Language and the Internet

DAVID CRYSTAL

CAMBRIDGE
UNIVERSITY PRESS

Contents

1   *A linguistic perspective*

> Will the English-dominated Internet
> spell the end of other tongues?

> Quite e-vil: the mobile phone
> whisperers

> A major risk for humanity

These quotations illustrate widely held anxieties about the effect
of the Internet on language and languages. The first is the sub-
heading of a magazine article on millennial issues.[1] The second is
the headline of an article on the rise of new forms of impoliteness in
communication among people using the short messaging service
on their mobile phones.[2] The third is a remark from the President of
France, Jacques Chirac, commenting on the impact of the Internet
on language, and especially on French.[3] My collection of press clip-
pings has dozens more in similar vein, all with a focus on language.
The authors are always ready to acknowledge the immense tech-
nological achievement, communicative power, and social potential
of the Internet; but within a few lines their tone changes, as they
express their concerns. It is a distinctive genre of worry. But unlike
sociologists, political commentators, economists, and others who
draw attention to the dangers of the Internet with respect to such
matters as pornography, intellectual property rights, privacy, se-
curity, libel, and crime, these authors are worried primarily about
linguistic issues. For them, it is language in general, and individual
languages in particular, which are going to end up as Internet

---

[1] Used in an article by Jim Erickson, 'Cyberspeak: the death of diversity', *Asiaweek*, 3 July 1998, 15.
[2] Lydia Slater, in *The Sunday Times*, 30 January 2000, 10.
[3] 'Language and electronics: the coming global tongue', *The Economist*, 21 December 1996, 37.

casualties, and their specific questions raise a profusion of spectres. Do the relaxed standards of e-mails augur the end of literacy and spelling as we know it? Will the Internet herald a new era of technobabble? Will linguistic creativity and flexibility be lost as globalization imposes sameness?

There is of course nothing new about fears accompanying the emergence of a new communications technology. In the fifteenth century, the arrival of printing was widely perceived by the Church as an invention of Satan, the hierarchy fearing that the dissemination of uncensored ideas would lead to a breakdown of social order and put innumerable souls at risk of damnation. Steps were quickly taken to limit its potentially evil effects. Within half a century of Gutenberg's first Bible (1455), Frankfurt had established a state censorship office to suppress unorthodox biblical translations and tracts (1486), and soon after, Pope Alexander VI extended censorship to secular books (1501). Around 400 years later, similar concerns about censorship and control were widespread when society began to cope with the political consequences of the arrival of the telegraph, the telephone, and broadcasting technology. The telegraph would destroy the family and promote crime.[4] The telephone would undermine society. Broadcasting would be the voice of propaganda. In each case, the anxiety generated specifically linguistic controversy. Printing enabled vernacular translations of the Bible to be placed before thousands, adding fuel to an argument about the use of local languages in religious settings which continues to resonate today. And when broadcasting enabled selected voices to be heard by millions, there was an immediate debate over which norms to use as correct pronunciation, how to achieve clarity and intelligibility, and whether to permit local accents and dialects, which remains as lively a debate in the twenty-first century as it was in the twentieth.

The Internet is an association of computer networks with common standards which enable messages to be sent from any central

---

[4] The parallels between the arrival of the Internet and the arrival of the telegraph are explored in Standage (1999).

computer (or *host*) on one network to any host on any other. It developed in the 1960s in the USA as an experimental network which quickly grew to include military, federal, regional, university, business, and personal users. It is now the world's largest computer network, with over 100 million hosts connected by the year 2000, providing an increasing range of services and enabling unprecedented numbers of people to be in touch with each other through electronic mail (*e-mail*), discussion groups, and the provision of digital 'pages' on any topic. Functional information, such as electronic shopping, business data, advertisements, and bulletins, can be found alongside creative works, such as poems and scripts, with the availability of movies, TV programmes, and other kinds of entertainment steadily growing. Some commentators have likened the Internet to an amalgam of television, telephone, and conventional publishing, and the term *cyberspace* has been coined to capture the notion of a world of information present or possible in digital form (the *information superhighway*). The potential of the Internet is currently limited by relatively slow data-transmission speeds, and by the problems of management and retrieval posed by the existence of such a vast amount of information (see chapter 7); but there is no denying the unprecedented scale and significance of the Net, as a global medium. The extra significance is even reflected in the spelling, in languages which use capital letters: this is the first such technology to be conventionally identified with an initial capital. We do not give typographical enhancement to such developments as 'Printing', 'Publishing', 'Broadcasting', 'Radio', or 'Television', but we do write 'Internet' and 'Net'.[5]

What is it like to be a regular citizen of the Internet, a *netizen*? Those who already spend appreciable amounts of time online need

---

[5] In its sense as a global network of computers. When the term is used to refer to a local network, or some local set of connected networks, it is usually given a lower-case initial – though usage is uncertain in both contexts. The abbreviated form, Net, is generally capitalized. Private networks within organizations, or *intranets*, are always lower-case. It is important to note that other networks exist. A chatgroup system, such as the Usenet newsgroups (pp. 131–3), may be carried by other networks than the Internet (such as UUCP). Although the focus of this book is the Internet, its conclusions apply just as much to these other nets.

only self-reflect; for those who do not, the self-descriptions of a 'day in a netizen's life' are informative. Here is Shawn Wilbur's, as he describes what a 'virtual community' means to him:[6]

> For me it is the work of a few hours a day, carved up into minutes and carried on from before dawn until long after dark. I venture out onto the Net when I wake in the night, while coffee water boils, or bath water runs, between manuscript sections or student appointments. Or I keep a network connection open in the background while I do other work. Once or twice a day, I log on for longer periods of time, mostly to engage in more demanding realtime communication, but I find that is not enough. My friends and colleagues express similar needs for frequent connection, either in conversation or through the covetous looks they cast at occupied terminals in the office. Virtual community is this work, this immersion, and also the connections it represents. Sometimes it is realtime communication. More often it is asynchronous and mostly solitary, a sort of textual flirtation that only occasionally aims at any direct confrontation of voices or bodies.

And there are now several sites which will advise you of the symptoms to look out for if you want to know whether you are Internet-driven. Here is a short selection from various pages headed 'addicted to the Internet':

> You wake up at 3 a.m. to go to the bathroom and stop to check your e-mail on the way back to bed.
> You sign off and your screen says you were on for 3 days and 45 minutes.
> You placed the refrigerator beside your computer.
> You say 'scroll up' when someone asks what it was you said.
> All of your friends have an @ in their names.
> You tell the cab driver you live at http://123.elm.street/house/bluetrim.html
> You check your mail. It says 'no new messages'. So you check it again.
> Your phone bill comes to your doorstep in a box.

---

[6] Wilbur (1996: 13–14). See also Naughton's account (1999: 143ff.).

It is not the aim of this book to reflect on the consequences for individuals or for society of lives that are lived largely in cyberspace. My aim is much more modest: it is to explore the ways in which the nature of the electronic medium as such, along with the Internet's global scale and intensity of use, is having an effect on language in general, and on individual languages in particular. It seems likely that these effects will be as pervasive and momentous as in the case of the previous communication technologies, mentioned above, which gave language printed and broadcast dimensions that generated many new distinctive varieties and usages, from the telegrammatic graphic prominence of newspaper headlines to the hyperverbal sonic prominence of sports commentaries. The electronic medium, to begin with, presents us with a channel which facilitates and constrains our ability to communicate in ways that are fundamentally different from those found in other semiotic situations. Many of the expectations and practices which we associate with spoken and written language, as we shall see (chapter 2), no longer obtain. The first task is therefore to investigate the linguistic properties of the so-called 'electronic revolution', and to take a view on whether the way in which we use language on the Internet is becoming so different from our previous linguistic behaviour that it might genuinely be described as revolutionary.

The linguistic consequences of evolving a medium in which the whole world participates – at least in principle, once their countries' infrastructure and internal economy allow them to gain access – are also bound to be far-reaching. We must not overstate the global nature of the Internet: it is still largely in the hands of the better-off citizens of the developed countries. But it is the principle which matters. What happens, linguistically, when the members of the human race use a technology enabling any of them to be in routine contact with anyone else? There has been much talk of the notion of a 'global village', which is at first sight a persuasive metaphor. Yet such a concept raises all kinds of linguistic questions. A village is a close-knit community, traditionally identified by a local dialect or

language which distinguishes its members from those elsewhere: 'That's not how we say things round here.' If there is to be a genuine global village,[7] then we need to ask 'What is its dialect?', 'What are the shared features of language which give the world community of users their sense of identity?' And, if we cannot discern any unifying dialect or language, or a trend towards such a unity, we need to ask ourselves if this 'global village' is anything more than a media fiction. Similar questions might be asked of related notions, such as 'digital citizens', 'the virtual community', and the 'Net generation'. The linguistic perspective is a critical part of this debate. As Derek Foster puts it, reflecting on the notion of a virtual community, 'the fullest understanding of the term is gained by grounding it in the communicative act itself'.[8] So the second task is to investigate whether the Internet is emerging as a homogenous linguistic medium, whether it is a collection of distinct dialects, reflecting the different backgrounds, needs, purposes, and attitudes of its users, or whether it is an aggregation of trends and idiosyncratic usages which as yet defy classification.

Internet situations

In a setting where linguistic differences are likely to loom large, the concept of a *language variety* will be helpful. A variety of language is a system of linguistic expression whose use is governed by situational factors.[9] In its broadest sense, the notion includes speech and writing, regional and class dialects, occupational genres (such as legal and scientific language), creative linguistic expression (as

---

[7] McLuhan (1962: 31), and elsewhere.
[8] Foster (1996: 35).
[9] Within linguistics, several terms have been used, over the years, for talking about language which varies according to situation, such as *speech community*, *register*, *genre*, *text*, and *discourse type*, each of which operates in its own theoretical frame of reference (see Crystal and Davy, 1969). As Internet linguistics develops, more sophisticated models will be needed to capture all elements of the variation found. For the present book, which is only a 'first approximation', I have avoided a more complex terminological system, and used the term *variety* without further qualification for all kinds of situationally influenced language. I also sometimes refer to *genres* within a variety. Within the Internet literature, terminology also varies a great deal when discussing the different kinds of Internet situation, such as *environment*, *interactive setting*, and *virtual space*.

in literature), and a wide range of other styles of expression. Varieties are, in principle, systematic and predictable. It is possible to say, with some degree of certainty in a given language, how people from a particular region will speak, how lawyers will write, or how television commentators will present a type of sport. Notions such as 'British English' or 'Liverpool English', 'legal French', and 'sports commentary' are the result. To change an important element in any situation is to motivate a change in the language people use there, if they wish to behave conventionally – whether the change is from one region to another, from law court to the street, from home to pub, from one listener to many, or from face-to-face to distant conversation. Sometimes the features of a variety are highly constrained by the situation: there are strict rules governing the kind of language we may use in court, for example, and if we break them we are likely to be criticized or even charged with contempt. In other situations there may be an element of choice in what we say or write, as when we choose to adopt a formal or an informal tone in an after-dinner speech, or a combination of the two. But all language-using situations present us with constraints which we must be aware of and must obey if our contribution is to be judged acceptable. Factors such as politeness, interest, and intelligibility govern what we dare to introduce into an after-dinner speech, and such criteria apply in all situations. 'Anything goes' is never an option – or, at least, if people do decide to speak or write without paying any attention to the sociolinguistic expectations and mores of their interlocutors, and of the community as a whole, they must expect to be judged accordingly.[10]

The distinctive features of a language variety are of several kinds. Many stylistic approaches recognize five main types, for written language.[11]

- *graphic* features: the general presentation and organization of the written language, defined in terms of such factors as

---

[10] Allowances can sometimes be made – as with some kinds of psychiatric disturbance and linguistic pathology, or the utterances of very young children.
[11] For the application of a model of this kind to several varieties of English, see Crystal and Davy (1969).

distinctive typography, page design, spacing, use of illus-
trations, and colour; for example, the variety of newspaper
English would be chiefly identified at this level through the
use of such notions as headlines, columns, and captions.

- *orthographic* (or *graphological*) features: the writing system
  of an individual language, defined in terms of such factors as
  distinctive use of the alphabet, capital letters, spelling, punc-
  tuation, and ways of expressing emphasis (italics, boldface,
  etc.); for example, American and British English are distin-
  guished by many spelling differences (e.g. *colour* vs. *color*),
  and advertising English allows spelling modifications that
  would be excluded from most other varieties (e.g. *Beanz
  Meanz Heinz*).

- *grammatical* features: the many possibilities of syntax and
  morphology, defined in terms of such factors as the distinctive
  use of sentence structure, word order, and word inflections;
  for example, religious English makes use of an unusual
  vocative construction (*O God, who knows . . .*) and allows a
  second-person singular set of pronouns (*thou, thee, thine*).

- *lexical* features: the vocabulary of a language, defined in terms
  of the set of words and idioms given distinctive use within a
  variety; for example, legal English employs such expressions
  as *heretofore*, *easement*, and *alleged*, as well as such phrases as
  signed sealed and delivered and Latin expressions such as *ex
  post facto.*

- *discourse* features: the structural organization of a text,
  defined in terms of such factors as coherence, relevance,
  paragraph structure, and the logical progression of ideas;
  for example, a journal paper within scientific English ty-
  pically consists of a fixed sequence of sections including the
  abstract, introduction, methodology, results, discussion, and
  conclusion.

'Whatever else Internet culture may be, it is still largely a text-based
affair.'[12] Spoken language currently has only a limited presence on

---

[12] Wilbur (1996: 6).

the Internet, through the use of sound clips, films, and video; but the use of speech will undoubtedly grow as technology develops, and it will not be long before we see the routine use of interactive voice (and video) dialogues, speech synthesis to provide a spoken representation of what is on a screen or to give vocal support to a graphic presentation, and automatic speech recognition to enable users to interact verbally with sites (see further, chapter 8). In addition to the above five types, therefore, we need to recognize two more:

- *phonetic* features: the general auditory characteristics of spoken language, defined in terms of such factors as the distinctive use of voice quality, vocal register (e.g. tenor vs. bass), and voice modality (e.g. speaking, singing, chanting); for example, in TV commentary, different sports make use of different vocal norms (e.g. the loud enthusiastic crescendos of football vs. the hushed monastic tones of snooker).
- *phonological* features: the sound system of an individual language, defined in terms of such factors as the distinctive use of vowels, consonants, intonation, stress, and pause; for example, regional accents are defined by the way they make different use of sounds, and distinctive pronunciation is also a notable feature of such varieties as newsreading, preaching, and television advertising.

Grammatical, lexical, and discourse features of course play a distinctive role in all spoken varieties of a language, as they do in the written. A television commentary is not distinctive solely in its pronunciation, but in its use of grammar, vocabulary, and general organization as well.

So the initial question for the person interested in Internet linguistics to ask is: is the Net a homogenous language-using electronic situation, likely to generate a single variety of language, defined using such variables as those listed above? Will all users of the Internet present themselves, through their messages, contributions, and pages, with the same kind of graphic, orthographic, grammatical, lexical, and discourse features? To answer these questions we need first to establish how many different situations the Internet

contains. We then need to describe the salient linguistic features of each situation, and to identify variations in the way they are used. This will help us talk more precisely about the strategies that people employ and the linguistic attitudes they hold, and thus enable us to begin evaluating their beliefs and concerns about Internet language. Some of these situations are easy to identify, because they have been around a relatively long time and have begun to settle down. Some are still in their infancy, with their situational status totally bound up with emerging technology, and therefore subject to rapid change: an example is the linking of the Internet to mobile phone technology, where the small screen size immediately motivated a fresh range of linguistic expression (see p. 228). Given the speed of technological change, doubtless new situational variables will emerge which will make any attempt at classification quickly outdated. But, as of the beginning of 2001, it is possible to identify five broad Internet-using situations which are sufficiently different to mean that the language they contain is likely to be significantly distinctive.

Electronic mail (*e-mail*)

E-mail is the use of computer systems to transfer messages between users – now chiefly used to refer to messages sent between private mailboxes (as opposed to those posted to a chatgroup). Although it takes up only a relatively small domain of Internet 'space', by comparison with the billions of pages on the World Wide Web, it far exceeds the Web in terms of the number of daily individual transactions made. As John Naughton says, 'The Net was built on electronic mail. . . . It's the oil which lubricates the system.'[13] Today, for example, I called up pages on the Web three times but sent twenty e-mails. My contacts included family, friends, and colleagues, as well as a range of new and long-standing business associates. My incoming e-mails included several of these, along with a sporadic sampling of 'junk' mail from organizations that had got hold of

[13] Naughton (1999: 150).

my e-address, some of which had attachments that were indistin-
guishable from a Web page in their linguistic character. Many of
the messages, incoming and outgoing, varied greatly in length and
style. The diversity of e-mail contexts is immediately apparent. So
here, too, the chief issue must be to determine the linguistic coher-
ence of the situation. Do the requirements of immediate and rapid
e-messaging promote the use of certain linguistic features which
transcend its many variations in audience and purpose? Indeed,
can we generalize about the language of e-mail at all? This question
is addressed in chapter 4.

Chatgroups

Chatgroups are continuous discussions on a particular topic, or-
ganized in 'rooms' at particular Internet sites, in which computer
users interested in the topic can participate. There are two situa-
tions here, depending on whether the interaction takes place in real
time (*synchronous*) or in postponed time (*asynchronous*).

- In a synchronous situation, a user enters a chat room and
  joins an ongoing conversation in real time, sending named
  contributions which are inserted into a permanently scrolling
  screen along with the contributions from other participants.
  Internet Relay Chat (IRC) is an example of one of the main
  systems available to users, consisting of thousands of rooms
  dealing with different topics. Although most people enter just
  one room at a time, there is nothing to stop them opening
  more than one chat window and engaging in two or more
  conversations simultaneously, if they have the requisite cog-
  nitive and linguistic skills.
- In an asynchronous situation, the interactions are stored in
  some format, and made available to users upon demand, so
  that they can catch up with the discussion, or add to it, at any
  time – even after an appreciable period has passed. The
  bulletin boards, a popular feature of 1980s computer-
  mediated communication, are one example. The thousands

of *newsgroups* on Usenet, covering a vast number of topics, provide another. Another is the *mailing list*, such as LIST-SERV®, to which users subscribe, knowing that all messages sent in to the list will reach everyone on that list.

Some chatgroups are global, receiving contributions from any geographical location; some are local, restricted to a particular country or region. Some are moderated, in the hands of an owner or editor; others are uncontrolled, other than by internal forces (see p. 146). Although the chatgroup situation would seem, at first sight, to promote the use of a highly distinctive and consistent language variety, the different factors involved – especially the factor of synchronicity – make it likely that it will contain significant diversity. This question is addressed in chapter 5.

Virtual worlds

Virtual worlds are imaginary environments which people can enter to engage in text-based fantasy social interaction. From the early notion of a *MUD* (originally 'multi-user dungeon', a derivation from the 1970s role-playing adventure game 'Dungeons and Dragons'), several adventure genres developed, offering players the opportunity to experience imaginary and vividly described environments in which they adopt new identities, explore fantasy worlds, engage in novel exploits, and use their guises to interact with other participants. Many MUDs, while reliant on the use of a shared virtual space and role-playing identities, move away from the creation of adventure worlds – for example, constructing worlds within education or business contexts, or using them for elaborate chat sessions. As a result, the acronym is also glossed as 'multi-user domain' or 'multi-user dimension'. Later technological developments enabled multimedia elements to be added to this genre, sound and video functions supplementing or replacing text to enable participants to take up an on-screen visual presence as avatars (a term from Hindu mythology, referring to an incarnation of a deity in earthly form) in

what some commentators have called *metaworlds*.[14] A range of sub-genres, with differing emphases, technical options, and of course acronym-like names, now exists, such as MOOs (MUD, Object-Orientated), MUSHes, MUCKs, MUSEs, and TinyMUDs (p. 173). The linguistic possibilities, in such imagination-governed worlds, are plainly immense, but – as with all games – there need to be constraints guiding the play, without which the interactions would be chaotic. These will be addressed in chapter 6.

World Wide Web (*WWW*)

The World Wide Web is the full collection of all the computers linked to the Internet which hold documents that are mutually accessible through the use of a standard protocol (the HyperText Transfer Protocol, or HTTP),[15] usually abbreviated to *Web* or *W3* and, in site addresses, presented as the acronym *www*. The creator of the Web, computer scientist Tim Berners-Lee, has defined it as 'the universe of network-accessible information, an embodiment of human knowledge'.[16] It was devised in 1990 as a means of enabling high-energy physicists in different institutions to share information within their field, but it rapidly spread to other fields, and is now all-inclusive in subject-matter, and designed for multimedia interaction between computer users anywhere in the world. Its many functions include encyclopedic reference, archiving, cataloguing, 'Yellow Pages' listing, advertising, self-publishing, games, news reporting, creative writing, and commercial transactions of all kinds, with movies and other types of entertainment becoming increasingly available. With such an enormous range of topic and purpose, the chief linguistic issues

[14] For example, Wallace (1999: 8).
[15] A protocol is a set of rules which enables computers to communicate with each other or other devices; the Transmission Control Protocol / Internet Protocol, TCI/IP, was made the Internet standard in 1985; *Wired Style* calls it 'the mother tongue of the Internet' (Hale and Scanlon, 1999: 159).
[16] Berners-Lee (1999). It should be evident that the popular practice of using the terms *Internet* and *Web* interchangeably is very misleading. The Web is one of several Internet situations.

here must be whether the Web can be said to have any coherence, as a linguistic variety, and whether it is possible to make useful or valid generalizations about its use of language at all. This question is addressed in chapter 7.

These five situations are not entirely mutually exclusive. It is possible to find sites in which all elements are combined, or where one situation is used within another. For example, many Web sites contain discussion groups and e-mail links; e-mails often contain Web attachments; and some MUDs include asynchronous chatgroups and permit participants to contact each other via e-mail. The Internet world is an extremely fluid one, with users exploring its possibilities of expression, introducing fresh combinations of elements, and reacting to technological developments. It seems to be in a permanent state of transition, lacking precedent, struggling for standards, and searching for direction. About the only thing that is clear is that people are unclear about what is going to happen. As John Naughton puts it, at the end of his book, *A brief history of the future*, 'The openness of the Net also applies to its future. The protocols which govern it leave the course of its evolution open.'[17] For example, it is likely that my five situations will need to be supplemented very soon by a sixth, as interactive voice dialogue becomes increasingly available, and conversationalists make decisions about what kind of spoken language to use to exploit the new medium. But there is no way of predicting whether this new language-using situation will make use of old conversational norms or invent fresh stylistic techniques to facilitate interaction, or what particular combination of new and old will prove to be most effective. This will doubtless add an extra chapter to some later edition of this book.

For each of the five situations outlined above, it is evident that people are still getting to grips with the communicative potential made available to them. They are in a learning situation of a rather special kind. They are having to acquire the rules (of how

---

[17] Naughton (1999: 271).

to communicate via e-mail, of how to talk in chatgroups, of how to construct an effective Web page, of how to socialize in fantasy roles), and yet there are no rules, in the sense of universally agreed modes of behaviour established by generations of usage. There is a clear contrast with the world of paper-based communication. Letter-writing, for instance, is routinely taught in school; and because there is widespread agreement on how letters are to be written, supported by the recommendations of usage manuals, we feel secure in that knowledge. We know such conventions as how to use opening and closing formulae (*Dear Sir/Madam*, *Yours faithfully*), where to put the address and date, and how to break up the text into paragraphs. Adults make use of this knowledge almost without thinking, and on occasion, as in informal letter-writing, they dare to break the rules with confidence. But with the Internet equivalent of letter-writing – e-mails – there is no such long tradition. Most people have been using e-mails for less than a decade, and they are unaware of the factors which have to be respected if their messages are not to be misunderstood. Often, the first indication that they have misconstrued a message comes when they receive an unpalatable response from the recipient.

Nobody knows all the communicative problems which lurk within e-discourses of all five kinds. Recommendations about approach and style are only beginning to be formulated, and many are tentative (see chapter 2). Market research companies are investing a great deal to discover how people react to different Web page configurations. Psychologists are beginning to probe the kinds of problem which affect individuals who engage in unconstrained fantasy play. There is an enormous amount of idiosyncrasy and variation seen in e-encounters. At the same time, the detailed studies which have taken place have begun to identify levels of shared usage within individual e-situations. Lynn Cherny, for example, having studied the language found in one kind of MUD (ElseMOO, p. 174), concludes that 'the linguistic interactions in ElseMOO are most amenable to description in terms of register', and Boyd Davis and Jeutonne Brewer, in their study of a chatgroup, although initially tentative, conclude that it 'may come to be seen as a

register . . . [an] emergent register'.[18] Certainly the participants
themselves seem to be aware that their language is distinctive.
Cherny in fact reports an attempt by ElseMOO in 1994 to doc-
ument its distinctive language.[19] Although it did not get very far –
being criticized by some members as going against the 'insider'
ethos of the community – the argument suggests some clear intu-
itions about the status of its usage as a variety.

   The language of Internet users is plainly in a state of transition.
As Patricia Wallace puts it, in her discussion of the false impressions
Net participants gain about each other during encounters: 'On the
Internet we are struggling with a very odd set of tools and pushing
them as hard as we can. Homo sapiens are both set in their ways
and amazingly adaptable, and right now, all of us are learning some
painful and awkward lessons about impression formation online.'
And she adds: 'I look forward to the time when the kinds of "in-
teraction rituals" that Goffman described will stabilize on the net
and the business of forming impressions will be more predictable,
reliable, and familiar, and much less prone to those hazardous
misperceptions.'[20] The need for greater predictability, reliability,
and familiarity is something which affects all Internet situations,
and also the language which is found there. It is a world where
individuals have tried to solve the problem of an electronically
constrained communications medium (see chapter 2) in countless
idiosyncratic ways. It is also a world where many of the partici-
pants are highly motivated individualists, intent on exploring the
potential of a new medium, knowledgeable about its procedures,
and holding firm views about the way it should be used. The most
informed of this population are routinely referred to as *geeks* –
defined by *Wired Style*, an influential Internet manual, as 'someone
who codes for fun, speaks Unix among friends, and reads Slashdot
daily'.[21] We might expect a great deal of linguistic innovation and

[18] Cherny (1999: 27), Davis and Brewer (1997: 28–9, 157).
[19] Cherny (1999: 85). She introduces the relevant chapter with an epigraph from a character
    called Damon, who says, 'anyone who doesn't think we speak some strange separate dialect
    has been smoking crack'.
[20] Wallace (1999: 36); see, also, Goffman (1959).
[21] Hale and Scanlon (1999: 88). Slashdot is a Website created in 1997 to provide 'News

ingenuity in their usage, accordingly. At the same time, everyone is aware that too much idiosyncrasy causes problems of intelligibility. Also, the pressure towards conformity is strong in those participatory activities to which the label 'community' has often been applied. As one contributor to a discussion about aggressive language (*flaming*, p. 55) said: 'You and I can talk any way we want on Internet; the question is what kind of conversation are we looking for.'[22] So, what kind of conversations are there, online, and how does one participate in them? Do we have to learn a new kind of language – 'Netspeak', as I shall call it – in order to be a netizen?

Netspeak

The term 'Netspeak' is an alternative to 'Netlish', 'Weblish', 'Internet language', 'cyberspeak', 'electronic discourse', 'electronic language', 'interactive written discourse', 'computer-mediated communication' (CMC), and other more cumbersome locutions. Each term has a different implication: 'Netlish', for example, is plainly derived from 'English', and is of decreasing usefulness as the Net becomes more multilingual (p. 216); 'electronic discourse' emphasizes the interactive and dialogue elements; 'CMC' focuses on the medium itself. It is perhaps unsurprising to see 'Netspeak', as a term, being given some popular currency – following the Orwellian introduction of *Newspeak* and *Oldspeak* in *1984*, later developments such as *Doublespeak* and *Seaspeak*, and media labels such as *Royalspeak* and *Blairspeak*. From the perspective of this book, it is broader than *Webspeak*, which has also had some use. As a name, *Netspeak* is succinct, and functional enough, as long as we remember that 'speak' here involves writing as well as talking, and that any 'speak'

---

for Nerds. Stuff that Matters': <http://www.slashdot.com>. If you have just learned something from this footnote, you are not a geek.

[22] Millard (1996: 154–5). Other references which focus on the linguistic identity of various e-situations include: Ferrara, Brunner, and Whittemore (1991), Baym (1993), Maynor (1994), Collot and Belmore (1996), and Baron (1998b). The notion of 'virtual speech community' is encountered in various forms, such as 'discourse community' (Gurak, 1997).

suffix also has a receptive element, including 'listening and reading'. The first of these points hardly seems worth the reminder, given that the Internet is so clearly a predominantly written medium (for its spoken dimension, see chapter 8), and yet, as we shall see, the question of how speech is related to writing is at the heart of the matter. But the second point is sometimes ignored, so its acknowledgement is salutary. On the Internet, as with traditional[23] speaking and writing, the language that individuals produce is far exceeded by the language they receive; and as the Internet is a medium almost entirely dependent on reactions to written messages, awareness of audience must hold a primary place in any discussion. The core feature of the Internet is its real or potential interactivity.

There is a widely held intuition that some sort of Netspeak exists – a type of language displaying features that are unique to the Internet, and encountered in all the above situations, arising out of its character as a medium which is electronic, global, and interactive. The linguistic basis for this intuition is examined in detail in chapters 2 and 3; but the fact that people are conscious of something 'out there' is demonstrated by the way other varieties of language are being affected by it. It is always a sure sign that a new variety has 'arrived' when people in other linguistic situations start alluding to it. For example, a comic courtroom sketch on television will borrow freely from legal language, assuming that viewers will recognize the linguistic allusions; and individuals can introduce references to legal language into their speech even if they have never been inside a courtroom in their lives – 'the tooth, the whole tooth, and nothing but the tooth' was one particularly bad dental pun I encountered recently. It is therefore of considerable

---

[23] The terms 'traditional' and 'conventional' are often used to refer to non-electronically mediated linguistic communication – old-style speech and writing – but there is no standard usage. More generally, there is no standard terminology for the distinction between the electronic and non-electronic worlds – though commonly used is the opposition *VR* ('virtual reality') and *RL* ('real life') or the adverbial *IRL* ('in real life'), the 'physical world', and other such locutions. Ihnatko (1997: 160) defines 'real world' as 'That which cannot be accessed via a keyboard. A nice place to visit, a good place to swing by when you're out of Coke, but you wouldn't want to live there.'

interest to note the way in which salient features of Netspeak, taken from one or other of its situational manifestations, have already begun to be used outside of the situation of computer-mediated communication, even though the medium has become available to most people only in the past decade or so. The influence is mainly on vocabulary, with graphology affected in some written varieties.[24]

   In everyday conversation, terms from the underlying computer technology are given a new application among people who want their talk to have a cool cutting-edge. Examples from recent overheard conversations include:

> It's my turn to download now (i.e. I've heard all your gossip, now hear mine)
> I need more bandwidth to handle that point (i.e. I can't take it all in at once)
> She's multitasking (said of someone doing two things at once)
> Let's go offline for a few minutes (i.e. let's talk in private)
> Give me a brain dump on that (i.e. tell me all you know)
> I'll ping you later (i.e. get in touch to see if you're around)
> He's 404 (i.e. he's not around; see p. 82)
> He started flaming me for no reason at all (i.e. shouting at me; see p. 55)
> That's an alt.dot way of looking at things (i.e. a cool way; see p. 83)
> Are you wired? (i.e. ready to handle this)
> Get with the programme (i.e. keep up)
> I got a pile of spam in the post today (i.e. junk-mail; see p. 53)
> He's living in hypertext (i.e. he's got a lot to hide; see p. 202)
> E you later (said as a farewell)

Programmers have long needed special vocabulary to talk about their lines of code, and some of this has now spilled over into

---

[24] An interesting influence occurs in those languages, such as Spanish and Portuguese, which lack the letter *w*, and where the existence of *WWW* in effect adds an extra letter to their alphabet. The influence of English on the vocabulary of other languages is also growing, such as *hack* and *scroll* (as verbs in Dutch), *scrollare* and *deletare* (Italian), *debugear* and *lockear* (Spanish).

everyday speech, especially to handle the punctuation present in an electronic address. For example, radio and television presenters commonly add e-addresses when telling listeners and viewers how they might write in to a programme, using *at*, *dot*, and *forward slash* to punctuate their utterance. *Dot com* is now a commonly heard phrase, as well as appearing ubiquitously in writing in all kinds of advertising and promotional material.

   In fact, written English shows developments well beyond the stage of the literal use of *.com*. This suffix is one of several domain names (with some US/UK variation) showing what kind of organization an electronic address belongs to:[25] *.com* (commercial), *.edu* or *.ac* (educational), *.gov* (governmental), *.mil* (military), *.net* (network organizations), and *.org* or *.co* (everything else). *Dotcom* has come to be used as a general adjective (with or without the period, and sometimes hyphenated), as in *dotcom organizations* and *dotcom crisis.* It has, however, come to be used in a variety of ludic ways, especially in those varieties where language play is a dominant motif – newspaper headlines and advertising.[26] It has been expanded into other words: a computer hardware store advertises itself as *SHOPNAME.computer.* Similarly, *www* became *web without worry* in a British Telecom advertising campaign. The similarity of *com* to *come* has been noticed, and doubtless there are similar links made in other languages. An offer to win a car on the Internet is headed *.com and get it.* A headline in the *Independent Graduate* on openings still available on the Web is headed: *Dot.com all ye faithful.* A phonetic similarity motivated a food-outlet advertisement: *lunch@Boots.yum.* The 'dot' element is now introduced into all kinds of phrases: *Learnhow.to* and *launch.anything*, are names of sites. The phrase *un.complicated* introduced an ad for personal finance. One company uses the slogan *Get around the www.orld*; another has the slogan *www.alk this way.*

[25] As of 2000. Other domain names are under consideration, such as .rec and .shop, allocated by such organizations as Network Solutions in the USA until 2000 and Nominet in the UK; the US role was taken over by Internet Corporation for Assigned Names and Numbers, established in 1998.

[26] Crystal (1998). Interestingly, when *dot.com* is written with a period, as here, the punctuation mark is never spoken aloud: we do not say 'dot dot com'.

A similar ludic trend applies to the symbol @, now the universal link between recipient and address. It was chosen pragmatically by a computer engineer, Ray Tomlinson, who sent the first network e-mail in 1972. He needed a character which did not occur in names, and this typewriter keyboard symbol stood out, with the bonus of having an appropriate meaning (of someone being 'at' somewhere).[27] A subsequent irony is that many firms and organizations have replaced the letter *a* or *at* in their name by an @: *@llgood*, *@tractions*, *@cafe*, *@Home*, *@pex*. And it has been seen turning up in other settings where traditionally the word *at* would be used: *This is where it's @* is one slogan; Bill Gates' 1999 book is called *Business @ the speed of thought*; and an academic article concludes a review of the interaction between literary and everyday language through the device *language @ literature* and *literature @ language*.[28] It has even been added to text where the word *at* would not normally appear – a postcard to my house read: *Crystals @. . .* followed by the address.

By now the *e*-prefix must have been used in hundreds of expressions. *The Oxford dictionary of new words* (1997)[29] had already noted *e-text*, *e-zine*, *e-cash*, and *e-money*, and in 1998 the American Dialect Society named *e-* 'Word [*sic*] of the Year' as well as 'Most Useful and Most Likely to Succeed'. Examples since noted include *e-tailing* and *e-tailers* ['retailing on the Internet'], *e-lance* ['electronic free-lance'] and *e-lancers*, *e-therapy* and *e-therapists*, *e-management* and *e-managers*, *e-government*, *e-bandwagon*, *e-books*, *e-conferences*, *e-voting*, *e-loan*, *e-newsletters*, *e-security*, *e-cards*, *e-pinions*, *e-shop*, *e-list*, *e-rage*, *e-crap*, and (Spanish) *e-moción*. Awareness of the form, though in the reverse direction, appeared on the side of a London taxi: *Watrloo No Problm* – glossed beneath by *no-e.anything*. A bookmaker developing a Net presence called the firm *e-we go*. Journalistic headlines and captions often

---

[27] Though some languages have borrowed the English word 'at' for this symbol, several have their own name for it: for example, @ is a 'snail' in Italian, a 'little mouse' in Chinese, an 'elephant's trunk' in Swedish, a 'worm' in Hungarian, and a 'spider monkey' in German.
[28] Crystal (1999).
[29] Knowles (1997).

play with terms in search of eye-catching effects, so it is not surprising to find e-motivated lexical formations in specialist newspapers and magazines, as well as in the general press. Examples include:

> MAJOR BREAKTHROUGH IN SEARCHITIS
> STOP INTERNET CLICKTOSIS
> Dealing with the dot.com Brain Drain
> The Geekicon (headline of an *Economist* review of a computer dictionary)

How many of these developments will become a permanent feature of the language it is impossible to say. We can never predict language change, only recognize it once it has happened. There are already signs of a reaction against some of the above usages. The authors of *Wired Style*, for example, beg, in relation to the use of *e-*: 'Please, resist the urge to use this vowel-as-cliche', citing such 'too-facile coinages' as *e-lapse*, *e-merge*, and *e-quip*.[30] A Silicon Valley company, Persistence Software, is reported to have established The Society for the Preservation of the Other 25 Letters of the Alphabet, in order to campaign against the proliferation of e-words. There have been similar complaints about the use of *dot.com* in advertising. A United States company-names specialist, Neil Cohen, is quoted as saying (in mid-2000), 'Using "e", "i", and ".com" will make the company seem like a dinosaur even five years from now.'[31] But this only makes the general hypothesis more compelling, that a notion of Netspeak has begun to evolve which is rapidly becoming a part of popular linguistic consciousness, and evoking strong language attitudes. The next step, accordingly, is to determine what its chief linguistic properties are. If Netspeak exists, the above examples will prove to be pointing to the tip of a large iceberg. Moreover, there will prove to be more fundamental linguistic strategies at work than these anecdotal illustrations suggest. If, then, people are worried about the effect of the Internet

---

[30] Hale and Scanlon (1999: 76).
[31] In *Language International*, 12 (4), August 2000, 48. See also Koizumi (2000), who reports that in 1999 the Japanese Patent Bureau accepted 50 names starting with *i-* (prompted by such names as *iMac* and *ipaq*) and 190 with *e-*.

on language in general and on their own language in particular –
as the quotations at the beginning of this chapter suggest – a first
step is to explore Netspeak in its various situational manifestations
to see what actually happens there. As John Paolillo puts it, in his
introduction to a paper on the virtual speech community:[32] 'If we
are to understand truly how the Internet might shape our language,
then it is essential that we seek to understand how different varieties
of language are used on the Internet.' Chapters 4–7, accordingly,
investigate the kind of language used in each of the five situations
described above. But all five have certain linguistic properties in
common, and these form the subject-matter of chapters 2 and 3.

[32] Paolillo (1999).