used language in the world.

An inevitable consequence of being spoken by the better part of a thousand million people is that the English spoken is not always the same English. There is an assumption that all Englishes are nevertheless mutually intelligible (otherwise it wouldn't be English) or at least that problems of intelligibility are confined to spoken communication, with text remaining relatively standard. I do not have space in this paper to argue against these assump-

tions, but see Horvath (1997) and the references cited therein. My own experience of different varieties of British English (Manchester, Newcastle, North London, Edinburgh, Dublin, Standard Scottish English, Hebridean English, Scots, Geordie, Hiberno-English and others) shows that unintelligibility is quite common and that problems do indeed carry over to text: examples include different vocabulary items, different syntactic constructions, and different meanings for the same word or construction. The variation in the English spoken or written by non-native English speakers is understandably broader: again, space is insufficient to argue that non-native speakers form the majority of users of English, or that the level of competence acquired by non-native speakers is generally quite low (secondary education only in most cases, and no formal education at all in many). In Europe, for instance, there are perhaps 70 million native English speakers and perhaps ten times as many learners.

Up until now, all the different varieties of English have survived side by side without too much destructive interference, because the language communities involved have been self-contained and to some extent isolated from each other. Generally speaking, American newspapers have predominantly been read by Americans, British radio and television has been broadcast to the British only, African poets and song-writers have commanded largely African audiences, and so forth. There have been a few notable exceptions, ranging from the serious to the trivial, and these in themselves are interesting:

1. In many fields of scientific endeavour, the language of international publication is English. Scientists from hundreds of different non-English-speaking countries and cultures are obliged to present their results in English. The result is a language which resembles English in many respects, but which is not any native variety. This new «International English» generally has an impoverished common vocabulary and syntax, and may be more intelligible to speakers of the author's native language than to native speakers of English.
2. English is the language of commercial pop music across most of the globe. This particular use of English seems to involve a variety spoken somewhere between Nashville and Hollywood. Native and non-native English speakers alike have adopted this variety of English for performance throughout the past three or four decades. Even in a forum such as the Eurovision Song Contest, which is specifically European and multilingual, this «Stage English» is the predominant language variety.
3. More recently, English-language television channels in the Northern Hemisphere have been broadcasting large amounts of Australian English, particularly soap operas during the peak children's viewing hours. This has resulted in an entire generation of British and Irish native speakers of English acquiring elements of Australian pronunciation. The most obvious acquisition has been of Australian declarative prosody, which ends with a rising intonation. In this variety, statements sound like questions?

This can be very confusing? It is certainly not a phenomenon native to most European English?

These three examples illustrate what happens when a more varied community decides to use English as a common language. In the case of the academic community, the need for intelligible communication of quite complex information leads to the creation of a semi-artificial language based on English which preserves the ability of the language to convey factual information but discards most of the nuances of English. In the case of pop music, the information content is not so important but a variety is chosen which has the desired tone —North American, earthy, quite urban, and racially neutral. In the case of Australian TV, we have an in-between case: those aspects of the variety, such as prosody, which can be assimilated into every-day life are readily adopted, but specific Australian vocabulary and syntax are not acquired (or at least not as readily). In the first and third cases, we have a blurring of boundaries between varieties of English, to the point of creating a new stripped-down variety in the first case. The second case is probably less representative of what happens when different varieties come into contact, but is presumably the result of one variety —controlling the US recording industry and the largest English-speaking market for pop music— being so dominant that it becomes impossible not to use it.

So, what will happen when everyone in the English-speaking world, native and non-native speakers, suddenly has access to billions of words of English from all over the world, written by native and non-native speakers of every variety of English? This is something which has never happened before. Its impact is expected to be much greater than that of radio or television, or even than the advent of literacy. This is the dream of those who advocate the Information Society: the availability of all the information you could wish for —texts, images, films, speech, multimedia— for a price, via the ever-growing network of computers known as the World Wide Web. Within ten years, it is suggested, almost every household in the industrialised world will have access to the Web.

2. English on the Web

Students, stockbrokers, scaffies, and senior citizens all surf the Web in pursuit of information and entertainment. It is notoriously difficult to obtain information about the overall size or content of the Web: it is not controlled by anyone, largely uncensored, contains multiple copies of much of its content, and is only beginning to be commercially exploited. It is estimated that the Web is growing by about a million words a day, that over 80% of its contents are in English, and that most of this is based in North America. Although the Web is more widely used in the USA than elsewhere, it is almost as widespread in Europe and Asia and is becoming increasingly common in Australasia and Africa.

The fact that at least 80% of the text on the Web is in English, and that usage of the Web is increasing rapidly worldwide, has caused great concern among speakers of languages other than English. The French, for example, have insisted on French language versions of Web sites within France, and the threat posed by the Web to the French language is taken very seriously (NoTIAL, 1997). In the Scandinavian countries, strategies for the transition to a society where all information is exchanged electronically (the so-called Information Society) include important actions to preserve native languages. The Norwegian government, for example, has stated that «Norwegian language and Norwegian user interfaces must be available in digital media and communications systems», and that «Stimulation of the production of Norwegian content must be a task worthy of priority by both the public and the private sectors», on the grounds that Norwegian «cultural development is dependent on a conscious approach to our national presence in the global electronic information networks». (Norwegian Government, 1996)

So far, there has not been any significant expression of concern for the future of English in the Information Society. It has generally been assumed that increased use and availability of the English language is a good thing for English speakers. This assumption may be correct, but it depends on two other assumptions:

1. The assumption that English text is relatively homogeneous, or at least that there are only a very small number (2, i.e. US and European) of distinct varieties, each of which will be provided for by the Web.
2. The assumption that the suppliers or authors of the English text on the Web will be mainly competent speakers of one of these varieties.

The first assumption is clearly false. Major varieties of English, with very different vocabularies, include at least European, American, Australasian, Indian, African, and Asian. Within each of these, there are important sub-groups: American splits into US, Canadian and Central/South American varieties; Australasian splits into Australian English, New Zealand English, and possibly others; and so on. In a continent the size of Africa, with some areas (e.g. South Africa) having large numbers of native English speakers, other areas (e.g. Kenya) having English as the main language of instruction in schools, and still other areas (e.g. Tanzania) having English as just one of several major languages, the number of different varieties and the chances of mutual incomprehension must be very high.

Let us take an example from closer to home: I, as a native speaker of a variety of British English, am confronted by the English of Southern Ireland, known as Hiberno-English. After four years of living in Dublin, I meet a man from a South Dublin suburb who asks me a question in English. He repeats the question several times, but I still have no idea what he is saying, and neither does the woman sitting next to me who is a native of County Cork. This is not an unusual occurrence: people from Donegal have great

difficulty understanding people from Kerry, and Dubliners frequently cannot comprehend speakers from the west coast of Ireland which is some 200km from Dublin.

The second assumption above, that the suppliers or authors of the English text on the Web will be mainly competent speakers of English, is equally false. Well-known examples of non-native English text include academic journals, tourism brochures, technical manuals and official forms. All of these texts are increasingly available on the Web. Here are some examples of text from these and other Web sources:

1. «the constant state of receptiveness in which speakers are found as they produce leads them to...» (International academic journal)
2. «University of XYZ suggest you ...», «This WWW server is a realisation of the CRI» (French university home page)
3. «somethings for the ladys» (East European entertainment site)
4. «a number of research projects that combine the remits of those for chairs or which stretch somewhat outside», (German university research site)
5. «Many airline companies provide direct services from main ports all over the world to Amsterdam Schiphol International Airport». (Dutch tourist site)
6. «nam» and «adress» (Official EU forms, English version)
7. «Concertation with other agents» (Document title, EU site)
8. «If you can not understand this page because of the poor English, would you try to review and correct it for us? We will be very glad if you do so. Please mail to ...» (Official Japanese Government site)
9. «Basic research that fosters buds of future science and technology»; «Advanced technologies for the utilization of space environment»; «Generation of channeling radiation and it's application; and other 7 subjects». (Japanese research site)

The proportion of non-native English speakers on the Web is high and increasing: with 80% of the Web in English, and use of the Web worldwide growing at an amazing rate, more and more non-native English speakers are writing and accessing Web documents in English. Within Europe, the European Commission estimates that 60% of businessmen accessing the Web are not native English speakers. Since English speakers were the first to make widespread use of the Web (hence the preponderance of English text), it is likely that the proportion of English native speakers will decline as the Web becomes more and more global. The effects of this decline are impossible to predict, but some possible futures are discussed in the next section.

3. Futures for English on the Web

The current situation, with most of the Web only accessible through English and an increasing amount of information only readily available on the Web,

is indeed increasing the use of English but it is also changing the English which is used. English on the Web is a hotch-potch of several native and non-native varieties. This makes communication on the Web unnatural even for native speakers of English. Whether this situation will continue is impossible to tell, but it is possible to imagine a range of possible futures for English on the Web:

1. A widely-used Web which is predominantly in English.
2. A Web which is predominantly in English, BUT which allows access using a range of languages through Language Engineering.
3. A Web where English is not the predominant language.
4. A Web which only appeals to professional users.

The first possibility, a widely-used Web predominantly in English, is the closest scenario to the present situation. The probable outcomes of such a situation are the blurring of boundaries between varieties of English, the development of an international *lingua franca* based on English but greatly simplified, and in the extreme case the marginalisation of native English speakers. This is a familiar pattern in the evolution of languages: the standardisation of Italian, for instance, involved picking one variety over others. Indeed, the evolution of English has included similar events: the choice of South-Eastern British English as a standard during the Middle Ages is one example, and other varieties of British English are still recovering from their loss of status. The difference in the case of English on the Web is that the variety which is chosen will be a mixture of native varieties and non-native varieties determined by the character of the electronic traffic on the Web. This variety could become a global pidgin or creole, and could even be adopted by native English speakers.

A second possible future situation might allow speakers of most languages to access the Web using their native language, at least in written form and possibly using speech input. This would require the availability of Language Engineering (LE) products for a range of languages, so that only monolingual speakers of English were obliged to access the Web using English. This is the optimal solution for all languages: for English it ensures that the language does not become a homogenised *lingua franca*, and for other languages it allows them to be used freely on the Web. Unfortunately, the current capabilities of Language Engineering for English and other languages are very limited (Truchot 1995, European Commission 1996), and considerable research and development work is necessary before LE can provide a multilingual Web.

A third possibility would be to make the Web less English. If the Web were the main repository of texts in French, Spanish, Chinese or even Old Irish there would be a strong motivation for users of those languages to access that information. Again, there is work to be done in the de-anglicisation of the Web.

A fourth possibility is to reduce the importance of the Web as a source of information. Most people still spend more time watching television than surfing the Web, and the promotion of languages other than English through the medium of TV has been very successful for Scottish Gaelic, Welsh, Catalan and other «minority languages». The current availability of a range of Englishes through media other than the Web is likely to delay the homogenisation of English in any case, but the survival of native English varieties can be helped by providing motivation for the use of languages other than English in as many contexts as possible.

The actual future importance of the Web in global terms may turn out to be much less than currently predicted: the size of the Web is already making communication slow, and the mass-market multimedia applications which are expected to commercialise the Web may never actually appear. The Web may simply fall into disuse except among the specialist and academic users for whom it was originally founded. This would perhaps be a less desirable future than the second possibility above, but it would still avoid the development of a global «basic English». However, since the likelihood is that the Web will continue to grow for several years to come, there are a number of steps which should be taken to promote the more desirable futures outlined above.

The factors determining the languages employed in human communication are basically twofold: practicality and motivation. Practicality will override motivation in cases where, for example, you wish to communicate with someone who doesn't understand the language you prefer to speak: thus, an Irish speaker may choose to use English when communicating with a tourist. Motivation will override practicality when there are several possible languages of communication: thus, a tourist learning Irish may make the effort to communicate in that language even when communication in English would be easier and quicker.

The effect of the Web, then, will depend on the possibilities available for communication and the motivations of the users. If the Web is the only means of access to essential or interesting information, and the only language understood by the network is English, then English will take over as the only possible language of communication and other languages will suffer. If the Web is accessible through a range of human languages, then the language of communication will depend on the motivation of the user. If information is only available through languages other than English, then users will be motivated to use those languages. If information sources such as newspapers, television and libraries remain more interesting than the Web, then the impact of the Web on global language use will be lessened.

A multilingual Web can be achieved in two ways: by providing technology which translates Web-based information into several languages, or by providing documents directly in several different languages. The first possibility depends on advanced LE products for each desired language, and even the most optimistic estimates only see this as a long-term option.

If the Web contained more information in languages other than English, this would increase the practicality for users to access information in a range of languages. Norway already has a policy of providing information in Norwegian on the Web, as do many other countries, and this is something which English speakers should encourage. Web sites based in English-speaking communities but intended for international consumption should have foreign language versions where possible, and university and other academic sites should ensure that their foreign language departments and book collections are well represented on the Web. As above, international collaboration would be most advantageous in this matter: commercial Web sites have already adopted mutual advertising whereby each site carries signposts to affiliated sites, and this could be done by academic sites with exchange partners or even research consortia.

The alternative to producing a multilingual Web is to reduce the importance and influence of the Web in global communication. Again, there are two aspects to the problem: reducing the attractiveness of the Web, and strengthening competing sources of information and entertainment.

The first aspect is unlikely to be resolved by anything other than the failure of communications technology. Unless the Web becomes too slow or too expensive to use, its attractiveness is likely to increase. Most countries see the Web as an opportunity rather than a threat: it is estimated that Web-based services will see an annual turnover of hundreds of billions of US dollars by the middle of the next decade, and that almost all official government business in many countries will be transacted electronically.

The second aspect appears to be much more readily solved, for speakers of all languages. There are many instances of «minority languages» whose use is on the increase because they have harnessed media other than the Web: I will briefly outline two such examples.

In Ireland, the minority language Irish is kept alive by a strong popular culture which is expressed in music, song, dance and literature. This culture has been supported by a national education policy which makes study of Irish compulsory from an early age and which favours Irish as the language of instruction. Ireland also produces television and radio programmes in Irish, and gives Irish equal status with English in all official communication. This policy has resulted in an increase in the use of Irish, despite massive cultural influence from the UK and the USA.

In the UK, the origin of the English language, there are several minority languages which thrive as a result of encouraging their cultural underpinnings. The most prominent of these is Welsh, which has equal status with English in Wales: it is the usual language of instruction in primary schools, and has its own television channel. Again, there is a popular culture of song and poetry which predates English. Welsh is on the increase, as are Scots Gaelic, Scots itself, Doric and even Cornish: there are plans for increased television coverage for Scots Gaelic, and cultural festivals in Doric and Cornish.

These are languages which are already minority languages, which are directly threatened by English (virtually all speakers of these languages are also native speakers of English), and which are nonetheless thriving: how have they achieved this? The reasons are superficially simple:

1. There is a strong popular culture which is distinct from the English-speaking world culture.
2. There is media exposure for the language.
3. There has been some official regional or national support for the language.

4. The Other View

Perhaps my arguments have not been convincing enough. Perhaps it might be a good thing if more people used English because of the Web. Look at the success of the Irish language, where most speakers are not native speakers but have learnt their Irish at school or after leaving school. The Irish language has undergone a process of smoothing over the past fifty years which has removed many of the local differences: standard Irish in schools, and national Irish language radio and television, have produced a common tongue where previously there were very different dialects. This has been good for Irish as a modern language, but it has produced a very different language from that spoken by octo- and septuagenarian native speakers, and it is questionable whether this new Irish will preserve the poetry, song and character of the language. Be that as it may, it is much harder to see the advantages of smoothing for English than for Irish. Merging the different varieties of a language with less than 100,000 mostly elderly native speakers on one small island may have saved that language from extinction, but merging the varieties of a language with 350 million native speakers all across the globe and an even larger number of non-native speakers will have quite a different effect. It may be that in the not too distant future most native speakers of English will find themselves in a similar position to the old people of rural Ireland: unable to understand the version of their language which is broadcast by the media and unable to communicate with their grandchildren because the language has changed so much in so short a time.

Bibliography

- EUROPEAN COMMISSION (1996). *Language technology: from the tower of Babel to the global village*. Luxembourg: Office for Official Publications of the European Communities.
- HORVATH, J. (1997). «The lingua franca of the Net». Telepolis archive, <http://www.heise.de/tp/english/inhalt/te/1149/1.html>
- NORWEGIAN GOVERNMENT (1996). *The Norwegian way to the information society: bit by bit*. Report from the State Secretary Committee for IT. Oslo: Ministry of Transport and Communication.

- NOTIAL (1997). *La normalisation: Clef de voute de la francisation des inforoutes*. ISBN 2-550-31674-6 (French Government report).
- TRUCHOT, C. (ed.) (1995). *Compte rendu du séminaire «Construire le Pluralisme Linguistique»*. Besançon: European Language Observatory.