18. Multilingualism on the Internet

Brenda Danet and Susan C. Herring

1. Introduction[1]

In recent years, the Internet has become a truly global communication network. According to a recent compilation (Almanac 2005), about one billion people, one-sixth of the world's population, are now online. Not all nations are equally represented: Notably absent are the countries on the African continent, and 15 countries account for about 70% of the total (Table 1). The United States has the largest single proportion online of any country, or 20% of the total. This reflects not only its relatively large population size and advanced technological infrastructure, but also the fact that the technology that makes the Internet possible was created in the 1960s in the United States (O'Neill 1995; Hafner and Lyon 1996; Cringely 1998).

Table 1. Top 15 countries on the Internet, 2004. Source: Computer Industry Almanac, September, 2004; URL http://www.c-i-a.com/pr0904.htm, retrieved August 12, 2005. Reproduced with permission.

Top 15 countries in internet usage		
Year-end 2004	Internet users (#K)	Share %
1. U.S.	185,550	19.86
2. China	99,800	10.68
3. Japan	78,050	8.35
4. Germany	41,880	4.48
5. India	36,970	3.96
6. UK	33,110	3.54
7. South Korea	31,670	3.39
8. Italy	25,530	2.73
9. France	25,470	2.73
10. Brazil	22,320	2.39
11. Russia	21,230	2.27
12. Canada	20,450	2.19
13. Mexico	13,880	1.49
14. Spain	13,440	1.44
15. Australia	13,010	1.39
Top 15 Countries	662,360	70.88
Worldwide Total	934,480	100

In the last decade, many people have expressed concern about the global dominance of English, and about the Internet as a new arena for its spread (Nunberg 2000; Pimienta 2002; Tsuda 2000; Mair 2002; Dor 2004). A 2002 survey found that over 56% of all Web pages were in English.[2] The English language is especially prominent in the commercial sphere: The Organization for Economic Co-operation and Development (OECD) reported that in July 2000 more than 94% of links to pages on secure servers were in English. "The only other languages to account for more than 1% of detected links to secure servers were German [...] and French [...], although Spanish and Japanese came close" (Organization for Economic and Social Development 2001). Some view the spread of English as a "natural", largely benign or even beneficial extension of globalization generally (Fishman 1998; Fishman, Conrad, and Rubal-Lopez 1996; Crystal 2001, 2003). Other authors take a dimmer view, writing of "linguistic imperialism" (Phillipson 1992, 2000; Phillipson and Skutnabb-Kangas 2001; see also Kirkpatrick, this vol.). Figure 1 contrasts the two views:

	Exploitation model	Grassroots model
political value of English	imperialist language	post-imperial language
chief cause for post World War II spread	organized/centralized language planning following Anglo-American master plan	demand-driven; decentralized rational choices by individuals and groups
English is the language of ...	Anglo-American capitalist interests	modernization and globalization
English is ...	a language that conveys an Anglo-Saxon/Western world view	an ideologically neutral lingua franca
English ...	transforms recipient societies (usually for the worse)	is transformed by recipient societies (rise of New Englishes)
chief beneficiary of "global English"	British and American capitalist interests	usually some segment of local users

Figure 1. Two models of the influence of English. Source: Mair 2002: 165. Reproduced with permission.

Although intended generally, these two models are arguably relevant to the current debate about English on the Internet.

Alongside concern about the dominance of English, there is evidence that the number of non-English speakers on the Internet is growing rapidly. Already by 2003, roughly two-thirds of all Internet users were not native speakers of

English (CyberAtlas 2003). In another estimate, about 800 million non-English speakers were online by 2005.[3] In only four of the 15 top countries online in 2004 (the US, the UK, Canada, Australia) was English the official or dominant language of users (Table 1). China and Japan together accounted for nearly another fifth of the total. Moreover, growth in the next few years is predicted to accelerate, especially in China and India.[4]

The growth of various language groups online reflects not only developments in technology and infrastructure, but demographic trends. Already in 1995, Chinese was spoken natively by over one billion people, far more than any other language; English was in second place, with under 400 million native speakers. Moreover, extrapolating from UN-based statistics, David Graddol predicts that the proportion of the global population speaking English natively will decline from nearly 9% in the mid-20th century to about 5% by 2050 (Graddol 1997/2000, 1999, 2004). In keeping with these developments, hundreds of millions of people are already participating online today in languages other than English, in some form of non-native English, or in a mixture of languages.

Academic research published in English on the language of computer-mediated communication (CMC) has only recently begun to take account of this complex empirical reality. Most researchers publishing in English venues have generalized about the language of computer-mediated *communication*, when in fact they were describing computer-mediated *English* (e.g., Ferrara, Whittemore, and Brunner 1991).[5] Some exceptions are publications by Naomi Baron (2000) and David Crystal (2001, 2004a), which contextualize English-based CMC within the history of the English language.

In recent years this situation has started to change. Increasingly, researchers have turned their attention to other languages used on the Internet, often – although not always – their native languages. In this chapter we survey this growing literature, treating multilingualism both macro-sociolinguistically and micro-sociolinguistically. We focus on issues of linguistic diversity and the fate of specific languages on the Internet at the macro level and on micro-level patterns of language use by individuals communicating via instant messaging, email, and chat.

Some of the research discussed here was published in a special issue of the *Journal of Computer-Mediated Communication* that we edited in October 2003 on the topic of "The Multilingual Internet: Language, Culture and Communication in Instant Messaging, Email and Chat" (Danet and Herring 2003a). Other articles cited appear in an expanded book on the same theme (Danet and Herring 2007).[6] The remainder of the studies were published in diverse venues, many of which are considered together for the first time here.[7] We discuss this emergent literature under five recurrent themes: writing systems, linguistic features of CMC, gender and culture, language choice, and language revitalization efforts.

2. Writing systems and online communication

2.1. ASCII encoding and its unintended consequences

Because early planners of the Internet were generally North American, and only had in mind how to facilitate communication in English, they did not anticipate problems that might arise when speakers of other languages tried to communicate online. The text-transmission protocol on the Internet is based on the ASCII character set (Figure 2). ASCII (pronounced AS-kee) is an acronym for "American Standard Code for Information Interchange". Established in the 1960s, it contains 128 seven-bit codes (unique combinations of 1's and 0's), 95 of which are available for use as graphical characters. This character set is based on the roman alphabet and the sounds of the English language. "Plain text", as in email and chat, usually refers to a format containing only basic ASCII characters, whether written in English or some other language.

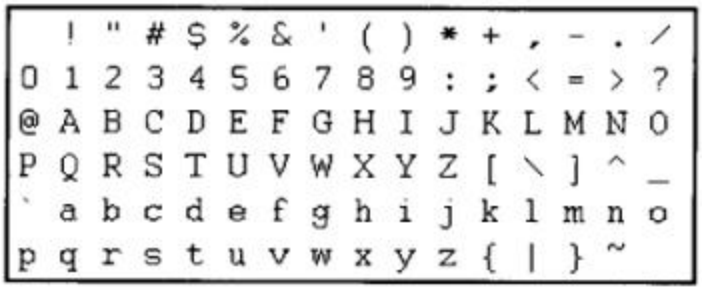


Figure 2. The ASCII character set. Source: http://www.cs.tut.fi/~jkorpela/chars.html; retrieved August 15, 2005.

There can be little doubt that the ASCII character set has privileged English on the Internet. Whether it concerns HTML (the markup language for Web pages), domain names on the Web (URLs), email addresses, or the content of instant messages, email, discussion list postings, or chat, speakers of languages other than English have faced varying degrees of difficulty.[8] Speakers of languages using the Latin or roman alphabet but with only a few characters missing, such as Maltese[9] or the Scandinavian languages, suffer the least disadvantage, though one that may produce embarrassing results. For example, ASCII does not include the last three letters of the Swedish alphabet, *å*, *ä*, and *ö*. In the past,

> Instead of using *å*, *ä*, and *ö*, [Swedes] managed with substitutions such as }, { and |, or *aa*, *ae* and *oe* in electronic communication. Or just "a" for *å* and *ä*, as well as for the letter *a*, and "*o*" both for *o* and *ö*.
> (Pargman 1998: 87)

The URL of a Swedish town called Hörby is http://www.horby.se. Swedes must live with the fact that without the two dots over the "o", the name of this town means "fornication village".[10] Fortunately, today, email sent within Sweden can use the full Swedish alphabet (Pargman 1998: 87).[11] Another language similarly challenged is Hawaiian, which is written in roman characters, but with additional use of macrons.[12] Warschauer and Donaghy (1997: 353) note that "Incorporation of diacritical marks is crucial, since they define meaning in Hawaiian; for example, *pau* means finished, *paÿu* means soot, paÿü means moist, and päÿü means skirt."

Speakers of languages with non-Latin writing systems, such as Greek, Russian, Arabic and Hebrew, and the East Asian languages (Chinese, Korean, Japanese) are more disadvantaged, being dependent on the development of special character sets to make word-processing and online communication possible.[13] While great progress has been made in these areas, truly multilingual global communication on the Internet is still plagued by many technical problems. As recently as the late 1990s, it was possible to claim that "English remains the only language that can be used without distortion on virtually every computer in the world" (Fishman 1998: 34).[14] Even in 2005, although speakers/writers of many languages could configure their computers to accommodate their own multiple or non-English language needs, they could not assume that others would be able to receive and read messages and longer texts including characters other than basic ASCII.

2.2. Modes of adaptation to the ASCII environment

Problems engendered by the dominance of the ASCII character set online, such as those just discussed, may lead some to speak of "typographic imperialism", as an extension of linguistic imperialism into the domain of the Internet (cf. Pargman and Palme 2004). In this section we ask: How have people communicating online in languages with different sounds and different writing systems adapted to the constraints of ASCII environments? What problems have they encountered, and what are the social, political, and economic consequences if they have or have not adapted? What progress has been made in solving these problems, and what remains to be done?

English-based research reveals that synchronous chat and even email and discussion list postings tend to have partially speech-like features (Yates 1996; Herring 1996b, 2001; Baron 2000; Crystal 2001; Danet 2001: chap. 2). What happens when people using formerly spoken-only varieties of languages participate in typed chat online? This question is especially pertinent with regard to Arab countries, which are characterized by *diglossia* (Ferguson 1972; Hudson 2001): High-prestige, written, literary, classical Arabic co-exists with a low-prestige, local spoken variety that is ordinarily not written – at least not until the advent of the Internet.

In a study of instant messaging among young female Gulf Arabic speakers in Dubai, David Palfreyman and Muhamed Al Khalil (2003, 2007) found a mixture of Arabic script, English and romanized Arabic – spoken Arabic written out in the roman alphabet (Figure 3). Whereas speaker D in Fig. 3 has typed her contributions in right-to-left Arabic script, speaker F "script-switches", writing in Arabic rendered left-to-right in the roman alphabet. Online romanization has also been documented for email in colloquial Egyptian Arabic (Warschauer, El Said, and Zohry 2002, 2007), and for chat on IRC in colloquial Moroccan Arabic (Berjaoui 2001). Palestinians write online messages using either roman or Hebrew letters, rather than Arabic ones (Myhill and Garra 2005).

D: وبركاته الله ورحمه عليكم السلام	*D: Hello there.*
D: شحنّج؟ حمده.. مرحب	*D: Hi Hamda, how are you doing?*
F: w 3laikom essalaaam asoomah ^__^	*F: Hi there Asooma ^__^*
F: b'7air allah eysallemch .. sh7aalech enty??	*F: Fine. God bless you. How about you?*
[pause]	[pause]
D: el7emdellah b'7eer w ne3meh	*D: Fine, great thanks.*
D: sorry kent adawwer scripts 7ag project eljava script w rasi dayer fee elcodes	*D: Sorry, I was looking for scripts for the java script project and my head is swarming with code.*
F: lol	*F: lol*

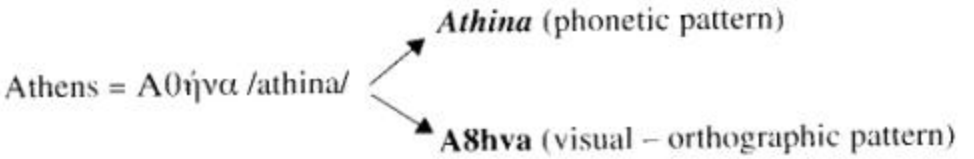
Figure 3. Opening of an IM (instant messaging) exchange in Dubai. Source: Palfreyman and Al Khalil 2003.

Palfreyman and Al Khalil also report use of numerals to represent sounds of Arabic that cannot otherwise be represented in the roman alphabet (third, fourth, fifth, and sixth contributions in Figure 3). These numerals are codified representations of sounds; for instance, the numeral 7 is used to represent /x/, as in <wa7ed> 'one'. The same phenomenon has been documented for colloquial Egyptian Arabic (Warschauer, El Said, and Zohry 2002, 2007) and colloquial Palestinian Arabic (Myhill and Garra 2005). A few Moroccan examples are discernible in IRC transcripts collected by Berjaoui (2001).[15]

Romanization is well documented for Greeklish – Greek written in roman letters in online communication. In the Greek case, Theodora Tseliga (2007)

found evidence of substituting a roman letter for a missing Greek one that resembles it in *graphic shape*. For instance, a person wishing to write the Greek name for the city of Athens in a romanized context has two options:

(1)



The second choice substitutes the numeral *8* for θ because the two are similar graphically (Tseliga 2007). Publications by Alexandra Georgakopoulou (1997 2004) and Jannis Androutsopoulos (1999a, 1999b, 2004) on aspects of email and chat in Greeklish also provide many examples of romanized Greek.[16] Androutsopoulos (1999a, 1999b, 2000) reports considerable inconsistency in Greeklish email because transliteration norms have not yet emerged. A study of code-switching on IRC by Androutsopoulos and Hinnenkamp (2001) contains material on romanized Greek alternating with German in the IRC channel *#hellas*, and alternation between German and the less typographically-challenged Turkish in *#turks*. The researchers also found instances of orthographic switching, i.e., portions of German written according to Turkish spelling rules. Androutsopoulos (2007) also documents romanized Persian, called "Fenglish", as well as Greeklish on German-based diasporic web forums.

There is little evidence (thus far, at least) that use of romanized Arabic is considered particularly controversial by Arabic speakers. Quite the opposite is the case for Greeklish, the use of which is hotly contested. Dimitris Koutsogiannis and Bessie Mitsikopoulou (2003, 2007) investigated attitudes towards Greeklish in the Greek press. Three main trends were identified: a retrospective trend that views Greeklish as a serious threat to the Greek language; a prospective trend arguing that Greeklish is a transitory phenomenon that will disappear as technology advances; and a resistive trend pointing to the negative effects of globalization.

2.3. Beyond ad hoc improvisation: The search for solutions

Over time, developers created partial solutions to the limitations imposed by the ASCII character set, expanding character sets to employ eight-bit profiles that facilitated use of specific languages and/or groups of languages online. For instance, the eight-bit character set known technically as ISO Latin 1 (alias ISO 8859–1), and more informally as the extended ASCII character set, added enough characters, including letters with diacritics, to accommodate the needs of many European languages, including Swedish (Figure 4).[17]

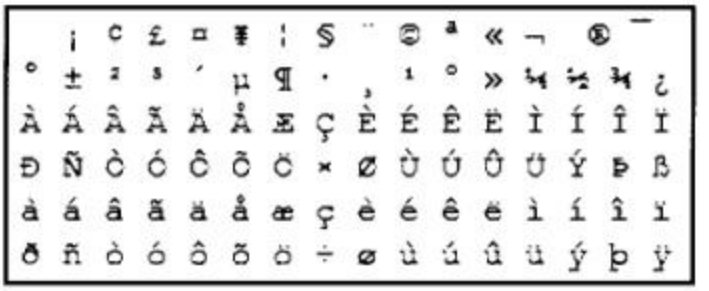


Figure 4. The ISO Latin 1 character set (alias ISO 8859–1). Source: http://www.cs.tut.fi/~jkorpela/chars.html#latin1, retrieved August 15, 2005.

Viewed globally, however, these solutions were unsatisfactory. Eventually, in a major step forward, the Unicode Standard was established (Everson 2002; Consortium 2003; Anderson 2004; Paolillo 2007).[18] Unicode is

> [...] the universal character encoding, maintained by the Unicode Consortium (http://www.unicode.org/). This encoding standard provides the basis for processing, storage and interchange of text data in any language in all modern software and information technology protocols.[19]

The vision behind Unicode is of one encoding for all the scripts in the world (Everson 2002). Whereas the original ASCII character set employs only seven bits per character, in Unicode, each character has a unique 16-bit profile.[20] Developments in the creation of Unicode are now greatly expanding the possibilities for multilingual word processing and communication online, in a broad range of languages and language families, making improvisational forms of adaptation less necessary.

Expansion of digital encoding possibilities is aimed both at archaic languages and at living ones, although living languages are of primary interest here. The Script Encoding Initiative, based at the University of California, Berkeley, funds proposals for scripts currently missing in Unicode.[21] The word *script* in the name of the organization reveals that only languages for which a script *already exists* can be candidates for digitization. In its latest version (4.1.0), Unicode can accommodate a million characters; thus far, over 97,000 characters have been encoded.[22] As of early 2005, Unicode accommodated over 50 scripts, just five of which now accommodate hundreds of the world's languages (see the Appendix). The Latin or roman alphabet, in particular, is used for many languages, sometimes with the addition of diacritics.

Although limitations imposed by the ASCII character set are slowly disappearing, there is a long way to go. More than 80 scripts remain uncoded in Unicode (Anderson 2005: 27). Fishman (1998: 32) estimates that only a small minority of the roughly 1,200 standard languages in the world, i.e. languages that have codified grammars and writing systems, are currently usable online.

3. Features of CMC in languages other than English

Studies of English CMC have often noted its special linguistic and typographic features. The best-known of these are the emoticon or "smiley face", along with abbreviations such as *LOL* ("laughing out loud"), *brb* ("be right back"), and rebus writing (e.g., *c u* for "see you"; Danet 2001; Herring 2001), and a tendency towards "speech-like" informality (Werry 1996). Crystal (2001) uses the term "netspeak" to describe what he considers to be a new, technologically-determined variety of English. Here we are interested in the effects of offline usage and culture on the features of CMC in languages other than English.

We illustrate these effects with reference to Japanese, an interesting case because it is a language with an unusually complex writing system. As Yukiko Nishimura (2003) explains, four scripts are used in standard Japanese orthography: 1) *kanji*, ideograms of Chinese origin; 2) *hiragana* and 3) *katakana*, systems for representing syllables; and 4) *romaji*, use of the roman alphabet to transliterate Japanese words and to represent originally foreign terms, such as *CD*, in otherwise Japanese contexts. Hiragana is used for grammatical endings, and to represent Japanese concepts and objects for which *kanji* do not exist, whereas katakana is used for foreign names and the representation of natural sounds, for instance ワンワン *wan wan*, the Japanese equivalent of "bow wow" for a dog barking.

3.1. Linguistic and interactional features of postings on electronic bulletin boards

In a study of electronic bulletin boards (conventionally abbreviated as BBSs), in this case for fans of popular culture idols, Nishimura (2003, 2007) compared linguistic and interactional aspects of Japanese postings with similar studies of English-language CMC. Her analysis revealed many similarities to English, but also distinctive differences. Among the similarities was evidence for multiple punctuation, eccentric spelling, use of all capital letters, written-out laughter, verbal descriptions of actions, and *kaomoji* ('face marks'), vertical analogs of Western-style "smiley" emoticons (e.g., **:-)** for a smile; **:-(** for a frown). The following is an example of a kaomoji from Nishimura's data:

(2) 復活おめでと～♪良かったね (*^▽^*)
hukkatu omedeto '♪yokatta ne (*^▽^*)
'Congratulations on your comeback♪[as if singing]That was good (*^▽^*)'

Nishimura glosses, "This face mark represents the mouth wide open, laughing loudly and cheerfully, with asterisks used to indicate (rosy) cheeks". Another Japanese-specific feature is the use of graphics such as musical notes to indicate mood, as in this example. Nishimura also found that users employed final par-

ticles, a feature usually associated with informal speech (e.g., the tag-question marker *ne*), interacting with other users online as if they were in face-to-face conversation.

3.2. Typography and playful performance

A tendency toward playful performance in online communication has often been documented for CMC in English (Danet 1995, 2001; Werry 1996; Danet, Ruedenberg, and Rosenbaum-Tamari 1998). Widespread play has been observed not only with identity (e.g., Bechar-Israeli 1995; Turkle 1995; Danet 1998), but with language and typography themselves. For example, elaborate play with the visual shape of typographic symbols to simulate smoking marijuana was documented in 1991 in an English-language IRC (Internet Relay Chat) channel; a portion is shown below:

(3) <Thunder>:I :I :\sssss :)
(inhales twice, exhales, smoke dissipates, smiles)
<Kang> hheeeheee
<Thunder>:-Q :I :I :\sssss :)
(puts marihuana reefer in mouth, inhales twice, exhales, smoke dissipates, smiles)
<Thunder>hehehehh
(Danet 2001: 106, gloss added)

Two participants, nicknamed <Thunder> and <Kang>, adapt "smiley" icons and other typographic symbols in real time to represent the cigarette and the process of smoking. In the third line, <Thunder> simulates the entire process, from taking the cigarette into his mouth (the letter "Q"), inhaling, exhaling, with multiple "s"s to simulate the sound and graphic shape of the smoke, to the final smile of pleasure.

Recent research by anthropologists Hirofumi Katsuno and Christine Yano suggests that playful performance via typography is even more elaborate in online Japanese than in English (Katsuno and Yano 2002, 2007). The deployment of kaomoji – 'face marks' – online has important connections with Japanese popular culture, including *manga* (comics; Schodt 1986; Kinsella 2000), a cult of cuteness (Kinsella 1995; McVeigh 2000; Hiorth 2003; Richie 2003; Allison 2004), and a tradition of feminized handwriting (Kataoka 1997, 2003a, 2003b). Whereas earlier kaomoji are typographically compact like Western "smileys" (albeit vertical, rather than horizontal, in orientation), Katsuno and Yano show that Japanese housewives have developed a repertoire of feminized, wider, "cuter" ones (Figure 5).

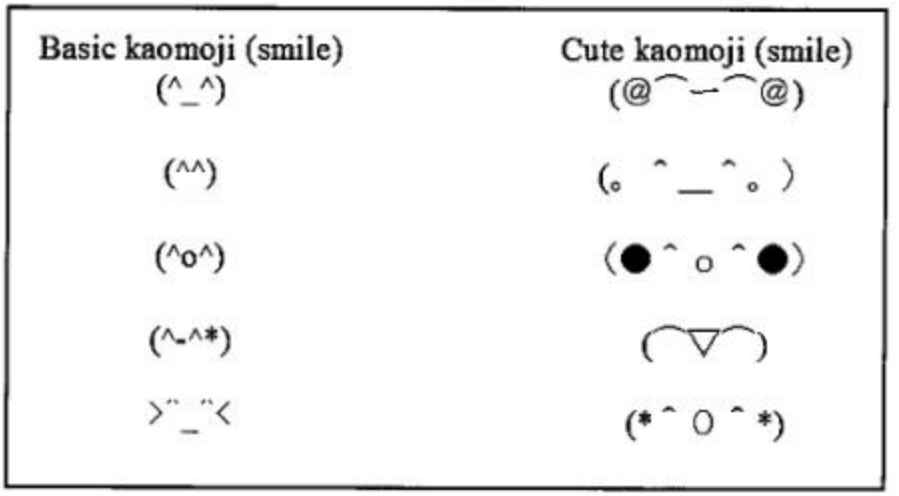


Figure 5. Basic and cute *kaomoji* (face marks). Source: Katsuno and Yano 2007.

The history and use of kaomoji in Japanese and of smileys in English CMC differ. Western-style "smileys" were originally a male phenomenon, created and circulated in the early 1980s by Scott Fahlman and others in the computer science community at Carnegie Mellon University (Fahlman n.d.). By the 1990s, however, more and more females became involved with computers. In the West today, the use of "smileys" is primarily associated with females and young people (Witmer and Katzman 1997). Moreover, they are often considered a tell-tale sign that one is a newcomer or "newbie", and are discouraged in serious online communication. In contrast, even Japanese seniors use kaomoji online (Kanayama 2003). Sugimoto and Levin (2000: 145) found that Japanese-style emoticons were more than four times more common in four Japanese newsgroups than were Western-style "smileys" in four American ones. This form of typographic expressivity is a distinctive form of emergent online culture in Japanese.

4. Gender, language, and culture online

Gender differentiation is an important aspect of culture that is often reflected in language use. English-language CMC research has shown that men and women use different discourse styles online much as they do offline (Herring 1996a, 1996b, 2003; cf. Lakoff 1975, 2004; Tannen 1990). We know of no studies yet to identify systematic "women's language" and "men's language" features in CMC in other languages. However, a growing number of case studies have examined issues associated with gender and Internet use in non-English contexts.

4.1. Politeness

One area of research involves interaction dynamics. Sandi de Oliveira (2003, 2007) analyzed politeness violations on the computer users' discussion list of a university in Portugal. Only Portuguese was used, and the grammar and spelling of the language were standard. However, the messages posted sometimes failed to observe the requirement – of utmost importance in Portuguese culture – to use the appropriate term of address. Thus, for example, a participant entitled by rank to be addressed as "*Professor Doutor*" [+ first name + last name] should not be addressed as "*Senhor*" (Mr.) [+ first name + last name]. Although women participated less often in discussions on the list, messages posted by women were more often treated as transgressions. Oliveira observes that men were quick to chastise transgressions, in contrast to English-based claims that men are less concerned than women with maintaining politeness norms (cf. Herring 1996b). At the same time, the behavior of the Portuguese men on the list asserted their traditional gender roles as interactionally dominant and representative of "authority".

In a study of gender and politeness in email in India, Asha Kaul and Vaibhavi Kulkarni (2005) analyzed 494 work- and task-related emails. Although all the messages were in English, reflecting the widespread use of English as a lingua franca and language of white-collar professionals in India, all were written by employees in Indian workplaces, and reflect the Indian cultural context. Kaul and Kulkarni found that women were more polite than men, as in previous studies of gender and politeness in English CMC. At the same time, the men in the Indian sample used flattery more than women, communicating praise and approval of the recipient's actions – a behavior more commonly associated with women in English CMC (Herring 1996b). Kaul and Kulkarni (2005) suggest that "this could be attributed to the cultural backdrop in which the emails were written where men take on the patronizing role and compliment frequently to motivate the team players/members". Also typical of Indian culture, women's emails were more likely to inquire about the well-being of the recipient and the recipient's family members before moving on to work-related topics.

4.2. Turn-taking

Focusing on the mechanics and power dynamics of interaction, Siriporn Panyametheekul and Susan Herring (2003, 2007) analyzed gender in relation to turn-allocation patterns in a popular Web-based Thai chat room. They found that females made greater use of turn-allocation strategies like those found in face-to-face conversation, and enjoyed greater interactional power in the chat room, chatting with whom they chose and receiving more responses to their messages, than did males. The authors also analyzed flirtatious initiations, find-

ing them to be infrequent and generally lacking in sexually explicit content. They interpreted their findings in relation to the gender demographics of the chat room, the norms of the Web site, and Thai cultural values of politeness and respect – all of which favor female participation.

These three studies demonstrate that gender interacts with culture online in ways that shape language and communication. It has also been suggested that the Internet has the potential to empower women and members of other traditionally subordinate groups (cf. Herring 2003). This potential takes on special significance for women living in traditional patriarchal cultures. According to Katsuno and Yano (2007), expressive use of kaomoji in chat online helps Japanese housewives defuse their real-world frustrations associated with meal preparation, child care, and boring husbands. The Middle East is another region in which gender roles are traditionally segregated. Deborah Wheeler (2001) studied women's use of the Internet in Kuwait, where Internet access is mainly through cybercafés in which – as in other public places in Kuwait – men and women sit in separate sections. Wheeler's evidence suggests that the greater freedom available online to chat with young people of the opposite sex could potentially break down traditional Islamic barriers to mixed-sex interaction.

5. Language choice: National, regional and global aspects

When participants have a choice of languages online, which ones do they choose and why? The factors affecting such choices vary depending on the technological, sociocultural and political context.

5.1. Language choice at the national level: The case of Switzerland

An interesting case in the European context is Switzerland, which has four national languages, German, French, Italian and Romansh, of which the first three are official languages used in government and federal administration. German is the mother tongue of the largest proportion of citizens, with French in second place. With regard to German, we have once again a situation of *diglossia*: A Swiss dialect of German is spoken in informal situations, but High German is used in writing and in formal spoken situations (Schiffman 1997). In addition, English has slowly gained ground as a lingua franca in Switzerland since World War II (Dürmüller 2002; Durham 2003, 2004; Demont-Heinrich 2005).

Mercedes Durham (2003, 2007) studied the languages used on an online mailing list for Swiss medical students during four calendar years, 1999 to 2002. In less than four years, English went from being used a little over 10% of the time to over 80% of the time (Figure 6). Most messages were monolingual and in English.

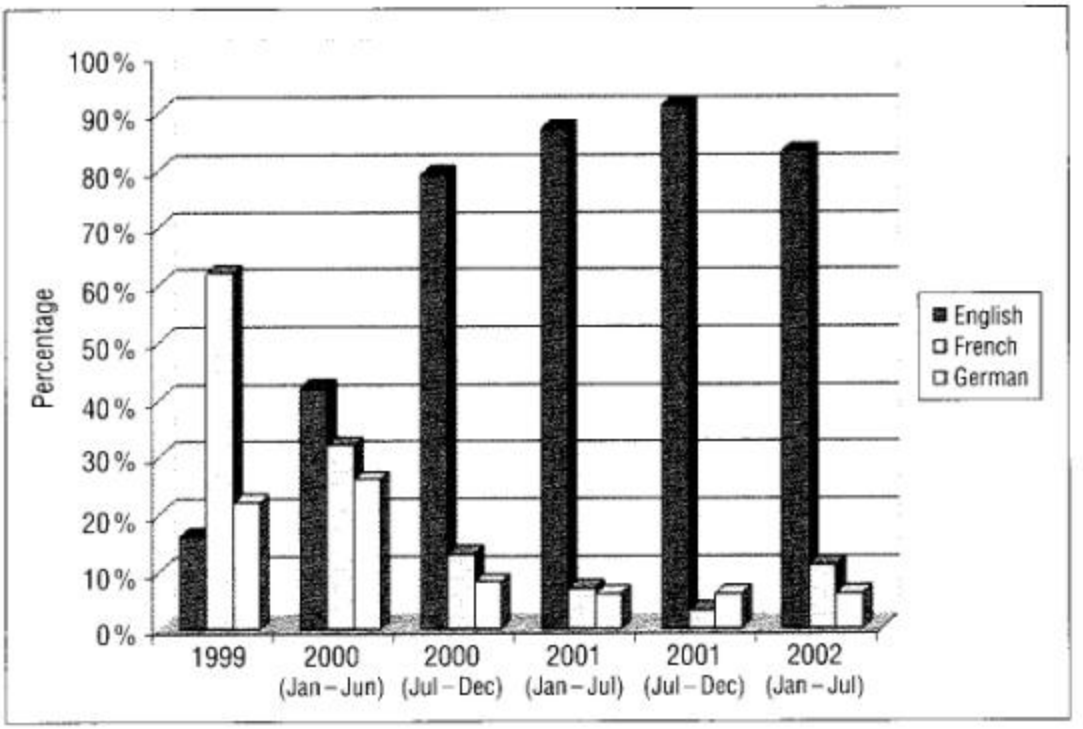


Figure 6. Spread of English on a mailing list for Swiss medical students, 1999–2002: emails by six-month period. Source: Durham 2003.

Medical education is not conducted in English in Switzerland; thus, this cannot be a reason for the shift to English. Durham argues that the main reason is that in Switzerland English is no one's native language, and it ensures the widest possible comprehension among subscribers to the discussion list. Consistent with this explanation, native Italian speakers were more likely to use English than were native speakers of either French or German, because fewer members of the list could be assumed to understand their language.

5.2. Language choice in the less developed world

With regard to language choice in the less developed world, the case of Tanzania is probably representative of many countries in Africa, where Internet penetration has been slow and large areas are not connected to the electricity grid. Safari Mafu, a member of a UNESCO team (Mafu 2004; Sue Wright 2004a), investigated the Tanzanian case. English played a central role in the country's colonial past; during British rule Africans were educated in Swahili, while Europeans and Asians were educated in English. After independence Swahili became the language of instruction but only at the elementary level. Swahili and English are both official languages.

While Internet use has grown in the last five years in Tanzania, only elites – government, universities – usually have access. For ordinary citizens, access is generally via a cybercafé, not a home computer. The government is not concerned that the main language used on the Web is English, although the level of proficiency in English in the general population is low.[23] While the students and professionals interviewed by Mafu (2004) reported some use of Swahili in email, English predominated in their Internet use, reflecting and perpetuating the elite status of users and the functions of English as a language of wider communication. Over one hundred minority languages also spoken in Tanzania are unrepresented online, and are likely to remain so.[24]

In many ways the situation in Africa regarding language choice is similar in the Arab world. English is the main language of email among young professionals in Egypt (Warschauer 2002; Warschauer, El Said, and Zohry 2002, 2007). A complicating factor is the lack of a single standard for communication in Arabic online, and many computers lack operating systems that can handle Arabic (Warschauer 2002). As noted earlier, many people type colloquial Egyptian Arabic in roman letters, both in email and chat. English is more common in formal email communication, reflecting once again its function as a language of wider communication.

5.3. Language choice in the European Union

Language choice issues also arise at the regional level. An interesting case is the European Union (EU), which at the time of this writing has 25 member states and 20 official languages. All legislation is published in all official languages, and the EU maintains a veritable army of translators for written documents and interpreters who perform direct and relay oral interpretation.[25]

Between July 2001 and October 2004, citizens were encouraged to participate in an online discussion about the EU constitution in a Web forum called Futurum,[26] in languages of their choosing. Ruth Wodak and Scott Wright (Wodak and Wright 2007; Scott Wright 2004) found that among the languages actually used in the forum, English dominated by far. Over 90% of all threads or topics introduced in English were conducted only in English. Threads introduced in other languages tended to use a greater diversity of languages, although such threads also tended to be shorter. These results are consistent with general surveys conducted in 1998 and 2000, which found that far more Europeans speak English than any other language.[27]

5.4. Negotiating language choice in global forums

The Futurum case raises a broader issue: How do participants negotiate language choice in global forums where participants are from many countries,

speak different native languages, and where there may be no overt link to a specific national or regional context and no official commitment to a given language – no FAQ (Frequently Asked Questions) file establishing the group's language, or moderator to police language choice? Under what conditions do participants accept English as the lingua franca? When are multiple languages and frequent code-switching tolerated? Under what circumstances does overt discrimination occur against speakers of particular languages? A small number of studies have begun to address such issues.

One of the first studies to investigate language choice online was John Paolillo's (1996) research on *soc.culture.punjab*, a Usenet newsgroup populated mainly by expatriate Punjabis living in Canada, the UK, and the US. In an analysis of a corpus of messages collected in 1994 and 1995 (when the group was still quite new), Paolillo found little use of Punjabi. English was the unmarked language, whereas Punjabi was functionally marginalized, a marked choice used primarily for expressive purposes. He argued that marginalization of Punjabi was fostered by a combination of intergenerational language shift, cultural ambivalence of expatriate Punjabis, the prestige of English in South Asia, and Usenet norms favoring English. Revisiting Paolillo's study a decade later, we can add that having to type Punjabi in roman characters may also have discouraged its use.[28]

Axelsson, Äbelin, and Schroeder (2003, 2007) studied efforts to switch languages in Active Worlds, a graphical chat system,[29] and the responses to such efforts. In this international context, English was the main language used. Non-English speakers, who were generally bilingual, tended to switch to English even in settings where the majority of users were non-English-speaking. English speakers accepted non-English languages more in themed settings (role-playing, games, religion, etc.) than in general, cosmopolitan ones without specific themes. The perceived intention of those attempting to switch languages – to find out if fellow speakers were present, to be playful, or in some cases to be intentionally disruptive – influenced the response.

Luis Fernandez (2001) reports more discouraging findings on the use of minority languages in online forums. The manager of a list discussing the future of Ireland warned those posting in Gaelic (rather than English) that their posts would be removed (Ostler 1999, cited in Fernandez 2001: 24). On Leonenet, a mailing list about current events in Sierra Leone, when some people started posting in Krio, the country's lingua franca, others thought this impolite vis-à-vis non-Sierra Leonean subscribers, or that the practice discriminated against speakers of the languages of other ethnic groups (Wright 1996: 24; cited in Fernandez 2001: 25).

In Fernandez's own research, he found almost no use of Basque in ostensibly Basque forums, even though many users were bilingual in Basque and Spanish or French. Only 8% of messages contained any Basque at all; most

were in Spanish. Spanish messages also tended to be much longer. The dominance of Spanish here resembles the findings of Wodak and Wright for English in Futurum. However, whereas the goal in the EU forum was communication among citizens of many nations, in the Basque forums the focus was specifically Basque and Basque issues.

Even in *Errelea*, the Basque-named mailing list for the Real Sociedad football team, in the Spanish province with the highest percentage of Basque speakers, the Basque language did not fare well.

> [...] there was colorful debate and exchange of ideas in the forum, but all messages were in Spanish. The presence of Basque words, terms, tags, salutations, and slogans was also common [...] but *they were always inserted into Spanish messages*, or used as signatures.
> (Fernandez 2001: 33; italics added)

Fernandez complained to the list about the absence of Basque. All responses were negative, arguing that wider communication in Spanish, which is known to all, is best; that using Basque would be disrespectful to those who do not know the language; and that politics should be kept out of the forum. Fernandez interprets this as an instance of "linguicism" – "ideologies and structures [...] used to legitimate, effectuate and reproduce an unequal division of power and resources [...] between groups which are defined on the basis of language" (Skutnabb-Kangas 1988: 13). More simply put, it is an instance of a minority group adopting the majority's stance toward its language. The uses of Basque described above – instances of code-mixing and code-switching where Spanish was the matrix language – appear to have been an expressive means of showing solidarity with fellow Basques, while simultaneously maintaining communication with others. Fernandez found only three Basque-only forums, and concluded that Basque could flourish only if a Basque-only policy is set beforehand.[30]

A more encouraging case is that of Assyrian/Neo-Aramaic/Syriac. Erica McClure (2001a, 2001b) undertook a comprehensive study of the role of language in the maintenance of the mainly diasporic community of Assyrians, an ancient people whose homeland is in the area of Turkey, Iran, Iraq, and Syria. A large group of Assyrians now lives in the United States, especially in the Chicago area; others live in Australia, Sweden, Lebanon, Iraq, and Canada (Gabrial 1998). Whereas in the homeland their Christian religion distinguished them from others, in the diaspora, the language, variously known as Assyrian, neo-Aramaic, or Syriac, is crucial. It is a Semitic language with a distinctive right-to-left script.[31]

McClure collected extensive samples from Usenet newsgroups including *(soc.culture.assyrian)*, chatrooms, and online publications, with special attention to the forms and functions of code-switching in these media genres (McClure 2001a). In the 1990s Assyrian was mostly transliterated into the roman alphabet

for online purposes, because of font difficulties. Even transliteration was problematic since it had not been standardized. Nevertheless, McClure (2001a: 186) reports a good deal of code-switching to Assyrian in mainly English-based chatrooms and newsgroup postings. Greetings and closings were frequently written in romanized Assyrian, to express solidarity with others.[32] She concludes: "Assyrians have found in the Internet a strong tool in the fight for the maintenance of their language" (McClure 2001b: 74). By the time of publication of her article (2001b), it was possible to post to the Assyrian Forum in an Assyrian font, as well as in English.[33]

6. The Internet and language revitalization efforts

As the brief review of the work by McClure on Assyrian suggests, an issue of ongoing concern is whether the Internet is accelerating the global spread of English and other "big" languages at the expense of local, indigenous, or minority languages (Phillipson 1992; Phillipson and Skutnabb-Kangas 2001; Sue Wright 2004b; Kirkpatrick, this vol.). Some scholars see the potential for this to result in eventual language death and an overall reduction in global linguistic diversity (Crystal 2000; Herring 2002; Paolillo 2007). Others believe that English can continue to spread and to serve as a useful lingua franca, while at the same time, increasing numbers of other languages and language groups can establish a viable presence on the Internet (Fishman 1998; Nunberg 2000; Crystal 2001, 2003; Dor 2004).[34]

There is serious cause for worry about the fate of "endangered" or minority languages, hundreds and thousands of which are rapidly becoming extinct, as the last speakers die off and institutional supports disappear (Krauss 1992, 1998; Grenoble and Whaley 1998, 2001; Nettle and Romaine 2000; Crystal 2000; Dalby 2002). Experts estimate that as many as half the approximately 6,900 languages of the world[35] will disappear in the 21st century (Nettle and Romaine 2000: 7; Crystal 2000: chap. 1). Both professional linguists and lay activists have been involved in efforts to save some of these languages.

The Internet offers a host of new tools to support these endeavors – databases, CD-ROMs, websites, and discussion forums that can be used for language learning, language advocacy, and other forms of communication (Buszard-Welcher 2001; Eisenlohr 2004). Laura Buszard-Welcher (2001) surveyed 50 websites for endangered native American languages, noting the extent to which they supplied community information, materials about writing and fonts, vocabulary and phrases, texts, reference materials such as dictionaries, and teaching materials. (At the time, few sites offered bulletin board or chat services.) Her paper was largely descriptive and prescriptive, concluding in favor of the positive potential of the Internet to revitalize endangered languages.

More recently, in an integrative literature review, Patrick Eisenlohr (2004) considers how electronic mediation technologies used to revitalize endangered languages both shape and are informed by linguistic ideologies. Lest we become overly enthusiastic about these technologies, he cautions that "many populations interested in reestablishing the practice of a lesser-used language often are least likely to engage in digital mediation practices" (Eisenlohr 2004: 26). Similarly, he expresses skepticism about the potential of digitally mediated language materials to facilitate community building (Eisenlohr 2004: 36–37). While online materials can foster virtual ties among people scattered around the globe in a diaspora, they may not necessarily foster face-to-face use of the language and its transmission to the younger generation (see also Fishman 2001: 458–459). Another issue is that indigenous peoples of the American Southwest are known to have objected in the past to the introduction of writing, during Mexican and Spanish rule, in order to keep cultural knowledge secret, and some groups continue to object to it today (Bielenberg 1999; Hinton 2001). Thus, these groups may be less likely to embrace Internet services.

In the late 1990s and early 2000s, linguists and activists began reporting case studies of Internet use in connection with specific endangered or minority languages.[35] Most of these were descriptive overviews, generally advocating online services enthusiastically; some also have an implicit or explicit prescriptive element, offering recommendations as to how to improve services, what new components might be added, and how they will enhance revitalization. Few case studies so far have been truly evaluative and empirical in orientation. Here we discuss two exceptions, the pioneering work of Mark Warschauer and his colleagues on Hawaiian, and action research on Sardinian by a team based in Germany.[36]

6.1. Hawaiian

"By the time Hawaii became a state in 1959, Hawaiian was spoken only by a few thousand elders" (Warschauer 1998: 141). As part of a cultural renaissance movement in the 1970s, Hawaiian pre-schools and kindergarten-to-12th-grade language immersion schools were created, as well as undergraduate and graduate programs in Hawaiian studies and language in universities. While microfiches and video recordings began to preserve cultural materials, there was a serious problem of access: "Native Hawaiians [were] dispersed in urban and rural communities on seven islands" (Warschauer 1998: 142).

In the early-to-mid-1990s, an Internet-accessible telecommunications system called *Leokī* ('Powerful Voice') was created, very likely the first in the world to operate entirely in an indigenous language.[37] The system includes private email, chat, discussion lists, a "newsline" (containing advertisements, announcements, information about language classes), vocabulary lists, current

and back issues of a newspaper, an area for cultural resources like stories and songs, and information about agencies supporting Hawaiian studies and language-learning – all in Hawaiian only (Figure 7).

Figure 7. *Leokī*, a telecommunications system for revitalization of Hawaiian. http://www.olelo.hawaii.edu/eng/resources/leoki.html, retrieved August 15, 2005. Reproduced with permission.

After two years of ethnographic research including participant observation and interviews, Warschauer concluded that "[t]he work of Hawaiians represents an excellent model of a group of people working to positively amplify existing cultural practices in an on-line environment" (Warschauer 1998: 157). At the same time, he noted that problems persist in relation to computer operating systems, and that many natives still could not access the system from their homes, because they lacked computers and even telephones (Warschauer 2002: 66).[38]

6.2. Sardinian

Sardinian is an unstandardized Romance language with about one million speakers, mainly concentrated on the Italian island of Sardinia. It is spoken mainly among friends and family, and more in rural than urban settings, co-existing

uneasily with standard Italian. Its status was improved somewhat when the island's legislature recognized it as an official language in 1997; in 1999 the State passed a law allowing it to be used in school (Grimaldi and Remberger n.d.). Despite these improvements in its legal status and the large number of speakers, Sardinian is considered an endangered language, since there is serious generational decline in its use (Mensching 2000).[39]

A team of linguists at two German universities, Guido Mensching and Lucia Grimaldi (Free University of Berlin) and Jürgen Rolshoven and Eva Remberger (University of Cologne), maintains a website housing a project called *Limba e Curtura de sa Sardigna* ('Language and Culture of Sardinia'; see Figure. 8). The project provides information about Sardinian; offers an asynchronous forum to discuss issues relating to the language as well as a chat mode (though little used, as yet); collects linguistic data with the aid of participants in the forum; documents the language and its texts; and archives and analyzes the linguistic data collected (Mensching 2000: 3).

Figure 8. *Limba e curtura de sa Sardigna* ('Language and Culture of Sardinia'); http://www.lingrom.fu-berlin.de/sardu/; retrieved August 15, 2005. Reproduced with permission.

Unlike Warschauer or McClure, this group does not merely study what others are doing to promote the languages at issue; they are themselves at the forefront of these efforts, although none of them is a native speaker of Sardinian. What began as an experiment a decade ago has taken on a life of its own, with subscribers to the mailing list becoming in effect collaborators of the researchers, supplying valuable data about Sardinian by their very efforts to write a formerly spoken-only language, and by serving as informants. As a result, the researchers are already able to report significant new findings. For instance, written communication on the mailing list between speakers/writers of two different dialects turns out to be mutually intelligible. Also, by interviewing subscribers on their use of three variants of the particle of assertion – *emmo*, *eja*, and *si* – the researchers were able to redraw the map for distribution of the three choices and to modify their explanation for the variation (Mensching 2000). In practical terms, they believe that activities related to the site will make an eventual contribution toward standardization of the language.

7. Conclusion

The research surveyed in this chapter reveals a complex and somewhat paradoxical state of affairs as regards linguistic diversity on the Internet at the present time. On the one hand, English has a historical advantage and continues to dominate many online contexts. Offline, as well, interest in English as a second language is growing: More young people around the world are learning English now than at any other time (Graddol 1997/2000). The Internet alone is not responsible for this spread, of course: Other political, economic, and cultural forces had already made English a global lingua franca (Crystal 2003). The Internet has, however, facilitated interaction among participants in multilingual nations, regions, and around the world, a number of whom employ English as a common language of communication. This, in turn, further strengthens the global position of English online and offline.

On the other hand, there is evidence that speakers of other languages are making the Internet their own. Warschauer (2001) calls this a situation of emerging diglossia online, with other languages besides English increasingly creating their own niche. ASCII is being modified to accommodate non-roman writing systems; fonts and operating systems are being localized, and Unicode is being expanded; the use of non-English CMC is growing; and the Web is being used to aid language revitalization efforts. Even when communicating in English, speakers of other languages may use the Internet to pursue the agendas of their own language and culture group, as in the case of the Basque forums described by Fernandez (2001). The current situation is far from one of absolute English dominance. Moreover, linguistic diversity online is on the rise (Paolillo 2007).

This survey is of necessity selective due to space limitations; we do not claim to have addressed all issues relevant to multilingualism on the Internet, and some crucial research has yet to be conducted. The following areas we feel are especially in need of attention in future research:

- Accurate assessments of what languages are used online (Paolillo 2007). Most estimates of language use on the Internet thus far have been based on population demographics, rather than on actual language use.
- Studies of the cognitive, social, and symbolic effects of technical (e.g., ASCII) bias on non-native English Internet users (cf. Pargman and Palme 2004). Software localization efforts notwithstanding, legacy systems will continue to be used for some time; to what extent are non-English speakers who must use them at a disadvantage?
- Comparative studies that use the same methods and concepts to investigate different languages and writing systems online, holding the mode of CMC constant (Danet and Herring 2003b). Most of the research produced to date has been case studies of individual languages.[40]
- More research into the effects of gender and culture on online language variation. To the extent that gender is culturally shaped, one would expect to find gender-based variation in CMC reflecting national, ethnic, and linguistic identities. Very little research has yet been done in this area.
- Systematic studies of language use and language choice in the design of websites. Most research on Internet multilingualism so far has focused on CMC modes such as email, chat, and discussion forums. How many languages are used on the Web today?[41] What proportion of Web sites is available in more than one language? In which languages? Do the versions in different languages communicate the same content and meanings?
- Research into the online status of other world languages besides English – for example, Spanish, Chinese, Russian, Malay and Hindi, each of which has a large regional base (cf. Graddol 1997/2000). In what languages are, for example, search engines available? How widespread is their use in these languages, and is their functionality comparable to that of search engines in English?
- Studies of automated translation. Can real-time machine translation during instant messaging, email, and chat facilitate communication among speakers of different languages? What obstacles stand in the way of such facilitation (cf. Climent et al. 2003, 2007)? How accurate is automatic translation of online content today, and what are the prospects for its improvement?

Like key transportation and telecommunication technologies before it, computer networking is leading the nations of the world into a new communication era, in which contact among speakers of different language varieties will in-

crease (Baron 2000; Crystal 2004b). The multilingual nature of the Internet is thus likely to become a phenomenon of even greater importance in the years to come. In a presentation on the subject of "Applied Linguistics for the 21st Century" prepared for the 1999 conference of the International Association for Applied Linguistics, Warschauer (2001) identified three issues of concern to applied linguists as a result of the spread of the Internet: literacy, second language learning and teaching, and international language use. Similarly, Crystal (2005) predicts the rise of what he terms an "Applied Internet Linguistics" – a linguistics that attends to language-related social problems arising from the emergence and prominence of the Internet as a new communication domain. We believe that multilingualism on the Internet should be a central focus of both agendas.

8. Appendix: Some languages accommodated by the five most widely used scripts incorporated in Unicode.*

Script	Languages
Latin	Afrikaans, Ainu, Albanian [1], Amo, Aymara, Azeri, Azerbaijani, Balear, Baluchi, Basque, Batak, Batak toba, Bosnian, Breton, Bahasa, Catalan, Chamorro, Cherokee, Cornish, Corsican, Cree, Croatian, Czech, Danish, Dutch, Edo, English, Esperanto, Estonian, Faroese, Fijian, Filipino, Finnish, French, Frisian, Gaelic, Gallegan, Gascon, German, Guarani, Haitian, Hanunóo, Hausa, Hawaiian, Hiri Motu, Hmong, Hopi, Hungarian, Ibibio, Icelandic, Indonesian, Ingush, Inuktitut, Iñupiaq, Irish, Italian, Javanese, Kalaallisut, Kanuri, Karelian, Khasi, Kinyarwanda, Kirghiz, Komi, Kurdish, Lapp, Latin, Latvian, Lingala, Lisu, Lithuanian, Lushootseed, Luxemburghish, Malay, Maltese, Manx, Mari, Mende, Malagasy, Maori, Marshallese, Moldavian, Naga, Nauru, Navajo, Niuean, Northern Sotho, North Ndebele, Norwegian, Nyanja, Oromo, Ossetic, Polish, Portuguese, Provençal, Prussian, Quechua, Rhaeto-Romance, Romanian, Romany, Rundi, Sami, Samoan, Sango, Serbian, Shona, Slovak, Slovenian, Somali, Songhai, Southern Sotho, Spanish, South Ndebele, Swahili, Swati, Swedish, Tagalog, Tagbanwa, Tahitian, Tajik, Tamashek, Tamazight, Time, Tetum, Tokelau, Tok Pisin, Tonga, Tsonga, Tswana, Turkish, Turkmen, Tuvalu, Udmurt, Uighur, Uzbek, Valencian, Venda, Vietnamese, Welsh, Wolof, Xhosa, Yi, Yoruba, Zulu

Cyrillic	Abaza, Abkhaz, Adygei, Aisor, Altai, Avar, Azeri, Azerbaijani, Balkar, Bashkir, Belarusian, Bulgarian, Buryat, Chechen, Chukchi, Chuvash, Dargwa, Dungan, Cyrillic, Gagauz, Inuktitut, Kabardian, Kalmyk, Khanty, Karachay, Karakalpak, Karelian, Kazakh, Khakass, Kirghiz, Komi, Koryak, Kurdish, Lak, Lezghian, Macedonian, Mansi, Mari, Moldavian, Mongolian, Mordvin, Nanai, Nenets, Netets, Nogai, Ossetic, Romanian [1], Romany, Russian, Sami, Selkup, Serbian, Shor, Tabasaran, Tajik, Tat, Tatar, Turkmen, Tuva, Udekhe, Udmurt, Uighur, Ukranian, Uzbek, Yakut
Devanagari	Awadhi, Bagheli, Balti, Bateri, Bhili, Bhojpuri, Bihari, Braj Bhasha, Chhattisgarhi, Garhwali, Gondi, Harauti, Hindi, Ho, Kachchi, Kanauji, Kankan, Kashmiri, Konkan, Limbu, Maithili, Marathi, Marwari, Mundari, Nepali, Newari, Pali, Sanskrit, Santali, Sherpa, Sindi
Arabic	Arabic, Azerbaijani, Baluchi, Farsi, Hausa [1], Indonesian [1], Ingush, Kashmiri, Kirghiz [1], Kurdish, Malay [1], Parsi-dari, Pashto, Punjabi, Sindhi, Tajik [1], Turkish [1], Turkmen [1], Uighur, Urdu, Uzbek, Wolof
Bengali	Assamese, Bengali, Chakma, Garo, Khasi, Meitei, Mundari, Naga, Riang, Santali, Sylhetti

* Compiled from "Languages and Scripts", http://www.unicode.org/onlinedat/languages-scripts.html, last updated June 3, 2005, retrieved September 22, 2005; additional languages provided by Deborah Anderson, personal email communication to Brenda Danet, August 29, 2005.

[1] Formerly or historically used this script, now uses another.

Notes

1. This is an expanded revision of a paper presented at the Annual Meeting of the American Association for the Advancement of Science, panel on "Language and the Internet: Usage Patterns, Global Issues, Future Trends", convened by Naomi S. Baron, Washington, DC, February 18, 2005. We thank Deborah Anderson and Keola Donaghy for helpful information, comments, and suggestions.
2. Internet Statistics: Distribution of Languages on the Internet, http://www.netz-tipp.de/languages.html; retrieved August 12, 2005. The figure refers to Web pages, not the Internet generally, despite the site's title.
3. See http://global-reach.biz/globstats/evol.html, retrieved August 12, 2005.
4. Although as of February 2005, only 7% of China's population of about 1,282,000,000 was online, this represented a growth rate of over 300% since 2000; similarly, whereas only a tiny 1.7% of the population of India (approximately 1,095,000,000) was online in early 2005, the rate of growth in the same five-year

period was 270% (Internet World Stats, Internet Usage in Asia, www.internetworldstats.com/stats3.htm, retrieved August 12, 2005).

5. The bias toward generalizing from materials in one's own language is probably also true of many CMC researchers writing in other languages. One of the first to point out this bias was Daniel Pargman (1998).
6. Another recent collection is the UNESCO-sponsored journal issue about multilingualism edited by Sue Wright (2004a), based on a survey administered to students of English in 10 countries (Tanzania, Indonesia, the United Arab Emirates, Oman, France, Italy, Poland, Macedonia, Japan, and Ukraine).
7. For additional references, see the introduction to Danet and Herring (2007).
8. An exception is Swahili, which is written in the standard Roman alphabet (The Kamusi Project); The Internet Living Swahili Dictionary, http://research.yale.edu/cgi-bin/swahili/main.cgi?right_frame_src=http%3A//www.yale.edu/swahili/home.html, retrieved August 15, 2005). For lists of languages written in variations of the roman alphabet, see http://www.omniglot.com/writing/languages.htm#latin, retrieved August 15, 2005, and the Appendix to this chapter.
9. Rare among Semitic languages, Maltese is written not in right-to-left non-roman script as is true of Hebrew, Arabic, and neo-Aramaic, but in roman characters.
10. *Hor* means "fornication" or "adultery", and *by* means "village". This example is from Pargman (1998) and Pargman and Palme (2004).
11. See the discussion of the extended ASCII character set in this chapter. A serious attempt is underway to make URLs multilingual. MINC: Multilingual International Names Consortium, http://www.minc.org/, is a non-profit international organization working to create "truly multilingual Internet domain names and keywords, [and] internationalization of Internet names, standards and protocols". Retrieved August 15, 2005.
12. See http://www.omniglot.com/writing/hawaiian.htm, retrieved August 15, 2005.
13. For explanations of word-processing in Japanese and Chinese, see Gottlieb (2000); Nishimura (2003); and Su (2003, 2007).
14. This claim was made not by Fishman, author of the article, but by the editors of *Foreign Policy*.
15. Fig. 3 also shows evidence of online code-mixing and code-switching. In the last contribution by speaker D, English words are mixed with Arabic ones within the same sentence. Also, there is a vertical "smiley" in the third line, typographically distinct from the familiar horizontal smiling face in use online in the West. For further discussion of vertical emoticons, see the section on *kaomoji*. On writing a very different, previously "oral" (Creole) language online and code-switching, see Hinrichs (2004, 2006).
16. In Georgakopoulou's work romanization per se is not discussed. See also Paolillo (1996) on a Usenet newsgroup called *soc.culture.punjab*, in which Punjabi is written in roman letters.
17. ISO is the acronym for the International Organization for Standards; see http://www.iso.org/iso/en/ISOOnline.frontpage, retrieved March 13, 2005. For further information on pre-Unicode, partial solutions for individual languages or language groups, see the Tutorial on Character Code Issues, http://www.cs.tut.fi/~jkorpela/chars.html, retrieved August 15, 2005; Everson (2002). The character set in Fig. 3 does not include several characters in the Maltese alphabet; see the URLs in Note 8.

18. The first recorded use of the term "Unicode" was by Joe Becker, in 1987, to refer to "unique, universal, and uniform character encoding". The Unicode Consortium was incorporated in the State of California in 1991. For more details on the history of Unicode, see http://www.unicode.org/history/, retrieved August 15, 2005.
19. This definition comes from the Unicode Home Pages, http://www.unicode.org/glossary/, retrieved August 15, 2005.
20. Strictly speaking, this is an over-simplification; see Note 22.
21. See http://www.linguistics.berkeley.edu/sei/, retrieved August 15, 2005. Volunteers do much of the work. Michael Everson is especially active in this effort; see his website, Everson Typography, http://www.evertype.com/, retrieved August 15, 2005.
22. This figure was cited by Deborah Anderson in a lecture at Yale University Library, June 30, 2005. Unicode Standard (4.0) accommodated 96,000 characters. Version 4.1.0, a minor update, added 1273 new ones. Contrary to intuition, Unicode is not "simply a 16-bit code where each character takes 16 bits and therefore there are 65,536 possible characters. This is [...] the single most common myth about Unicode" (Spolsky 2003).
23. Script per se is not a problem in Swahili, as the roman alphabet is used; see The United Republic of Tanzania Website, Frequently Asked Questions, http://www.tanzania.go.tz/learn_kiswahilif.html, retrieved August 15, 2005.
24. See The Languages of Tanzania: A Web Links Collection, http://www.african.gu.se/tanzania/weblinks.html, retrieved August 15, 2005.
25. The official languages are: Swedish, Danish, German, Czech, Greek, Estonian, French, Latvian, Italian, Lithuanian, Dutch, Hungarian, Portuguese, Maltese, Finnish, Polish, English, Slovak, Spanish, Slovene, as well as Irish Gaelic, which was added in June 2005 (Marsh 2005). See http://europa.eu.int/comm/translation/index_en.htm and http://europa.eu.int/comm/scic/interpreting/faq_en.htm, both retrieved August 15, 2005. In a note appended to Fishman (1998), the editors of *Foreign Policy* comment:

    > Despite these [...] initiatives [...] efforts to promote equal treatment of all official languages [...] have fallen flat. Although 15 percent of the European Commission's 17,000 personnel are translators, interpreters, and terminologists, EU institutions use only a handful of "working languages" to conduct daily business [...] although French is still used more frequently than English in the European Commission, English is preferred among younger officials. (Fishman 1998: 29)

26. The forum was closed after the constitution treaty was signed in October 2004. Material relating to Futurum debates is now archived at http://europa.eu.int/constitution/futurum/index_en.htm, retrieved August 15, 2005.
27. The 1998 survey found that nearly half of all Europeans spoke English, whereas only about one third spoke German, and fewer spoke French. The position of French and German was reversed in the 2000 survey, but English retained its dominance. See Eurobarometer (1999) and Eurobarometer (2001).
28. On the Punjabi script, see http://www.omniglot.com/writing/gurmuki.htm, retrieved August 15, 2005.
29. See http://www.activeworlds.com/, retrieved August 15, 2005. In these settings communication is mainly via text, but users are represented visually by graphical avatars and can move about the 3-D virtual environment.

30. On the Basque presence on the Internet, see also Uberuaga (2000) and Arbelaiz (2001). While Fernandez mentions various technical obstacles to using Basque online, the Basque alphabet is apparently not one of them. With the exception of *ñ* and *Ñ*, all letters can apparently be accommodated by the basic roman alphabet.
31. See, e.g. http://www.language-museum.com/a/assyrian-neo-aramaic.php; http://www.omniglot.com/writing/syriac.htm, both retrieved August 15, 2005.
32. For somewhat similar findings on switching into the home language for expressive purposes, but in a different context, see Androutsopoulos (2007).
33. See http://www.aina.org/bbs/index.cgi? (retrieved August 15, 2005), although it took some hunting to find a posting in Assyrian.
34. What constitutes a "viable presence" is an important research question in its own right.
35. The latest estimate is that there are over 6,900 languages in the world; see *Ethnologue*, http://www.ethnologue.com/, retrieved August 15, 2005.
36. See also Sperlich's (2005) case study of Niuean, a Pacific Island language.
37. The system began running in 1993, although used only by a small pilot group; it became Internet-accessible and started to be used by more people in 1994; immersion school access began in earnest in 1995. Personal email communication to Brenda Danet from Keola Donaghy, manager of the system, August 25, 2005.
38. Updating this overview in 2005, Keola Donaghy wrote, "The most significant development is our use of Leoki in Hawaiian language instruction. We've had over a hundred students in about 25 US states, as well as Europe and Japan over the past three years". Personal email communication to Brenda Danet, August 25, 2005.
39. For more on the Sardinian situation, see http://www.uoc.es/euromosaic/web/document/sard/an/i1/i1.html#3.7, retrieved August 15, 2005. On the endangered status of four dialects of Sardinian, see the UNESCO (2003) *Redbook on Endangered Languages*, http://www.helsinki.fi/~tasalmin/europe_index.html, retrieved August 15, 2005.
40. Pioneering exceptions are Fouser, Narahiko, and Chungmin (2000) and Androutsopoulos (2007).
41. With little effort, David Crystal found evidence for the presence of about 1,000 languages, and estimated that about 1,500 were on the Web at the time of his writing (Crystal 2004b: 88).

References

Allison, Anne
2004 Cuteness as Japan's millennial product. In: Joseph Tobin (ed.), *Pikachu's Global Adventure: The Rise and Fall of Pokemon*, 34–49. Durham, NC: Duke University Press.

Almanac, Computer Industry
2000 Population explosion. http://www.clickz.com/stats/sectors/geographics/article.php/5911_151151, retrieved August 12, 2005.

Anderson, Deborah
2004 The script encoding initiative (original English version of article published in German). *SIGNA* 6: 1–12. http://www.linguistics.berkeley.edu/~dwanders/SIGNAEnglish.pdf, retrieved August 12, 2005.

Anderson, Deborah
2005 Global linguistic diversity for the Internet. *Communications of the ACM* 48 (1): 27–28.

Androutsopoulos, Jannis
1999a Spelling variation in Latin alphabeted Greek email messages. Conference paper, 20th Working Meeting of the Linguistics Department, Aristoteles University of Thessaloniki, Thessaloniki, Greece, April 23–25, 1999. http://greekweb.archetype.de/greekmail/abstracts.htm, retrieved August 12, 2005 (abstract in English and Greek).

Androutsopoulos, Jannis
1999b Latin-Greek spelling in e-mail messages: Usage and attitudes (in Greek). *Studies in Greek Linguistics* 1: 75–86.

Androutsopoulos, Jannis
2000 From dieuthinsi to diey8ynsh: Orthographic variation in Latin-alphabeted Greek (in Greek). Conference paper, 4th International Conference on Greek Linguistics, University of Nicosia, Cyprus, September, 1999. http://greekweb.archetype.de/greekmail/icgl99.html, retrieved August 12, 2005 (in Greek; abstract in English).

Androutsopoulos, Jannis
2007 Language choice and code-switching in German-based diasporic web forums. In: Danet and Herring (eds.), *The Multilingual Internet: Language, Culture, and Communication Online*. 340–361.

Androutsopoulos, Jannis and Volker Hinnenkamp
2001 Code-switching in der bilingualen Chat-Kommunikation: Ein explorativer Blick auf #hellas und #turks. In: Michael Beißwenger (ed.), *Chat-Kommunikation. Sprache, Interaktion und Sozialität in synchroner computervermittelter Kommunikation. Perspektiven auf ein interdisziplinäres Forschungsfeld*, 367–401. Stuttgart: Ibidem.

Arbelaiz, Asunción Martínez
2001 Basque, the Internet, and new language policies. In: Christopher Moseley, Nicholas Ostler and Hassan Ouzzate (eds.), *Endangered Languages and the Media: Proceedings of the Fifth FEL (Foundation for Endangered Languages) Conference, Agadir, Morocco, 20–23 September 2000*, 98–103. Bath, UK: Foundation for Endangered Languages.

Axelsson, Ann-Sofie, Åsa Abelin and Ralph Schroeder
2003 Anyone speak Spanish? Language encounters in multi-user virtual environments and the influence of technology. *New Media & Society* 5(4): 475–498.

Axelsson, Ann-Sofie, Åsa Abelin and Ralph Schroeder
2007 Anyone speak Swedish? Language encounters in multi-user virtual environments and the influence of technology. In: Danet and Herring (eds.), *The Multilingual Internet: Language, Culture, and Communication Online*. 362–381.

Baron, Naomi S.
2000 *Alphabet to Email: How Written English Evolved and Where It's Heading*. New York/London: Routledge.

Bechar-Israeli, Haya
1995 From <Bonehead> to <clonehead>: Nicknames, play and identity on Internet Relay Chat. *Journal of Computer-Mediated Communication* 1(2). http://jcmc.indiana.edu/vol1/issue2/bechar.html, retrieved August 15, 2005.

Berjaoui, Nasser
2001 Aspects of the Moroccan Arabic orthography with preliminary insights from the Moroccan computer-mediated communication. In: Michael Beißwenger (ed.). *Chat-Kommunikation: Sprache, Interaktion, Sozialität & Identität in synchroner computervermittelter Kommunikation. Perspektiven auf ein interdisziplinäres Forschungsfeld*, 431–465. Stuttgart: Ibidem.

Bielenberg, Brian
1999 Indigenous language codification: Cultural effects. In: Jon Reyhner, Gina Cantoni, Robert N. St. Clair and Evangeline Parsons Yazzie (eds.). *Revitalizing Indigenous Languages*, 103–112. Flagstaff, AZ: Northern Arizona University. Online at http://jan.ucc.nau.edu/~jar/RIL_8.html, retrieved August 15, 2005.

Buszard-Welcher, Laura
2001 Can the Web save my language? In: Leanne Hinton and Kenneth Hale (eds.). *The Green Book of Language Revitalization in Practice*, 331–348. San Diego: Academic Press.

Climent, Salvador, Joaquim Moré, Antoni Oliver, Míriam Salvatierra, Imma Sànchez, Mariona Taulé and Lluïsa Vallmanya
2003 Bilingual newsgroups in Catalonia: A challenge for machine translation. *Journal of Computer-Mediated Communication* 9(1). http://jcmc.indiana.edu/vol9/issue1/climent.html, retrieved August 15, 2005.

Climent, Salvador, Joaquim Moré, Antoni Oliver, Míriam Salvatierra, Imma Sànchez and Mariona Taulé
2007 Can machine translation enhance the status of Catalan versus Spanish in online academic forums? In: Danet and Herring (eds.), *The Multilingual Internet: Language, Culture, and Communication Online*, 209–230.

Consortium, The Unicode
2003 *The Unicode standard, version 4.0 (online)*. Boston: Addison-Wesley. Also online at http://www.unicode.org/versions/Unicode4.0.0/, retrieved August 15, 2005.

Cringely, Bob
1998 *Nerds 2.0: History of the Internet*. Videotape. New York: Ambrose.

Crystal, David
2000 *Language Death*. Cambridge: Cambridge University Press.

Crystal, David
2001 *Language and the Internet*. Cambridge: Cambridge University Press.

Crystal, David
2003 *English as a Global Language*. 2nd ed. Cambridge: Cambridge University Press.

Crystal, David
2004a *The Stories of English*. New York: Overlook Press.

Crystal, David
2004b *The Language Revolution*. London: Polity Press.

Crystal, David
2005 The scope of Internet linguistics. Conference paper, Annual Meeting, American Association for the Advancement of Science, panel on "Language on the Internet: Usage Patterns, Global Issues, Future Trends," Naomi S. Baron, convenor. Washington, DC, February 17–21, 2005.

CyberAtlas
2003 June 6 Population explosion! http://www.clickz.com/stats/sectors/geographics/article.php/5911_151151, retrieved August 12, 2005.
Dalby, Andrew
2002 *Language in Danger*. London & New York: Allen Lane.
Danet, Brenda (ed.)
1995 Play and Performance in Computer-Mediated Communication. *Journal of Computer-Mediated Communication* [Special issue]. http://jcmc.indiana.edu/vol1/issue2/, retrieved August 15, 2005.
Danet, Brenda
1998 Text as mask: Gender, play and performance on the Internet. In: Steven G. Jones (ed.), *Cybersociety 2.0: Revisiting Computer-Mediated Communication and Community*, 129–158. Thousand Oaks, CA, and London: Sage.
Danet, Brenda
2001 *Cyberpl@y: Communicating Online*. Oxford: Berg Publishers. Companion Website: http://pluto.mscc.huji.ac.il/~msdanet/cyberpl@y/, retrieved August 15, 2005.
Danet, Brenda and Susan C. Herring (eds.)
2003a The Multilingual Internet: Language, Culture and Communication in Instant Messaging, Email and Chat. *Journal of Computer-Mediated Communication* 9(1) [Special issue]. http://jcmc.indiana.edu/vol9/issue1/, retrieved August 15, 2005.
Danet, Brenda and Susan C. Herring
2003b Introduction: The multilingual Internet. *Journal of Computer-Mediated Communication* 9(1). http://jcmc.indiana.edu/vol9/issue1/intro.html, retrieved August 15, 2005.
Danet, Brenda and Susan C. Herring (eds.)
2007 *The Multilingual Internet: Language, Culture, and Communication Online*. New York: Oxford University Press.
Danet, Brenda, Lucia Ruedenberg and Yehudit Rosenbaum-Tamari
1998 "Hmmm ... Where's that smoke coming from?" Writing, play and performance on Internet Relay Chat. In: Fay Sudweeks, Margaret McLaughlin and Sheizaf Rafaeli (eds.), *Network and Netplay: Virtual Groups on the Internet*, 47–85. Cambridge, MA: AAAI/MIT Press.
Demont-Heinrich, Christof
2005 Language and national identity in the era of globalization: The case of English in Switzerland. *Communication Inquiry* 29(1): 66–84.
de Oliveira, Sandi Michele
2003 Breaking conversational norms on a Portuguese users' network: Men as adjudicators of politeness? *Journal of Computer-Mediated Communication* 9 (1). http://jcmc.indiana.edu/vol9/issue1/oliveira.html, retrieved August 15, 2005.
de Oliveira, Sandi Michele
2007 Breaking conversational norms on a Portuguese users' network: Men as adjudicators of politeness? In: Danet and Herring (eds.), *The Multilingual Internet: Language, Culture, and Communication Online*, 256–277.

Dor, Daniel
2004 From Englishization to imposed multilingualism: Globalization, the Internet, and the political economy of the linguistic code. *Public Culture* 16(1): 97–118.

Durham, Mercedes
2003 Language choice on a Swiss mailing list. *Journal of Computer-Mediated Communication* 9(1). http://jcmc.indiana.edu/vol9/issue1/durham.html, retrieved August 15, 2005.

Durham, Mercedes
2004 The future of Swiss English. In: Britt-Louise Gunnarsson, Lena Bergström, Gerd Eklund, Staffan Fridell, Lisa H. Hansen, Angela Karstadt, Bengt Nordberg, Eva Sundgren and Mats Thelander (eds.), *Language Variation in Europe: Papers from Iclave 2 (Second International Conference on Language Variation in Europe, Uppsala University, Sweden, June 12–14, 2003)*, 156–167. Uppsala: Uppsala University.

Durham, Mercedes
2007 Language choice on a Swiss mailing list. In: Danet and Herring (eds.), *The Multilingual Internet: Language, Culture, and Communication Online*. 319–339.

Dürmüller, Urs
2002 English in Switzerland: From foreign language to lingua franca. In: D. J. Allerton, Paul Skandera and Cornelia Tschichold (eds.), *Perspectives on English as a World Language*, 115–123. Basel: Schwabe.

Eisenlohr, Patrick
2004 Language revitalization and new technologies: Cultures of electronic mediation and the refiguring of communities. *Annual Review of Anthropology* 33: 21–45.

Eurobarometer, European Union
1999 Eurobarometer Report Number 50: Public Opinion in the European Union. http://europa.eu.int/comm/education/policies/lang/languages/index_en.html, retrieved August 15, 2005.

Eurobarometer, European Union
2001 Eurobarometer Report Number 54: Europeans and Languages, Executive Summary. http://europa.eu.int/comm/public_opinion/archives/eb/ebs_147_summ_en.pdf, retrieved August 15, 2005.

Everson, Michael
2002 Leaks in the Unicode pipeline: Script, script, script ... (A Unicode Technical Note). Conference paper, the 21st International Unicode Conference, Dublin, Ireland, May 14–17, 2002. http://www.unicode.org/notes/tn4/, retrieved September 25, 2005.

Fahlman, Scott E.
n.d. Smiley lore :-). http://www-2.cs.cmu.edu/~sef/sefSmiley.htm, retrieved September 25, 2005.

Ferguson, Charles
1972 Diglossia. In: Pier Paolo Giglioli (ed.), *Language and Social Context*, 232–252. Harmondsworth: Penguin.

Fernandez, Luis
2001 Patterns of linguistic discrimination in Internet discussion forums. *Mercator Media Forum* 5: 22–41.

Ferrara, Kathleen, Hans Brunner and Greg Whittemore
1991 Interactive written discourse as an emergent register. *Written Communication* 8: 8–34.
Fishman, Joshua
1998 *Foreign Policy* 113, Winter 1998/99: 26–40.
Fishman, Joshua (ed.)
2001 *Can Threatened Languages Be Saved? Reversing Language Shift, Revisited: A 21st Century Perspective*. Clevedon: Multilingual Matters.
Fishman, Joshua A., Andrew W. Conrad and Alma Rubal-Lopez (eds.)
1996 *Post-imperial English: Status Change in Former British and American Colonies, 1940–1990*. Berlin/New York: Mouton de Gruyter.
Fouser, Robert, Narahiko Inoue and Chungmin Lee
2000 The pragmatics of orality in English, Japanese and Korean computer-mediated communication. In: Lyn Pemberton and Simon Shurville (eds.), *Words on the Web: Computer-Mediated Communication*, 52–62. Bristol, UK: Intellect Books.
Gabrial, Albert
1998 3,000 years of history, yet the Internet is our only home. *Cultural Survival Quarterly* 21(4): 42.
Georgakopoulou, Alexandra
1997 Self-presentation and interactional alliances in e-mail discourse: The style- and code-switches of Greek messages. *International Journal of Applied Linguistics* 7: 141–164.
Georgakopoulou, Alexandra
2004 To tell or not to tell? Email stories between on- and off-line interactions. *language@internet* 1. Retrieved July 24, 2006 from http://www.digijournals.de/languageatinternet/articles/36.
Gottlieb, Nanette
2000 *Word-processing Technology in Japan: Kanji and the Keyboard*. Richmond: Curzon.
Graddol, David
1997/2000 *The Future of English? A Guide to Forecasting the Popularity of the English Language in the 21st Century*. London: The British Council. Available online since 2000 at http://www.gre.ac.uk/~ds42/pages/future%20of%20English.pdf, retrieved August 15, 2005.
Graddol, David
1999 The decline of the native speaker. *AILA Review* 13: 57–68.
Graddol, David
2004 The future of language. *Science* 303(5662), 1329–1331.
Grenoble, Lenore A. and Lindsay J. Whaley (eds.)
1998 *Endangered Languages: Language Loss and Community Responses*. Cambridge/New York: Cambridge University Press.
Grenoble, Lenore A. and Lindsay J. Whaley
2001 Endangered languages. In: Rajend Mesthrie (ed.), *Concise Encyclopedia of Sociolinguistics*, 465–467. Amsterdam/New York: Elsevier.
Grimaldi, Lucia and Remberger, Eva-Maria
n.d. The Promotion of the Sardinian Language and Culture via the Internet: Fields of Activity and Perspectives. http://www.lingrom.fu-berlin.de/sardu/grimaldi_remberger.html, retrieved July 24, 2006.

Hafner, Katie and Matthew Lyon
1996 *Where Wizards Stay Up Late: The Origins of the Internet*. New York: Simon & Schuster.

Herring, Susan C. (ed.)
1996a *Computer-Mediated Communication: Linguistic, Social, and Cross-Cultural Perspectives*. Amsterdam: John Benjamins.

Herring, Susan C.
1996b Posting in a different voice: Gender and ethics in computer-mediated communication. In: Charles Ess (ed.), *Philosophical Perspectives on Computer-Mediated Communication*, 115–145. Albany: SUNY Press.

Herring, Susan C.
2001 Computer-mediated discourse. In: Deborah Tannen, Deborah Schiffrin and Heidi Hamilton (eds.), *Handbook of Discourse Analysis*, 612–634. Oxford: Blackwell.

Herring, Susan C.
2002 The language of the Internet: English dominance or heteroglossia? Keynote presentation, CATaC (Conference on Cultural Attitudes towards Technology and Communication). Montreal, University of Montreal, 2002. Available from http://ella.slis.indiana.edu/~herring/CATaC.ppt, retrieved September 24, 2005.

Herring, Susan C.
2003 Gender and power in online communication. In: Janet Holmes and Miriam Meyerhoff (eds.), *The Handbook of Language and Gender*, 202–228. Oxford: Blackwell.

Hinrichs, Lars
2004 Emerging orthographic conventions in written creole: Computer-mediated communication in Jamaica. *Arbeiten aus Anglistik und Amerikanistik* 29(1): 81–109.

Hinrichs, Lars
2006 *Codeswitching on the Web: English and Jamaican Creole in E-mail Communication*. Amsterdam: John Benjamins.

Hinton, Leanne
2001 Introduction to the Pueblo languages. In: Leanne Hinton and Kenneth L. Hale (eds.), *The Green Book of Language Revitalization in Practice*, 61–62. San Diego: Academic Press.

Hinton, Leanne and Kenneth L. Hale (eds.)
2001 *The Green Book of Language Revitalization in Practice*. San Diego: Academic Press.

Hjorth, Larissa
2003 Cute@keitai.com. In: Nanette Gottlieb and Mark J. McLelland (eds.), *Japanese Cybercultures*, 50–59. London/New York: Routledge.

Hudson, Alan
2001 Diglossia. In: Rajend Mesthrie (ed.), *Concise Encyclopedia of Sociolinguistics*, 226–231. Amsterdam/New York: Elsevier.

Kanayama, Tomoko
2003 Ethnographic research on the experience of Japanese elderly people online. *New Media & Society* 5(2): 267–288.

Kataoka, Kuniyoshi
1997 Affect and letter-writing: Unconventional conventions in personal writing by young Japanese women. *Language in Society* 26(1): 103–136.

Kataoka, Kuniyoshi
2003a Form and function of emotive pictorial signs in casual letter writing. *Written Language and Literacy* 6(1): 1–29.

Kataoka, Kuniyoshi
2003b Emotion and youth identities in personal letter writing: An analysis of pictorial signs and unconventional punctuation. In: Jannis K. Androutsopoulos and Alexandra Georgakopoulou (eds.), *Discourse Constructions of Youth Identities*, 121–149. Amsterdam: John Benjamins.

Katsuno, Hirofumi and Christine R. Yano
2002 Face to face: On-line subjectivity in contemporary Japan. *Asian Studies Review* 26(2): 205–231.

Katsuno, Hirofumi and Christine R. Yano
2007 *Kaomoji* and expressivity in a Japanese housewives' chatroom. In: Danet and Herring (eds.), *The Multilingual Internet: Language, Culture, and Communication Online*. 278–300.

Kaul, Asha and Vaibhavi Kulkarni
2005 Coffee, tea, or ...? Gender and politeness in computer-mediated communication (CMC). *Indian Institute of Management Ahmedabad Working Papers*. http://ideas.repec.org/p/iim/iimawp/2005-04-02.html, retrieved September 25, 2005.

Kinsella, Sharon
1995 Cuties in Japan. In: Lise Skov and Brian Moeran (eds.), *Women, Media and Consumption in Japan*, 220–254. Honolulu: University of Hawaii Press.

Kinsella, Sharon
2000 *Adult Manga: Culture and Power in Contemporary Japanese Society*. Honolulu: University of Hawaii Press.

Koutsogiannis, Dimitris and Bessie Mitsikopoulou
2003 Greeklish and Greekness: Trends and discourses of "glocalness." *Journal of Computer-Mediated Communication* 9(1). http://jcmc.indiana.edu/vol9/issue1/kouts_mits.html, retrieved August 15, 2005.

Koutsogiannis, Dimitris and Bessie Mitsikopoulou
2007 Greeklish and Greekness: Trends and discourses of "glocalness". In: Danet and Herring (eds.), *The Multilingual Internet: Language, Culture, and Communication Online*. 142–160.

Krauss, Michael
1992 The world's languages in crisis. *Language* 68: 4–10.

Krauss, Michael
1998 The scope of the language endangerment crisis and recent response to it. In: K. Matsumara (ed.), *Studies in Endangered Languages*, 101–114. Tokyo: Hituzi Syobo.

Lakoff, Robin Tolmach
1975 *Language and Woman's Place*. New York: Harper & Row.

Lakoff, Robin Tolmach
2004 *Language and Woman's Place: Text and Commentaries* (revised and expanded edition by Mary Bucholtz). New York/Oxford: Oxford University Press.

Mafu, Safari
2004 From the oral tradition to the information era: The case of Tanzania. *International Journal on Multicultural Societies* 6(1): 53–78.

Mair, Christian
2002 The continuing spread of English: Anglo-American conspiracy or global grassroots movement? In: D. J. Allerton, Paul Skandera and Cornelia Tschichold (eds.), *Perspectives on English as a World Language*, 159–169. Basel: Schwabe.

Marsh, Bill
2005 Livonian spoken here (for now). *New York Times*, June 19.

McClure, Erica
2001a The role of language in the construction of ethnic identity on the Internet: The case of Assyrian activists in diaspora. In: Christopher Moseley, Nicholas Ostler and Hassan Ouzzate (eds.), *Endangered Languages and the Media: Proceedings of the Fifth FEL (Foundation of Endangered Languages) Conference, Agadir, Morocco, 20–23 September, 2001*, 68–75. Bath, UK: Foundation for Endangered Languages.

McClure, Erica
2001b Oral and written Assyrian-English codeswitching. In: Rodolfo Jacobson (ed.), *Codeswitching Worldwide II*, 157–191. Berlin: Mouton de Gruyter.

McVeigh, Brian J.
2000 How Hello Kitty commodifies the cute, cool and camp: 'Consumutopia' versus 'control' in Japan. *Journal of Material Culture* 5(2): 225.

Mensching, Guido
2000 The Internet as a rescue tool of endangered languages: Sardinian. Presented at the conference *Multilinguae: Multimedia and Minority Languages*, San Sebastian, Spain, November 8–9, 2000. http://www.lingrom.fu-berlin.de/sardu/, retrieved February 26, 2005.

Moseley, Christopher, Nicholas Ostler and Hassan Ouzzate
2001 Endangered languages and the media. In: Christopher Moseley, Nicholas Ostler and Hassan Ouzzate (eds.), *Endangered Languages and the Media: Proceedings of the Fifth FEL (Foundation for Endangered Languages) Conference, Agadir, Morocco, 20–23 September 2001*, 7–8. Bath, UK: Foundation for Endangered Languages.

Myhill, John and Eman Garra
2005 Writing Palestinian Arabic in electronic messages. Paper presented at the fourth conference, Israel Association for the Study of Language in Society. Bar Ilan University, Ramat Gan, Israel, June 14, 2005. http://www.biu.ac.il/js/hb/ials/abs.htm, retrieved September 21, 2005 (paper title in Hebrew and English; abstract in English).

Nettle, Daniel and Suzanne Romaine
2000 *Vanishing Voices: The Extinction of the World's Languages*. Oxford: Oxford University Press.

Nishimura, Yukiko
2003 Linguistic innovations and interactional features of casual online communication in Japanese. *Journal of Computer-Mediated Communication* 9(1). http://jcmc.indiana.edu/vol9/issue1/nishimura.html, retrieved August 15, 2005.

Nishimura, Yukiko
2007 Linguistic innovations and interactional features of casual online communication in Japanese. In: Danet and Herring (eds.), *The Multilingual Internet: Language, Culture, and Communication Online*. 163–183.

Nunberg, Geoffrey
2000 Will the Internet always speak English? *The American Prospect Online* 11(10). http://www.prospect.org/print/V11/10/nunberg-g.html, retrieved September 25, 2005.

O'Neill, Judy
1995 The Role of ARPA in the development of the ARPANET, 1961–1972. *Annals of the History of Computing* 17(4): 76–81.

Organization for Economic Co-operation and Development
2001 *Understanding the digital divide*. http://www.oecd.org/dataoecd/38/57/1888451.pdf, retrieved September 20, 2005.

Ostler, Nicholas
1999 Fighting words: As the world gets smaller, minority languages struggle to stake their claim. *Language International* 11(2): 38–45.

Palfreyman, David and Muhamed Al Khalil
2003 "A funky language for teenzz to use": Representing Gulf Arabic in instant messaging. *Journal of Computer-Mediated Communication* 9(1). http://jcmc.indiana.edu/vol9/issue1/palfreyman.html, retrieved August 15, 2005.

Palfreyman, David and Muhamed Al Khalil
2007 "A funky language for teenzz to use": Representing Gulf Arabic in instant messaging. In: Danet and Herring (eds.), *The Multilingual Internet: Language, Culture, and Communication Online*. 43–63

Panyametheekul, Siriporn and Susan C. Herring
2003 Gender and turn allocation in a Thai chat room. *Journal of Computer-Mediated Communication* 9(1). http://jcmc.indiana.edu/vol9/issue1/panya_herring.html, retrieved August 15, 2005.

Panyametheekul, Siriporn and Susan C. Herring
2007 Gender and turn allocation in a Thai chat room. In: Danet and Herring (eds.), *The Multilingual Internet: Language, Culture, and Communication Online*. 233–255.

Paolillo, John C.
1996 Language choice on soc.culture.punjab. *Electronic Journal of Communication* 6(3). http://ella.slis.indiana.edu/~paolillo/research/paolillo.publish.txt, retrieved September 25, 2005.

Paolillo, John C.
2007 How much multilingualism? Linguistic diversity on the Internet. In: Danet and Herring (eds.), *The Multilingual Internet: Language, Culture, and Communication Online*. 408–430.

Pargman, Daniel
1998 Reflections on cultural bias and adaptation. Paper presented at CATaC (*Cultural Attitudes towards Technology and Communication*). London, August 1–3, 1998. http://www.it.murdoch.edu.au/~sudweeks/catac98/pdf/06_pargman.pdf, retrieved September 25, 2005.

Pargman, Daniel and Jacob Palme
2004 Linguistic standardization on the Internet. In: Fay Sudweeks and Charles Ess (eds.), *Proceedings of CATaC: Cultural Attitudes Towards Technology and Communication 2004*, 385–388. Murdoch, Australia: Murdoch University Press.

Phillipson, Robert
1992 *Linguistic Imperialism*. Oxford: Oxford University Press.

Phillipson, Robert
2000 English in the new world order: Variations on a theme of linguistic imperialism and "world" English. In: Thomas Ricento (ed.), *Ideology, Politics and Language Policies: Focus on English*, 87–106. Amsterdam: John Benjamins.

Phillipson, Robert and Tove Skutnabb-Kangas
2001 Linguistic imperialism. In: Rajend Mesthrie (ed.), *Concise Encyclopedia of Sociolinguistics*, 570–574. Oxford: Elsevier.

Pimienta, Daniel
2002 Put out your tongue and say "aaah": Is the Internet suffering from acute "Englishitis"? *UNESCO Newsletter on Communication and Information in the Knowledge Society*, January 30, 2002. http://www.unesco.org/webworld/points_of_views/300102_pimienta.shtml, retrieved May 24, 2005.

Richie, Donald
2003 *Kawaii*: Kingdom of the cute. In: Donald Richie, *The Image Factory: Fads & Fashions in Japan*, 53–62. London: Reaktion.

Schiffman, Harold F.
1997 Diglossia in a sociolinguistic situation. In: Florian Coulmas (ed.), *The Handbook of Sociolinguistics*, 205–216. Oxford: Blackwell.

Schodt, Frederik L.
1986 *Manga! Manga! The World of Japanese Comics*. Tokyo, New York and San Francisco: Kodansha International.

Skutnabb-Kangas, Tove
1988 Multilingualism and the education of minority children. In: Tove Skutnabb-Kangas and J. Cummins (eds.), *Minority Education: From Shame to Struggle*, 9–44. Philadelphia: Multilingual Matters.

Sperlich, Wolfgang B.
2005 Will cyberforums save endangered languages? A Niuean case study. *International Journal of the Sociology of Language* 172(1): 51–78.

Spolsky, Joel
2003 The absolute minimum every software developer absolutely, positively must know about Unicode and character sets (no excuses!), October 8, 2003. http://www.joelonsoftware.com/articles/Unicode.html, retrieved August 25, 2005.

Su, Hsi-Yao
2003 The multilingual and multi-orthographic Taiwan-based Internet: Creative uses of writing systems on college-affiliated BBSs. *Journal of Computer-Mediated Communication* 9(1). http://jcmc.indiana.edu/vol9/issue1/su.html, retrieved August 15, 2005.

Su, Hsi-Yao
2007 The multilingual and multi-orthographic Taiwan-based Internet: Creative uses of writing systems on college-affiliated BBSs. In: Danet and Herring

(eds.), *The Multilingual Internet: Language, Culture, and Communication Online*. 64–86.

Sugimoto, Taku and James A. Levin
2000 Multiple literacies and multimedia: A comparison of Japanese and American uses of the Internet. In: Gail E. Hawisher and Cynthia L. Selfe (eds.), *Global Literacies and the World-Wide Web*, 133–153. London: Routledge.

Tannen, Deborah
1990 *You Just Don't Understand: Women and Men in Conversation*. New York: William Morrow.

Tseliga, Theodora
2007 "It's all *Greeklish* to Me!" Linguistic and sociocultural perspectives on Greeklish (Roman-alphabeted Greek) in asynchronous CMC. In: Danet and Herring (eds.), *The Multilingual Internet: Language, Culture, and Communication Online*. 116–141.

Tsuda, Yukio
2000 Envisioning a democratic linguistic order. *TESL Reporter* 33(1): 32–38.

Turkle, Sherry
1995 *Life on the Screen: Identity in the Age of the Internet*. New York: Simon & Schuster.

Uberuaga, Blas Pedro
2000 The Basque presence on the Internet: Yesterday, today and tomorrow. *Journal of the Society for Basque Studies in America* 20.

UNESCO
2003 *UNESCO Red Book on Endangered Languages* [World Wide Web], last updated March 4, 2003. http://www.tooyoo.l.u-tokyo.ac.jp/Redbook/, retrieved August 25, 2005.

Warschauer, Mark
1998 Technology and indigenous language revitalization: Analyzing the experience of Hawai'i. *The Canadian Modern Language Review/ La Revue canadienne des langues vivantes* 55(1): 141–161.

Warschauer, Mark
2001 Millennialism and media: Language, literacy and technology in the 21st century. *AILA Review* 14: 49–59. Online edition available at http://www.aila.soton.ac.uk/pdfs/Aila14.pdf, retrieved August 15, 2005.

Warschauer, Mark
2002 Languages.com: The Internet and linguistic pluralism. In: Ilana Snyder (ed.), *Silicon Literacies: Communication, Innovation and Education in the Electronic Age*, 62–74. London/New York: Routledge.

Warschauer, Mark and Keola Donaghy
1997 Leoki: A powerful voice of Hawaiian language revitalization. *Computer-Assisted Language Learning* 10(4): 349–361.

Warschauer, Mark, Ghada R. El Said and Ayman Zohry
2002 Language choice online: Globalization and identity in Egypt. *Journal of Computer-Mediated Communication* 7(4). http://jcmc.indiana.edu/vol7/issue4/warschauer.html, retrieved August 15, 2005.

Warschauer, Mark, Ghada R. El Said, and Ayman Zohry
2007 Language choice online: Globalization and identity in Egypt. In: Danet and Herring (eds.), *The Multilingual Internet: Language, Culture, and Communication Online*. 303–318.

Werry, Christopher C.
1996 Linguistic and interactional features of Internet Relay Chat. In: Susan C. Herring (ed.), *Computer-Mediated Communication: Linguistic, Social, and Cross-Cultural Perspectives*, 47–64. Amsterdam: John Benjamins.

Wheeler, Deborah
2001 Women, Islam, and the Internet: Findings in Kuwait. In: Charles Ess and Fay Sudweeks (eds.), *Culture, Technology, Communication: Towards an Intercultural Global Village*, 158–182. Albany, NY: SUNY Press.

Witmer, Diane F. and Sandra Lee Katzman
1997 On-line smiles: Does gender make a difference in the use of graphic accents? *Journal of Computer-Mediated Communication* 2(4). http://jcmc.indiana.edu/vol2/issue4/witmer1.html, retrieved August 15, 2005.

Wodak, Ruth and Scott Wright
2007 The European Union in cyberspace: Democratic participation via online multilingual discussion boards? In: Danet and Herring (eds.), *The Multilingual Internet: Language, Culture, and Communication Online*, 383–407.

Wright, Handel Kashope
1996 Email in African studies. *Convergence: The Journal of Research into New Media Technologies* 2(1): 20–30.

Wright, Scott
2004 A comparative analysis of government-run discussion boards at the local, national and European Union levels. Ph.D. thesis, University of East Anglia, Norwich.

Wright, Sue (ed.)
2004a Multilingualism on the Internet. *International Journal on Multicultural Societies* 6(1). http://portal.unesco.org/shs/en/ev.php-URL_ID=7142&URL_DO=DO_TOPIC&URL_SECTION=-465.html, retrieved August 25, 2005 (also available in French).

Wright, Sue
2004b Thematic introduction. *International Journal on Multicultural Societies* 6 (1): 3–11. http://portal.unesco.org/shs/en/ev.php-URL_ID=7142&URL_DO=DO_TOPIC&URL_SECTION=-465.html, retrieved August 25, 2005 (also available in French).

Yates, Simeon J.
1996 Oral and written linguistic aspects of computer-conferencing. In: Susan C. Herring (ed.), *Computer-Mediated Communication: Linguistic, Social, and Cross-Cultural Perspectives*, 29–46. Amsterdam: John Benjamins.