Online Englishes

Mark Warschauer, Rebecca Black and Yen-Lin Chou

Introduction

The last few decades have witnessed some of the most rapid changes in human communication in world history. Though the internet was barely known a quarter of a century ago, today some 1.5 billion people around the world read, write and communicate online (Miniwatts Marketing Research 2008). An estimated 55 billion emails are sent every day, not including spam (Grossman 2008), and the blogging search engine Technorati is tracking some 133 million blogs around the world (Technorati 2008). From knowledge workers to office staff to teenage youth, large numbers of people around the world rely extensively on computer-mediated communication.

A disproportionate amount of this global communication is conducted in English. More than half of the .com and .net internet sites in the world are hosted in the US (Paolillo 2005), and the nine most heavily visited websites in the world are all in English (Alexa 2008). An estimated 29.4 per cent of world internet users are native speakers of English (Miniwatts Marketing Group 2008), and English has become the dominant lingua franca for cross-language communication online (Crystal 2001; Paolillo 2005).

Online communication is different than previous forms of interaction in many important ways. Online, large numbers of people from around the world can interact at the same time in a single forum. While interacting at a fast pace, they can still maintain a written archive of their communication. People can quickly encounter and get to know large numbers of strangers, and they can stay in constant close communication with friends at almost all hours of the day. They can publish their reports or multimedia documents for virtually free, and they can hotlink parts of their texts to link to the words of others.

For all of these reasons, online communication is engendering its own styles, genres and forms of English. Some people contend that it is resulting in the bastardization of English, the ruining of standards, and the misinformation of the public, while others contend that it is democratizing English by extending new forms of low-cost interaction, collaboration, and publishing to native and non-native English speakers around the world. While there are certainly elements of truth in both arguments, there is no doubt
that online Englishes are challenging prior notions of who the language belongs to, whose voices are heard, and who contributes to knowledge formation and dissemination.

Whose language?

Anatoly Voronov, then the director of a pioneering telecommunications network in Russia, spoke for many around the world when in 1996 he called the English-dominant World Wide Web ‘the ultimate act of intellectual colonialism’ (Spector 1996: 1). In France, Russia, China, the Middle East and elsewhere, government officials, language rights activists and others were up in arms about the fact that upwards of 80 per cent of the web’s content was in a single colonial language (Warschauer 2000a).

Since then, these concerns have subsided, as the internet has become much more multi-lingual. The percentage of English online has fallen by about half, with the amount of online content growing rapidly in both major languages and minor languages (Pimienta 2005). For example, Wikipedia alone has versions in 262 languages, 163 of which include 1,000 or more articles, and a number of which are endangered (Cohen 2006; Wikipedia 2009a).

The growth of multiple languages online undermines neither the internet’s use as a medium for communication across language groups nor the role of English as dominant lingua franca in such cross-linguistic contact. This lingua franca role both corresponds to, and has accelerated, the already prominent role of English in international media, political and business communication at the advent of the internet (Crystal 2003). At first glance, the pre-eminence of English as the de facto global lingua franca would seem to privilege native English speakers, who can participate effortlessly in international online fora. However, by simultaneously facilitating daily communication in English by hundreds of millions of non-native speakers around the world, this trend also calls into question who controls English and sets its standards. There is thus a growing movement around the world to teach a denationalized version of English based on local and regional standards of pronunciation, syntax and usage, rather than US or British English (Warschauer 2000b), and to use a simplified global English rather than US or British English in international business correspondence (McAlpine 2006). At the same time, the US- and British-dominated English language teaching (ELT) publishing industry is being increasingly challenged by competitors from outer or expanding circle (Kachru 1990) countries such as Singapore, Israel, Greece and Spain (Francis 2000). This internationalization of English was brought home to the first author of this chapter when I worked on a large ELT project in Cairo, and my Egyptian colleagues revised my English language correspondence to ensure that it met the standards of pragmatics and politeness of English language communication in Egypt, even when that communication was directed from one American to another (see discussion in Warschauer 2000b).

Stultified norms of what constitutes English are also being challenged by the widespread use of highly colloquial, informal and hybrid forms of interaction referred to as NetSpeak (Crystal 2004). These new forms are especially prominent in highly interactive forms of computer communication, such as electronic mailing (emailing), instant messaging (IMing), internet-relay chat (IRC or chatting) and short-messaging service (SMS, also known as texting). A great deal of public rhetoric is grounded in what Crystal (2001: 1) calls a ‘genre of worry’ that focuses on the potentially corruptive nature of
online registers and the idea that non-standard linguistic conventions associated with electronic media are spilling over into offline writing and conversation. Scholarly research surrounding these forms of computer-mediated communication (CMC) has tended to fall into two distinct camps: studies celebrating the unique nature of online registers; and studies disavowing any significant difference between on- and offline communication, save for the medium. Of late, such scholarship has turned towards a more holistic approach to understanding online discourse, emphasizing the interplay of technical and contextual factors.

Electronic mail

Email, which predates the internet, is an asynchronous form of online communication that allows users to write, send, save and sort electronic messages. When it came into common use in the 1990s, email was heralded as a revolutionary medium that would change the face of communication. Early examinations of the linguistic features of email suggested that users’ language tended to be less formal, less lexically sophisticated, and less grammatically and orthographically correct than paper-based prose (Baron 1998; Crystal 2001). Scholarly analysis of email and similar forms of CMC also gave rise to preliminary discussions about electronic text as a new hybrid communicative mode that blurred the distinction between spoken and written language (Ferrara et al. 1991; Werry 1996; Yates 1996).

In spite of this auspicious beginning, in terms of transformative linguistic and generic potential, email has continued, in Herring’s (2004: 27) words, ‘slouching toward the ordinary’. No longer on the cutting edge of information and communication technologies (ICTs), email is viewed as passé by youth, and is often used by adults in lieu of paper letters, announcements and memos. The English language forms and grammatical conventions for personal and business interactions conducted via email have come to mimic their print-based counterparts to a great degree (Crystal 2001). Some exceptions include the aforementioned informality that often manifests in a lack of salutations, an extended range of punctuation (e.g. ‘!!!’), and a reduced use of capitalization (Crystal 2001).

Instant messaging and chatting

IMing and chatting are real-time or synchronous forms of online communication that came into popular use in the 1990s. The primary difference between IMing and chatting is that IMing only allows dyadic communication while chatting allows multiple users to exchange messages at the same time in what is known as a chat room. According to Pew Internet and American Life surveys, around 53 million online adults (Shiu and Lenhart 2004) and around 13 million online teens (Leinhart et al. 2001) use IM on a daily basis, with around 41 per cent of working internet users using IM in the workplace (Madden and Jones 2008). Recent studies have shown that IM is more than just a communicative medium; it also serves as a way for youth to strengthen and expand social networks (Lewis and Fabos 2005), and as a means of self-expression via customized user profiles, buddy icons and away messages (Shiu and Lenhart 2004).

Due to their synchronous nature, IM and chat interactions, more so than email, tend to take on a highly informal, conversational format and have been catalysts for a great
deal of public concern surrounding the possible deleterious effects of online communication on the English language. For example, in a New York Times article, a teacher expressed concern over abbreviations such as ‘u, r, ur, b4, wuz, cuz, 2’, appearing of late in student writing. According to the article, such abbreviations are part of an ‘online lingua franca: English adapted for the spitfire conversational style of internet instant messaging’ (Lee 2002: eighth paragraph). However, the media also have described this ‘online lingua franca’ as ‘the bastardization of language’ (O’Connor 2005, cited in Tagliamonte and Denis 2008: 4) and ‘the linguistic ruin’ (Axtman 2002, cited in Tagliamonte and Denis 2008: 4) of modern times.

Public concern about language change seems to stem from several discourse features that are commonly used in IM and other forms of online communication. One such feature is the tendency towards the aforementioned abbreviations. Other common features include acronyms and initialisms, which are abbreviations formed using the initial letters or syllables of a phrase. Abbreviations typically associated with IM and chat are lol (laugh out loud), b4 (be right back), afk (away from keyboard), asl (age, sex, location). America Online, provider of AIM, the first widely used IM program, hosts a website with a list of AIM acronyms (America Online 2008). Another discourse feature commonly associated with online communication is the emoticon. The word ‘emoticon’ comes from a portmanteau of the words emot (or emotion) and icon and it describes graphic or keyboard representations of facial and bodily expressions or emotional content. Common emoticons include :) (smiling face), ^_~ (Asian smiling face), :-( (face with tears), @_@ (surprised face), and XD (mischievous face). Rebus forms of writing are also commonly associated with IM and, as will be discussed in the next section, SMS. Common rebuses include aar8 (at any rate), b4n (by for now) and cul8r (see you later).

Linguists, on the other hand, have proposed that IM language use is much less radical than the press suggests. For example, Baron’s (2004) study, based on a corpus of US college students’ instant messages, found that only 0.3 per cent of words were common IM abbreviations, less than 0.8 per cent were common IM acronyms, only 0.4 per cent were emoticons, and that only 65.3 per cent of contracted word forms were used. A study based on a corpus of Canadian teens’ IM use findings yielded similar statistical results (Tagliamonte and Denis 2008). This latter study also examined “the extent to which IM language mirrors everyday language” by comparing the use of discourse features such as personal pronouns, quotatives and intensifiers in written text, IM and spoken youth language. According to the authors, the analysis revealed that IM language is characterized by a robust mix of features from both informal spoken registers and more formal written registers – in essence it is a hybrid register’ (Tagliamonte and Denis 2008: 5).

In a qualitative study of CMC, Lam (2004) investigated youths’ use of language in a Chinese–English bilingual chat room. According to Lam, youth in this chat room code-switched between English and Chinese in order to express modality, convey humour and emotion, and mark social roles and relationships in their conversations. Much like the previous study, Lam’s findings suggest that IM language is a hybrid register in several respects. First, the IM language of youth in the bilingual chat room incorporated features of spoken Chinese, as well as written English text. Moreover, Lam points out that use of Chinese discourse markers “could be a simple yet pervasive way in which a Cantonese conversational tone is introduced into an otherwise English dialogue” (2004: 54). Thus representing the global forms of English being used by adolescents in online spaces that attract interlocutors from around the world. She concludes that use of such hybrid forms serves to help create a ‘collective ethnic identity’ (2004: 45) for Chinese immigrants.
Finally, though research in this area has just begun, initial studies indicate that messaging on youth-oriented social network sites, such as MySpace and Facebook, features the same kind of informal elements found in instant messaging and chat rooms, such as written description of non-linguistic cues (e.g. ‘hug,’ ‘wink’), use of non-linguistic symbols to display emotions (e.g. 😊), shortened forms (e.g. bday, pic, luv) and extensive code-switching between multiple dialects and languages (Chou 2008).

Short-messaging service

Another electronic form of communication that is rapidly growing in popularity among youth and adults alike is short-messaging service (SMS), otherwise known as texting. Text messages are asynchronous and are constrained by a protocol that allows a maximum of 160 characters per message. This constraint on the number of characters has prompted widespread use of abbreviated forms of language often referred to as ‘textese’. Much like the language associated with IM and chat, textese consists of abbreviation, logographic spelling and rebus forms of writing. In recent years, there have been linguistic analyses of texting in several languages, including Swedish (Hård af Segerstad 2002), Norwegian (Ling 2005) and German (Döring 2002, cited in Ling and Baron 2007). Save for one study on British English (Thurlow 2003), there have been relatively few studies of the language forms associated with English-based texting. This can in part be attributed to the ubiquity of mobile phones and thus texting in Europe and Asia, versus the high percentage of personal computers and thus IM and chat in the US (see Ling and Baron 2007).

As the one exception, Thurlow (2003) examined the linguistic forms and communicative functions of youth’s text message use. Findings revealed that the primary linguistic changes that youth made (abbreviations, contractions, acronyms, misspellings and non-conventional spellings) were serving the sociolinguistic “maxims” of (a) brevity and speed, (b) paralinguistic restitution and (e) phonological approximation (Thurlow 2003: section 4). According to the authors, these changes were linguistically ‘unremarkable’ and ‘would not be out of place on a scribbled note left on the fridge door’ (2003: section 4). Thurlow’s discussion highlights a theme that runs through much of the academic research and commentary on the potential linguistic changes associated with new ICTs – that technologies such as email, IM, chat and SMS do not, for the most part, bring about changes in language forms, but rather amplify trends already underway. Studies consistently show that levels of informality and the use of non-standard linguistic forms vary according to context and purpose. As Crystal (2008) points out in the following passage, rebuses and other abbreviated forms of writing have been around for centuries:

Similarly, the use of initial letters for whole words (r for ‘no’, gf for ‘girlfriend’, cmb ‘call me back’) is not at all new. People have been initialising common phrases for ages. IOU is known from 1618. There is no difference, apart from the medium of communication, between a modern kid’s lol (‘laughing out loud’) and an earlier generation’s Swalk (‘sealed with a loving kiss’).

(Crystal 2008: 14th paragraph)

In summary, electronic interaction today features many of the same types of abbreviations and colloquialisms as those that occurred previously when conversational English
was put into writing. However, thanks to the sheer size and volume of the internet, and the amount of time many people spend chatting or texting online, such forms have become more widespread and controversial. Overall, they represent an expansion of the written use of colloquial English vs formal or academic English. As such, they enable many people on the margins of power, including youth and immigrants, to communicate in a form that better expresses their sense of identity and community.

Whose voice?

The principal inventor of the web, Timothy Berners-Lee, intended it to be a read-write medium in which it was as easy to create and publish material as it was to read and browse (Berners-Lee 1999). However, the web that emerged frustrated that vision, as online publishing in the web’s early days necessitated mastery of complex coding processes. The development of specialized web design software partially solved this problem, but it was the development and diffusion of free blogging software and host sites that truly allowed web-based publishing to become a mass phenomenon.

Blogs fall within a range of categories, each with its own antecedents in other media (for an overview, see Miller and Shepherd 2004). As Chesser (2005) notes, the majority appear to fall within two general types. First, there are personal journals, which fall within the pre-internet tradition of diaries and personal letters. They largely describe people’s personal thoughts, feelings and day-to-day experiences, and serve the dual purpose for the writer of keeping friends or family informed and reflecting on one’s identity through writing. A second type of blog falls within the tradition of the newspaper column or pamphlet. It seeks to inform, agitate and persuade, most frequently on political topics.

Herring and colleagues carried out content and genre analysis of several hundred randomly selected blogs in a series of studies published in 2004–6 (Herring, Kouper, Scheidt and Wright 2004; Herring, Scheidt, Bonas and Wright 2004; Herring et al. 2005; Herring and Paolillo 2006; Herring et al. 2006). They found that personal journals constituted 70.4 per cent of their sample (Herring et al. 2005). The next largest group, constituting 12.6 per cent of the sample, was what they called filter blogs, because they often filtered news and information from the broader web (see Blood 2002 for original use of the term). These filter blogs primarily contained observations and were of external, typically public, events and tended to correspond to the informational/agitational purpose described above. A third type of blog was identified that seeks to provide information and observations on a topic, project or product; this category, referred to as k-log (knowledge log), constituted 4.5 per cent of the blogs they examined.

Though Herring and her colleagues did not match blog purpose with blog topic in their analyses, the sample blogs they chose as illustrations for each of the three main purpose categories match exactly with the topical categorization suggested by Stone (2004), with personal journal blogs typified by personal experience topics, filter blogs typified by political topics, and knowledge blogs typified by technology topics.

There is wide variation in blog structure, from single-author blogs with few links to external sites, few if any comments and infrequent updates, to complex multi-author blogs with extensive linking and tagging, constant updates and voluminous commenting. The majority of blogs analysed by Herring’s group fell on the simple side. A total of 90.8 per cent of the randomly selected blogs they analysed were single-authored, and blogs in their sample were updated on an average of every 2.2 days. The typical blog

495
entry contained 0.65 links to other material, and only 43 per cent of blogs allowed comments by others. A total of 9.2 per cent of blog entries contained images (Herring et al. 2005).

However, what is typical in a random sample of blogs is quite different than what is typical in people’s experiences with blogs. That is because the majority of blogs are rarely visited, while a small number of ‘a-list’ blogs dominate the traffic on the blogosphere (Shirky 2003; Herring et al. 2005). Many of these high-traffic blogs feature complex networking features that enable highly innovative forms of communication and advocacy. For example, Daily Kos, a left-of-centre activist political blog, has a main editor and 15 contributing editors who write front-page postings known as stories; hundreds of people who write additional postings linked from the front page known as diaries; thousands of people that write threaded comments on stories and diaries; extensive linking to other blogs and websites from within comments, stories, diaries and user signature lines; tagging of all diaries and stories to create a folksonomy (i.e. user-generated taxonomy) of blog topics; a search mechanism to find stories, diaries or comments by tag, content or author; an elaborate user recommendation system so that the most highly recommended diaries rise to the top of the list while the most negatively rated comments disappear; a hierarchical system of participants so that those who receive the most positive comments achieve greater privileges to negatively rate others; a main blogroll linking to other link-minded blogs on the front page and distinct blogrolls on other pages created by users; and a collaboratively edited political encyclopedia (Kos Media 2009). Launched by a Salvadoran immigrant in 1982, Daily Kos now has more than 175,000 registered users (Kos Media 2005), receives nearly a million daily visits (The Truth Laid Bear 2009), and has established itself as a major force in US politics (Chait 2007).

Examining the overall blogosphere, Herring and her colleagues suggest that blogs fill an intermediary role within online genres, about midway between standard HTML documents, such as personal home pages, and asynchronous computer-mediated communication (CMC), such as newsgroups, bulletin boards or email discussion lists (see Figure 28.1; Herring et al. 2005). They are more frequently updated and include more exchange among people and a higher percentage of text (as opposed to multimedia) than standard web pages. But the exchanges on them tend to be more asymmetric (i.e. dominated by main authors) and less frequently updated (with sites such as Daily Kos an exception) than CMC sites such as newsgroups.

Herring and her colleagues have also begun to examine whether gender differences exist in the language used on weblogs, as they do in other written texts. Their initial research suggests that within particular blog genres, little differences exist between males’ and females’ use of language, with both men and women using more formal

![Figure 28.1](image-url) The overall blogosphere.
typically 'male' language on filter blogs and more typically 'female' language on personal journal blogs (Herring and Paolillo 2006). However, since filter blogs are mostly written by men and personal journal blogs are mostly written by women (Herring, Kouper, Scheidt and Wright 2004; Herring and Paolillo 2006), the overall use of language in the blogosphere is still gendered.

Scholars have just begun to explore authorship and social participation in the blogosphere (for a discussion of some research questions, see Lankshear and Knobel 2006). A number of initial reports have examined the motivations and personal experiences of bloggers, either from a third-person (e.g. Nardi et al. 2004) or first-person perspective (e.g. Krause 2004; Davies and Merchant 2007).

C hesher (2005) analysed authorship on blogs, comparing the conventions of authorship in the blogosphere to those in other electronic or print genres. Authorship in blogs tends to be strongly identified to a real or pseudonymous person through a username or display name for each blog and blog entry, or through an about or profile section that gives more information about the writer. In contrast, older web documents, such as standard web pages, often lack this information. The visual consistency of a blog, compared to a typical HTML web page, also highlights personal ownership and authorship, and the reverse chronological order and specific time stamp on postings create a temporal link between author and reader. Blogs that are most successful, whether in reaching out to a few readers or hundreds of thousands, tend to have a strong authorial voice. In most cases, this personal voice is more easily achieved in blogs than in print journalism, such as newspapers, since blogging encourages an informal, idiosyncratic style and content. In addition, the sheer ease of publishing a blog, as compared, for example, to either setting up and maintaining a frequently revised standard website or becoming a writer for a print newspaper or magazine, makes authorship accessible to a greatly expanded number of people. Chesher concludes that the 'death of the author', which was originally predicted by post-structuralists (Barthes 1977), and which was supposedly going to be hastened by the decentralised and collaborative nature of hypertext (Poster 1990; Landow 1992), is greatly exaggerated. As he states, 'the author is alive and well, and has a blog' (2005: first paragraph).

Beyond giving tens of millions of people new opportunities for authorship, the blogosphere also offers a political voice to those on the margins of power. This is due in part to its structural features discussed earlier. By occupying an intermediary format between the highly interactive form of computer-mediated communication and the more permanent forms of traditional online publishing, blogs can simultaneously replace both institutions pointed to by political theorist Tocqueville as vital for citizen participation: the meeting hall and the newspaper (Tocqueville 1937; Warschauer 2003). Thus, in authoritarian countries such as Iran, blogging has emerged as an important, if risky, form of citizen advocacy to challenge both the censored media and the restricted space for traditional organizing (see, e.g., Hendelman-Baavur 2007). In the United States, the grassroots left, which was relatively dormant from the 1970s to 1990s, has found the blogosphere to be a particularly potent organizing tool, using it more successfully than the right to mobilize funds and support for its favoured candidates and causes, and thus counterbalancing the right's dominance over talk radio (Chait 2007). In 2008, online mobilizing and fundraising campaigns played a critical role in the election of the first African-American president, helping Obama first overcome a heavily favored Democratic competitor for the nomination, and then defeat a popular Republican war hero (Lister 2008).
Of course, blogging is not a silver bullet for achieving social change. The success of the US Democratic Party in the 2006 and 2008 elections was due to many factors beyond successful online mobilization. Also, authoritarian regimes have the power to censor or block blogs and arrest bloggers (see, e.g., Gray 2008), or publish their own misinformation on blogs. The blogosphere is a complex and competitive social and political environment, with those seeking to spark, resist or co-opt social reform movements all fighting for influence, together with millions of others without political agendas (for mappings of the Iranian and US blogospheres, see Kelly and Etling 2008; Linkfluence 2008).

Whose knowledge?

If blogs create new opportunities for expressing voice, then wikis create new opportunities for sharing and producing knowledge. Wikis are simply websites that any visitor can contribute to or edit (Richardson 2006). Though there is no authoritative listing or account of the number of wikis, they are surely far fewer than blogs. They have been principally established so that groups of people can contribute their knowledge and writing skills to collaboratively create informational documents. For example, some of the largest wikis (based on statistics from S23 2007) include Richdex (an open source directory on a wide range of topics), WowWiki (an information source about the World of Warcraft online game), and wikiHow (a collaborative how-to manual).

By far the largest wiki, and one of the ten most visited websites in the world (see Alexa 2008), is Wikipedia. Its English version alone includes more than 2,600,000 articles totalling some 1 billion words, more than 25 times as many as are in the next largest English language encyclopedia, the Encyclopædia Britannica (Wikipedia 2009e). Most remarkably, there have been some 236 million edits to Wikipedia since its inception in 2001 made by 5.77 million contributors (Wilkinson and Huberman 2007).

Most of the textual analysis of wikis has been directed at Wikipedia, with much of the research focus on its accuracy. Its breadth of content, ease of access, free cost and links to external material make Wikipedia potentially highly useful to a vast online audience. The foremost question for casual users and researchers alike has been whether a collaborative process that welcomes the participation of novices as well as experts can produce satisfactorily accurate results. In a widely cited study on this topic, Nature (Giles 2005) had a panel of experts compare content from 42 entries of approximately the same length on scientific topics from Wikipedia and the Encyclopædia Britannica. The experts identified 162 errors in the Wikipedia content (four of which were serious) and 123 in the Encyclopædia Britannica content (four of which were serious), thus suggesting that neither encyclopedia is infallible, and that the six-year-old open-source Wikipedia is only slightly less accurate than the 238-year-old professionally edited Britannica. In a related study, Chesney (2006) had 258 research staff judge the credibility of two Wikipedia articles, one in their area of expertise and one chosen randomly. In general, the researchers found the articles to be credible, and even more so in their own area of expertise.

Though anyone can accidentally or purposely introduce errors into Wikipedia, they are usually found and corrected quickly by the site’s large number of volunteer editors. In one experiment, a professor of communication intentionally introduced 13 errors, some obvious and some subtle, in a range of Wikipedia articles. He checked back on the articles three hours later and all 13 had been corrected (Reed 2006).
Focusing on linguistic features rather than accuracy, Bell (2007) compared articles in Wikipedia and the online version of Encyclopaedia Britannica on three measures: readability, syntax (specifically nominal vs verbal nature) and use of fact statements vs value statements. He found that the two encyclopaedias were roughly comparable on all three measures. A similar study by Elia, focusing on lexical density, use of formal nouns and impersonal pronouns, and average word length, concurred that the language in Wikipedia 'shows a formal and standardized style similar to that found in Britannica' (2007: 18), even though its articles were twice as long on average and had far more hypertextual links.

If blogs served to suggest that the author is well and alive, wikis fulfill the prophecy of authorship facing away. In essence, the distance between the author and audience is eliminated when the audience can directly edit the author's work. In many Wikipedia articles, it is difficult to discern a principal author. For example, a review of the history (posted with each article) for the Wikipedia entry on the innocuous topic of asparagus indicates it has been edited hundreds of times by dozens of people over the last five years.

Wikipedia provides a fruitful source for researching the nature of collaborative authorship and editing. A study by Wilkinson and Huberman (2007) analyzed the impact of cooperation among editors on Wikipedia on article quality. Specifically, when controlling for age and visibility of articles, they found that both the numbers of edits and the numbers of editors were strongly correlated with article quality. On the one hand, this seems intuitive, in that more attention should result in higher quality. However, the authors point out that in other areas, such as software development, industrial design and cooperative problem-solving, large collaborative efforts are known to produce ambiguous results.

In a study on the Hebrew version of Wikipedia, Ravid (2007, cited in Warschauer and Grimes 2007) analyzed how this collaboration worked, and how it differed between featured articles (which are generally recognized as being higher quality) and non-featured articles. Using a variety of social network analyses, he compared structures of dominance and heterogeneity among contributors in 432 featured articles and 410 non-featured articles. In general, he found a greater degree of inequality of participation in the featured articles. In other words both featured and non-featured articles had large numbers of contributors, but a smaller circle of presumably more expert authors contributed a larger portion of the articles selected for their high quality.

One controversy surrounding Wikipedia has focused on it as a source for student research. The founder of Wikipedia, Jim Wales, provides the most commonsense answer to this, suggesting that although Wikipedia can help provide an overview of issues and a starting point for identifying primary sources, students are better off using primary sources as definitive sources in their research. 'For God's sake, you're in college; don't cite the encyclopedia,' Wales told one college student (Young 2006: second paragraph.)

A more interesting question is how writing for wikis can affect the learning process. The potential of wikis for teaching and learning is hinted at by Ward Cunningham, inventor of the wiki, who commented that 'The blogosphere is a community that might produce a work, whereas a wiki is a work that might produce a community' (Warschauer and Grimes 2007: 12). Cunningham's statement illuminates a central contradiction of CMC since its inception: it has served as a powerful medium for exploring identity, expressing one's voice, airing diverse views and developing community, yet has proven a very unsuitable medium for accomplishing many kinds of collaborative work through the inherent difficulty of arriving at decisions in groups dispersed by
space and time (see meta-analysis comparing face-to-face and computer-mediated decision-making by Baltes et al. 2002).

Wikis turn traditional CMC activity around in several respects. Whereas email and chat, the most traditional CMC genres, facilitate informal, author-centric, personal exchange, writing on a wiki facilitates more formal, topic-centric, depersonalized exchange. Each edit makes a concrete contribution to a collaborative written product, with authorships relegated to a separate page that only the most serious of readers are likely to notice. Wikis are thus an especially powerful digital tool for knowledge development, and thus for education (for examples, see Mader 2007; Wikipedia 2009b).

Finally, the existence of a 'simple English Wikipedia' – with more basic vocabulary and grammatical structures, fewer idioms and jargon, and shorter articles – further democratizes this knowledge tool, as it makes the process of accessing and disseminating information more accessible to learners of English, people with learning difficulties, and children (Simple English Wikipedia 2009).

Wikis, and Wikipedia, are just one way that control of the knowledge production process is being challenged. For example, in the area of scholarly and scientific research, online research databases and journals are also threatening academic publishers’ control of knowledge dissemination (Willinsky 2006).

**Conclusion**

When the internet first emerged, there were simplistic notions of a single online English, which contrasted with both spoken and written English. In fact, there are many varieties and genres of online English, just as such diversity exists in the spoken and written realms. However, there are some commonalities across this diversity, and one important common trend involves the challenge to traditional gatekeepers of English language use, as exemplified by Wikipedia challenging the *Encyclopaedia Britannica*, the blogosphere challenging the mainstream media, or tens of millions of youth challenging notions of correct English.

None of these challenges are, in and of themselves, revolutionary. Non-standard varieties and usages of English have existed for centuries, and new media have continually emerged to either complement or replace the old. The significance of these changes in language and communication will in the future, as in the past, depend on the broader social circumstances in which they unfold. Kaplan’s comments on the matter, first made in the early days of the World Wide Web and published in a then-new online magazine, still ring true today:

> The proclivities of electronic texts - at least to the extent that we can determine what they are - manifest themselves only as fully as human beings and their institutions allow, that they are in fact sites of struggle among competing interests and ideological forces.

*(Kaplan 1995: 28)*

Youth, immigrants and others may seize on new hybrid forms of online Englishes to express their identity, but they will require mastery of sanctioned varieties of English for social or economic advancement. Bloggers can challenge state authority, and can be thrown in jail for doing so. And the viability of new sources of online knowledge,
whether in Wikipedia or non-commercial journals, will be called into question by traditional gatekeepers.

Finally, we have only scratched the surface in this chapter of the ways that Englishes are evolving online. Multiplayer games, podcasting and video publishing will all have their own impact on the evolution and use of English. And, in these audiovisual domains, as in the textual domains discussed in this chapter, the proclivities of new Englishes will manifest themselves as human beings and their institutions allow. However, that discussion will have to await another chapter, perhaps to be published on YouTube.

Acknowledgments

This paper draws in part on the first author’s previous discussion of blogs and wikis in ‘Audience, authorship, and artifact: the emergent semiotics of Web 2.0’ (Warschauer and Grimes 2007).

Suggestions for further reading


References


