12

Electronic English

The vision I have for the Web is about anyth potentially connected with anything.

Tim Berners-Lee, Weaving the

Multilingualism on the Internet can be seen as a labove all irreversible inevitability. In this perspective to make fun of the wet blankets who only speak to about the supremacy of English.

Marcel Grangier, in email to Marie Lebert (14 Jan

A major risk for humanity.

Jacques Chirac, The Economist (21 Decen

It is obvious that new electronic channels such as email ar Web have brought an enormous quantitative increase in the around the world. English had a headstart over other e-revolution, which took place chiefly in the 1980s and ea because of the way the technology developed in an English (the United States), but also because the Internet was tailor-n using the Roman alphabet with no diacritics (such as the rêvé and the umlaut in German Frühling). English became its However, the electronic revolution has also boosted the use and the proportion of English on the Internet has been ste the 1990s. Since 2004, when Facebook began (quickly follow 2005 and Twitter in 2006), development of social media has factor in altering the linguistic character of the Internet, as the majority of users are not just reading content on the Inadding to it, and most of these additions are expressed in the Even endangered languages can benefit from the Internet, networks of users who can now converse regularly around the world in a lesser-used tongue.

It is too soon to say what permanent effect the Internet will have on languages. Electronically mediated communication has been in routine use only since the early 1990s - the World Wide Web was introduced in 1991 - and a period of two or three decades is an eye-blink in the history of a language. It takes a lot longer for a change to emerge, for individuals to get used to its novelty, for them to start using it in everyday speech and writing, and for it eventually to become so widely used that it becomes a permanent feature of a language, recorded in dictionaries, grammars and manuals of style. There are already some tell-tale signs of what may happen, but conclusions have to be tentative

General statements about Electronic English (EE) are inevitably tentative because of the nature of the medium. Its size, for a start, makes it difficult to manage - there has never been a corpus of language data as large as this one, containing more written language than all the libraries in the world combined. Then there is its diversity, which defies linguistic generalization – the stylistic range of EE includes the vast outputs found in email, chatrooms, the Web, virtual worlds, blogging, instant messaging, text messaging and Twitter, as well as the increasing amount of linguistic communication encountered in social networking forums such as Facebook, each output presenting different communicative perspectives, properties, strategies and expectations.

The speed of change makes it difficult to keep pace. How can we generalize about the style of emails, for example? When it first became prevalent, in the mid-90s, the average age of emailers was in the 20s, and it has steadily risen. Nielsen Surveys showed it to have reached the mid-30s by 2007. Many emailers are now senior citizens. Similarly, the average age of a Facebook user has sharply risen since 2004, from a predominantly young person's medium to a medium for everyone - in the United States the average is now around 40 years. The consequence is that the original colloquial and radical style of emails and forum posts (with their deviant spelling, punctuation and capitalization and a copious use of non-standard expressions) has been supplemented by a more conservative and formal style, as older people introduce norms derived from the standard language.

But it is not solely a matter of age. The pragmatic purpose of a piece of EE can alter, sometimes overnight. A good example is the short-messaging service Twitter, which, when it first arrived, used the prompt 'What are you doing?' The result was a range of tweets that were inward-looking, using lots of first-person pronouns and present tenses ('I'm really excited to be watching Hamlet at the Globe.'). Then in November 2009, Twitter changed its prompt to 'What's happening?' This made the tweets outward-looking, with lots of thirdperson pronouns and a wider range of tense-forms ('New Globe programme

was published yesterday. It's brilliant.'). The result was a shift in the aims and linguistic character of Twitter English, which took on more of the features of a news service, as well as attracting more advertising content. Although limited to 140 characters (20 less than the 160-character limit in text messages, due to the need to leave space for the tweeter's identity), this still allows quite a substantial message of 20 or 30 words (in English).

Is Electronic English changing the language?

When we consider EE as a species of written language and compare it with traditional modes of writing, certain novel properties are immediately apparent. However, these properties are nothing to do with the standard conception of writing as a combination of vocabulary, grammar and orthography. EE has certainly introduced a few thousand new words, but these make up only a tiny fraction of the million+ words that exist. New words enter English every day, from all kinds of sources, and the Internet accounts for only a fraction of them. There was a peak of novel word-creation when the

New words in Electronic English

In addition to the many new words or senses that have entered English as part of the language of computing (window, file, paste, search, options, software, 404 error, crash, multimedia ...), the Internet has extended English vocabulary in several ways:

Coinages: blog (a shortening of Web log), netiquette, netizen, tweet, unfriend Compounds: cost-per-click, dotcom, hotlink, webcam, twittersphere, voicemail

Prefixes: e- as in e-cash, e-commerce, e-courses, e-mail

cyber- as in cybercafe, cyberculture, cybersex, cyberspace hyper- as in hyperfiction, hyperlink, hypertext, hyperzine

Suffixes: -ware as in courseware, firmware, freeware, spyware

-bot ('a software application that runs automated tasks' - from robot) as in annoybot, chatterbot, spybot, mailbot

Words converted from one class to another, such as nouns becoming verbs: bookmark, flame, message, text

Metaphorical uses of existing words: browse, chat, cookie, hack, link, menu, portal, spam, surf, virus, wizard

Abbreviations: IP (Internet protocol), FAQ (Frequently answered questions), URL (Uniform resource locator), BTW or btw (by the way)

Playful word-creation: Twictionary (Twitter dictionary), vaxen (plural of 'VAX computer', on analogy with oxen), bloggerrhoea (the condition of someone who blogs too much)

Web was created, and subsequent technological innovations (such as blogging and texting) have each added to the lexicon, but there is nothing in the panel above that we have not seen before in the history of English. There is nothing revolutionary here.

Similarly, the grammar of written EE displays no novelty in comparison with what was used before. We do not see different word-orders or wordendings. There have been occasional experiments, such as the use of a novel -z plural to identify illegal files (tunes vs. tunez, films vs. filmz), but these were short-lived and hardly justify the application of the label 'revolutionary' to a description of EE grammar. Certainly, EE is often stylistically novel in its use of grammatical constructions - for example, using short and abbreviated sentences in texting and tweeting, or colloquial syntax in blogs - but the history of English shows many earlier examples of such features. Very short sentences can be seen in Old English texts, such as the Colloquy of Ælfric, written around the year 1000.

The most distinctive feature of EE lies in the way people manipulate certain features of the orthography. Some Internet genres motivate a minimalist use of punctuation and capitalization, as seen in emails, texts or tweets that have few or no punctuation marks or capital letters. But these same genres also display the opposite trend, using repeated characters and capitalization, as in Fantastic!!! and NOOOooh!!!! These are the nearest equivalent written language has to spoken-language features like voice quality, pitch range and loudness what phoneticians refer to as the prosodic and paralinguistic dimensions that the human voice can draw on as an expressive channel of communication. They are a consequence of the way Internet users have tried to make the written language of their new medium more conversational in tone (see Chapter 10).

Another distinctive feature of EE orthography is that, apart from in audio/ video interactions (such as Skype or iChat), it lacks the facial expressions, gestures and conventions of body posture and distance that are so critical in expressing personal opinions and attitudes and in moderating social relationships. The limitation was noted early in the development of the medium and led to the introduction of smileys or emoticons (a blend of 'emotive icon'), such as:

- :(or :-(unhappy
- :)) or :-)) very happy
- ;) or ;-) wink, smirk

Today, dozens of emoticons are offered by message exchange systems and these have been extended by the use of Japanese emoji ('picture letters'), which

present a wide range of everyday objects and actions in pictographic (and usually colourful) form, such as:



It is plain that emoticons are a potentially helpful way of capturing some of the basic features of facial expression and body movement, but their semantic role is limited. They can forestall a gross misperception of a speaker's intent, but an individual emoticon still allows a large number of readings - the basic smile:) could mean happiness, joke, sympathy, good mood, delight, amusement, sarcasm, etc. - which can only be disambiguated by referring to the verbal context. Without care, moreover, they can foster misunderstanding - adding a smile to an utterance that is plainly angry can increase rather than decrease the force of the 'flame'. So it is not surprising to see that the use of emoticons has lessened, as time goes by. People have realized that they do not solve all communication problems in EE, and may even add to them.

There are a few other noticeable orthographic features, such as the use of bicapitalization - or bicaps (AskJeeves, YouTube) and proper names that begin with a lower-case letter (eBay, iMac). Some punctuation marks have acquired new values, such as the use of asterisks to express emphasis (a *very* important point), @ in email addresses, and underbars to replace underlining (a production of _Macbeth_) which is not possible in all types of electronic input. Some of these features have begun to be used outside their originally Internet settings. The hashtag (#), for example, achieved a new lease of life after 2006, following its use on Twitter as a way of identifying a search (one would find tweets to do with dictionaries by going to #dictionary). It then became more widely used as a way of adding a comment to anything one might write, whether on the Internet or not. So here, for example, we might add: #interesting point. In 2015, BBC Radio 2 and Oxford University Press made hashtag the Children's Word of the Year, having observed a remarkable increase in its comment-function frequency in schoolchildren's narrative writing.

Despite these novelties, the writing system of EE on the whole looks very similar to what existed before the Internet arrived, and this should be a point

of reassurance to those who feel that the arrival of electronically mediated communication is somehow going to undermine the standard language and the educational goal of learning to write good English prose. All that has happened is that the various genres of Internet communication have contributed lively and versatile additions to the already rich tapestry of English language varieties, especially offering the opportunity to communicate in more informal, vernacular styles than were available in traditional writing. Young people in fact are reading and writing more, and in a wider range of styles, than their pre-Internet counterparts did. Similar stylistic expansions have accompanied every technological innovation in the history of language, such as the evolution of new graphic conventions following the arrival of printing and of new genres of oral expression following the introduction of broadcasting. The Internet has similarly extended the language's stylistic range - or perhaps we should say, is extending, for the technological revolution that began with the Internet is by no means over.

Is EE a revolution?

If EE is a revolution, its novelty lies not in vocabulary, grammar and orthography, nor in the development of new styles of discourse, but in the opportunities it presents for fresh kinds of communicative activity. There are several fundamental differences between writing on the page and writing on the screen.

First, there is a contrast with the space-bound character of traditional writing - the fact that a piece of text is static and permanent on the page. If something is written down, repeated reference to it will encounter an unchanged text. Putting it like this, we can see immediately that EE is not by any means like conventional writing. A 'page' on the Web often varies from encounter to encounter (and all have the option of varying, even if pageowners choose not to take it) for several possible reasons: its factual content might have been updated, its advertising sponsor might have changed or its graphic designer might have added new features. Nor is the writing that we see necessarily static, given the technical options available that allow text to move around the screen, disappear/reappear, change colour and so on. From a user point of view, there are opportunities to 'interfere' with the text in all kinds of ways that are not possible in traditional writing. A page, once downloaded to the user's screen, may have its text cut, added to, revised, annotated, even totally restructured, in ways that nonetheless retain the character of the original. The possibilities are causing not a little anxiety among those concerned about issues of ownership, copyright and forgery.

Second, we see differences between some varieties of EE and traditional writing when we ask how complex, elaborate or contrived they are. Certain varieties are very similar to what happened before the Internet arrived. In particular, the Web allows the same range of planning and structural complexity as would be seen in writing and printing offline. But for chatgroups, virtual worlds, instant messaging, and social media generally, where the pressure is strong to communicate rapidly, there is much less complexity and forward planning (see the notions of interactiveness, emotive expression and online processing, pp. 206-7). Blogs vary greatly in their constructional complexity - some are highly crafted; others are wildly erratic, when compared with the norms of the standard written language. Emails also vary - some people are happy to send messages with no revision at all, not caring if typing errors, spelling mistakes and other anomalies are included in their messages; others take as many pains to revise their messages as they would in non-EE settings.

Third, we can compare the factual content of EE and traditional writing. The majority of the latter is factually communicative, as is evident from the vast amount of reference material in libraries. A focus on fact is also evident on the Web, where most varieties of written language (legal, religious and so on) can be found with little stylistic change other than an adaptation to the electronic medium. Many blogs and emails are also full of factual content. But other EE situations are less so. Chatgroups and social media forums vary enormously the more academic and professional they are, the more likely they are to be factual in aim; social and ludic chatgroups, on the other hand, routinely contain sequences that have negligible factual content. Instant message exchanges are also highly variable, sometimes containing a great deal of information, sometimes being wholly devoted to social chit-chat.

Fourth, traditional writing is graphically rich, as we can immediately see from the pages of many a fashion magazine. The Web has reflected this richness, but greatly increased it, the technology putting into the hands of the ordinary user a range of typographic and colour variation that far exceeds the pen, the typewriter and the early word processor, and allowing further options not available to conventional publishing, such as animated text, hypertext links and multimedia support (sound, video, film). On the other hand, as typographers and graphic designers have repeatedly pointed out, just because a new visual language is available to everyone does not mean that everyone can use it well. Despite the provision of a wide range of guides to Internet design and desk-top publishing, examples of illegibility, visual confusion, over-ornamentation and other inadequacies abound. They are compounded by the limitations of the medium, which cause no problem if respected, but which are often ignored, as when we encounter screenfuls of unbroken text, paragraphs that scroll downwards interminably or text that scrolls awkwardly off the right-hand side of the screen. The problems of graphic translatability are only beginning to be appreciated - that it is not possible to take a paper-based text and put it on a screen without rethinking the graphic presentation and even, sometimes, the content of the message.

The Internet, then, offers new communicative possibilities in the way people can manipulate written language, and it is this that justifies those who say it is a revolution. One of the consequences has been the emergence of new kinds of text.

New kinds of text

Every time a new technology arrives, we see the growth of new kinds of discourse, reflecting the aims and intentions of the users. Printing introduced us to such notions as newspapers, chapter-organization and indexes. Broadcasting brought sports commentary, news-reading and weather-forecasting. EE is no different. The content displayed on a screen presents a variety of textual spaces whose purpose varies. There is a scale of online adaptability. At one extreme, we find texts where no adaptation to EE has been made - a pdf of an article on screen, for example, with no search or other facilities - in which case, any linguistic analysis would be identical with that of the corresponding offline text. At the other extreme, we find written texts that have no counterpart in the offline world. Here are some examples.

Texts whose aim is to defeat spam filters

We only have to look in our email junk folder to discover a world of novel texts whose linguistic properties are hugely idiosyncratic:

supr vi-agra online now znwygghsxp VI @ GRA 75% off regular xxp wybzz lusfg fully stocked online pharmac'y Great deals, prescription d[rugs

The aim, clearly, is to make the language sufficiently deviant that the messages will not be excluded by spam filters, but not so deviant that they will be unintelligible. It is possible to see a linguistic rationale in the graphological variations in the word Viagra, for example, introduced to ensure that it avoids the word-matching function in a filter. We may find the letters spaced (V i a g r a), transposed (Viarga), duplicated (Viaggra) or separated by arbitrary symbols (Vi*agra). However, there are only so many options and these can to a large extent be predicted by sophisticated filtering techniques. There have been huge advances here since the early days when linguistically naive software, having been told to ban anything containing the string S-E-X, disallowed messages about Sussex, Essex and many another innocent term.

Texts whose aim is to save time, energy or money

Text-messaging (a different sense of the term text, note) is a good example of a genre whose linguistic characteristics have evolved partly as a response



to technological limitations. The limitation to 160 characters (for Roman alphabets) motivated an increased use of non-standard words, using logograms (C for see, 2 for to), initialisms (BTW 'by the way', LOL 'laughing out loud'), shortenings (bro 'brother', hol 'holiday') and other abbreviatory conventions (xInt 'excellent', msg 'message'). The important word is 'partly'. Most of these abbreviations were being used in EE long before mobile phones became a routine part of our lives. Indeed, abbreviations such as c for see and u for you have a long-standing presence in English. They were a popular feature of the word-games played in the Victorian England of the nineteenth century. The motivation to use them clearly goes well beyond the ergonomic, as their playful character provides entertainment value as an end in itself as well as increasing rapport between participants.

Multi-authored texts

Wiki pages, such as those seen on Wikipedia, are also a new textual phenomenon. They are the result of an indefinite number of interventions by an indefinite number of individuals over an indefinite number of periods of time (which become increasingly evidenced as time goes by). From a linguistic point of view, the result is pages that are temporally and stylistically heterogeneous. We find noticeable differences, such as standard and non-standard language coexisting on the same page, often because some of the contributors are communicating in a second language in which they are non-fluent. Tenses go all over the place, as this example illustrates (reproduced exactly as it appeared in Wikipedia):

Following his resignation, Mubarak did not make any media appearances. With the exception of family and a close circle of aides, he reportedly refused to talk to anyone, even his supporters. His health was speculated to be rapidly deteriorating with some reports even alleging him to be in a coma. Most sources claim that he is not longer interested in performing any duties and wants to 'die in Sharm El-Sheikh.' [59][60]

On 28 February 2011, the General Prosecutor of Egypt issued an order prohibiting Mubarak and his family from leaving Egypt. It was reported that the former president was in contact with his lawyer in case of possible criminal charges against him. [61] As a result, Mubarak and his family had been under house arrest at a presidential palace in the Red Sea resort of Sharm el-Sheikh. [62] On Wednesday 13 April 2011 Egyptian prosecutors said they had detained former president Hosni Mubarak for 15 days, facing questioning about corruption and abuse of power, few hours after he was hospitalized in the resort of Sharm el Sheik. [63]

Note the way for example, we move from past tense to present tense in paragraph 1, and from was to had in paragraph 2. Note also the way former president Hosni Mubarak is introduced in the last sentence, as if this were a new topic in the discourse. Note the three different spellings of the Red Sea resort. And how are we to interpret such non-standard usages as was speculated, in case of and few hours?

In pages like this, traditional notions of stylistic coherence, with respect to level of formality, technicality and individuality, no longer apply, though a certain amount of accommodation is apparent, either because contributors sense the properties of each other's style, or a piece of software alters contributions (e.g. removing obscenities), or a moderator introduces a degree of levelling. The pages are also semantically and pragmatically heterogeneous, as the intentions behind the various contributions vary greatly. Wiki articles on sensitive topics illustrate this most clearly, with judicious observations competing with contributions that range from mild through moderate to severe in the subjectivity of their opinions. And one never knows whether a change introduced in a wiki context is factual or fictitious, innocent or malicious.

Anonymous texts

Of course, EE is not the first medium to allow interaction between individuals who wish to remain anonymous, as we know from the history of telephone and amateur radio; but it is certainly unprecedented in the scale and range of situations in which people can hide their identity in the written language, especially in chatgroups, blogging and social networking. These situations routinely contain individuals who are talking to each other under nicknames (nicks), which may be an assumed first-name, a fantasy description (topdude, sexstar), or a mythical character or role (rockman, elfslayer). Operating behind a false persona seems to make people less inhibited - they may feel emboldened to write more and in different ways from their real-world linguistic repertoire. They must also expect to receive messages from others who are likewise less inhibited and be prepared for negative outcomes. There are obviously inherent risks in talking to someone we do not know, and instances of harassment, insulting or aggressive language, and subterfuge are legion. Terminology has evolved to identify them, such as flaming, spoofing, trolling and lurking. New conventions have evolved, such as the use of CAPITALS to express 'shouting'. While all of these phenomena have a history in traditional mediums, EE makes them present in the public domain to an extent that was not encountered before.

Automatically translated texts

The last few years have seen a marked increase in the number of automatic translation services, such as Google Translate, available for use by the general public. This follows a period when 'machine translation', as it was often called, was considered to be either impossible or too expensive to be taken



seriously. The change came about as a result of the development of faster and more powerful computers, new computational techniques of pattern matching and a more sophisticated awareness of the linguistic issues. A huge amount of progress has been made, so that for most of the world's widely used languages it is now possible to find software that will translate a text sufficiently well to enable users to obtain a general sense of the content. Total accuracy and acceptability is still some way off, however. Here is one system's English translation of an extract from the Swedish government's website:

An important manifestation of that we in Sweden live in a democracy is the general elections held every four years. Then the country's seven million voters an opportunity to influence which parties will represent the Swedish people in parliamentary, county councils and municipalities.

For comparison, here is the original:

Ett viktigt uttryck för att vi i Sverige lever i en demokrati är de allmänna valen som hålls vart fjärde år. Då har landets sju miljoner röstberättigade en möjlighet att påverka vilka partier som ska representera svenska folket i riksdag, landsting och kommuner.

The grammatical errors (underlined) are obvious, to anyone who knows standard English. Any errors in the lexical equivalents – such as whether the choices made in English for the Swedish political bodies are the best ones – would be less easy to see. But the translation does succeed in conveying what the text is about and it would not take a human translator long to polish the text to make it acceptable. The official website in English looks quite different, suggesting that a good translation often requires considerable structural text changes:

General elections which are held every four years are an important expression of the fact that we in Sweden live in a democracy. As one of the approximately 7 million people in the country entitled to vote, you are given an opportunity to influence which parties are to represent you in the Riksdag, county council and municipal council.

Rather more difficult to handle are texts that present colloquial or literary content, with informal vocabulary, idioms and unusual structures. Here, computer-aided translation is more likely to yield good results - human translators use a computer to provide a first approximation and then refine what they see. But for all texts, it is wise at present to treat translations by machine with caution, especially if important decisions (such as of a legal or financial kind) depend on the outcome.

Permanently evolving texts

All the texts so far illustrated in this book have one thing in common: they are easily identifiable and determinate. They have definable physical boundaries, which can be spatial (e.g. letters and books) or temporal (e.g. broadcasts and interviews). They are created at a specific point in time; and once created, they are static and permanent. Each text has a single authorial or presenting voice (even in cases of multiple authorship of books and papers), and that authorship is either known or can easily be established (except in some historical contexts). It is a stable, familiar, comfortable world. And what the Internet has done is remove the stability, familiarity and comfort.

Written texts are defined by their physical boundaries: the edges of the page, the covers of the book, the border of the road sign ... Spoken texts are defined by their temporal boundaries: the arrival and departure of participants in a conversation, the beginning and end of a broadcast, the opening and closing of a lecture ... Internet texts are more problematic. Sometimes, as with a text message or an instant-message exchange, we can clearly identify the start and the finish. But with most Internet outputs there are awkward decisions to be made. In particular, how are we to define the boundaries of a text that is ongoing? People can now routinely add to a text posted online, either short-term (as in the immediate response to a news story), or medium- or long-term, as in comments posted to a blog, bulletin board or other forum. The linguist Ferdinand de Saussure's classical distinction between synchronic and diachronic does not adapt well to electronically mediated communication, where everything is diachronic, time-stampable to a micro-level. Texts are classically treated as synchronic entities, by which we mean we disregard the changes that were made during the process of composition and treat the finished product as if time did not exist. But with many e-texts there is no finished product. I can today post a message to a forum discussion on page X from 2004. From a linguistic point of view, we cannot say that we now have a new synchronic iteration of X, because the language has changed in the interim. I might comment that the discussion reads like something 'out of Facebook' - which is a comment that could be made only after 2004, when that network began. Issues of this kind present linguists with unprecedented problems.

The future of EE

The biggest question marks to do with linguistic change on the Internet relate to the way the technology is rapidly developing. Most of the observations about written language in this chapter are based on what can be seen on the large screen of a computer. But it is a fact that Internet access is becoming increasingly mobile. Indeed, in some parts of the world, where a wired electricity supply is unreliable or absent (such as a great deal of Africa), the only way of reaching the Internet is via mobile phones. So what happens, in terms of legibility, when a page containing a large amount of visually encoded information is presented on a small mobile screen? How is the information re-organized? What is lost and what is gained? If, as the mobile phone industry is predicting, the majority of Internet access will soon be through handheld devices, then how relevant will be all the generalizations about the character of EE that have hitherto been based only on an analysis of large-screen displays?

This chapter has largely focused on written language. The main issue for the future will be how to deal with the increased presence of spoken outputs, as a result of growth in Voice over the Internet and mobile communication. There are several new kinds of speech situation here, such as automatic speech-to-text translation (as when voicemail is turned into text messages), text-to-speech translation (as when a web page is read aloud), voice recognition interaction (as when we tell the washing machine what to do) and voice synthesis (as when we listen to GPS driving instructions). The techniques are a long way from perfection - systems have recurrent problems with regional accents, speed of speech, background noise and the interpretation of proper names. But these will reduce as time goes by. The Internet will become a spoken as well as a written medium, and EE will expand further as it copes with new kinds of speech output. It will also enable us to have intimate contact with the way new varieties of English are emerging globally, both in speech and writing - a late twentieth-century development to which we turn in our final chapter.