Teaching and learning the reading of homepages

JENNIFER THURSTUN – Macquarie University

ABSTRACT
Increasingly, ESL teachers are incorporating the World Wide Web (Web) and the Internet into their teaching, taking advantage of its seemingly boundless capacity to present language and information. This article will briefly consider the rationale for using the Web in language teaching, and then focus on ways to facilitate the teaching of, and learning about, the reading of homepages, which is the entry point for Web reading. I suggest that homepages be considered a text type, and that an understanding of the structure and characteristics of this text type can help teachers and students in their approach to the reading of homepages so they can use the Web more effectively.

The Web in the classroom
In justifying the use of the Web in teaching, we should keep in mind that new information technologies in education are driven largely by commercial rather than educational interests. In some ways, the information superhighway is ‘a solution in search of a problem’ (Kenway 1995: 40), and education is targeted as a promising application. The so-called experts in educational software are often also marketing directors. Bigum (2002) notes that computing and communication technology has been promoted for education by vendors since the late 1970s, with promises of improved learning and employment opportunities and access to boundless information. Although such claims have yet to be proven, they tend to be taken for granted by educators and parents, as well as by policy makers. The computer and the Internet are facts of life in the classroom, and education is undergoing a huge and significant technological revolution. A great deal of experimentation is taking place and there are many challenges to be faced both by teachers and learners. A degree of caution is appropriate – teachers need to be sure that the technology serves both their purposes and those of their students, rather than the reverse.

Leading TESOL teachers (Hawisher and Selfe 1989; Warschauer 1999) have stressed the benefits of including computer and Internet use in the language classroom. TESOL courses offer instruction in computer use for language teaching (Kamhi-Stein et al 2002), and teachers and teacher trainers
in both the first and second language environment are responding to government policies that link literacy with technology and the Web. There are suggestions that failure to teach students Web literacy may result in their disadvantage within the global literacy community (Sutherland-Smith 2002), and that the literacies of the new technologies are a subset of those needed in a learner’s personal, work and community life (McPherson and Murray 2003). There is no doubt that teachers and learners need to be involved in learning about and using the Web, since it is an undeniable and increasingly important part of our world, and because ‘literacy involves gaining the skills and knowledge to read and interpret the text of the world and to successfully navigate and negotiate its challenges, conflicts, and crises’ (Kellner 2002: 157).

Successful use of the Internet and the Web, however, requires the development of special skills and understandings. Electronic text differs radically from print text (Thurstun 2000) and creates quite different challenges both for the reader and the teacher of reading. According to Snyder (2002: 3), ‘People have to learn to make sense of the iconic systems evident in computer displays – with all the combinations of signs, symbols, pictures, words and sounds’.

The Internet began as an English-language phenomenon, and English continues to dominate cyberspace (Warschauer 2002). So for speakers of English as a second or foreign language, the Internet experience combines the dual challenge of understanding not only text but also the conventions and symbols that have developed in the English-speaking world for presenting information via the computer, the Internet and the Web. The cyberlibrary of information made accessible through the Web is made up of websites, and the first experience of a particular website is usually its homepage.

Despite the anarchic nature of the Internet, and the constant change and experimentation occurring in cyberspace, the design and structure of the webpage, particularly the homepage, has consolidated in recent years. Usability studies, largely led by American usability guru Jakob Nielsen of Sun Microsystems, that have tested the ease or difficulty with which users read and negotiate websites have influenced the design of sites. Advice given to website designers (Nielsen 2000) indicates the emergence of a preferred structure for homepages that have clear design, technical and linguistic characteristics. This paper will describe these characteristics and suggest that the evolution of the homepage has consolidated to the extent that it can be considered a text type, and that such an insight can facilitate the teaching and learning of this aspect of online literacy. It will also look at the research on the way readers read webpages, and make suggestions to teachers about
choosing websites for student use and formulating strategies for effective ‘reading’ of webpages.

**Structure of homepages**

This article focuses on the reading of homepages. The term ‘homepages’ here refers to the external pages of a site, the pages that provide a map of the website and which are the user’s introduction to the site (not the more internal pages, which may not be text at all but pictures or sound clips). The term ‘text type’ refers to a discourse reflecting typical organisational patterns, characteristics and purposes (Johns 2002). It is suggested here that the homepage can, by and large, be considered a text type, and that by describing this text type in broad terms it is hoped that learners can be helped to anticipate its characteristics and use it more effectively.

The Internet is a moving, changing, evolving phenomenon, and the teaching and learning of the skills needed for handling it must also allow for a certain amount of flexibility and openness to the unexpected. However, there is a solid basis of research that indicates which design elements improve and hinder the readability and use of websites, and this enables us to make clear statements about the structure and characteristics found in most homepages. Teachers can use this research both to make informed choices about the websites they recommend for student use, and to help students understand and anticipate the characteristics of the text type, with a view to increasing the efficiency with which they read and use it.

The typical webpage referred to by Nielsen (2000) has evolved with a basic structure that tends to cross borders of language and culture; the homepages of the World Bank (http://www.worldbank.org) or Lonely Planet (www.lonelyplanet.com) are good examples. The page tends to be designed in a two- or three-column format, with a heading and perhaps a menu set out in a row at the top. In a three-column format the central column is likely to contain most detail and content, the left column usually serves as a table of contents and the right column can be expected to provide functional links (such as print, subscribe, download PDF file). Where there are only two columns, the functional links (if there are any) might be incorporated into either one of the two columns.

Not all homepages follow this format: some designers seem to have limited awareness of these ‘conventions’ and the advice of usability studies, and some experimental sites (music, film, art, literary sites, for example) purposely break the mould. However, Nielsen (1998) considers the established approach, despite its drawbacks (http://www.alertbox/980322.html), the most usable format, largely because users have come to expect this design.
Nielsen’s (1998) research shows that users are very conservative and, having become accustomed to the established design, reject innovations and advanced design concepts. He considers that, since users frequently move back and forth between pages and between sites,

the entire corpus of the Web constitutes a single interwoven user experience …

The Web as a whole is the foundation for the user interface, and any individual site is nothing but a speck in the Web universe

(see website, Nielsen 1998).

His studies show that users expect webpages to work as they do elsewhere. He therefore sees the value of webpages conforming to the particular structure that has evolved, and considers that designers should only attempt to bring in innovations gradually. The webpage has, it would seem, established itself as a text type, albeit one set within the limited confines of a screen but one which provides access through links to the vastness of cyberspace. In a study of 50 commercial homepages, Nielsen (2003) found that they were used as follows:

19% for operating system and browser overhead
20% for navigation
20% for content of interest to users
2% for advertisements
9% for self-promotions
5% for welcome, logo, tagline, other site identification
5% for filler (useless stock art)
20% unused

Of these, navigation and content are the only categories that are obviously useful and/or interesting to users, while the rest contain ‘fluff’ that users tend to ignore (promotions, site identity, filler clipart, and unused pixels as opposed to white space or separators). Navigation and content, the parts that are useful to users, account for only 40 per cent of space. It is, therefore, important to direct students’ attention to this 40 per cent, and help them realise that the other areas are not particularly useful.

The area containing content is usually the central column in a three-column homepage, or the right column in a two-column homepage. It is usually text based and contains links. The area for navigation can be scattered – to the right, left, above and/or below – but is generally found in the left-hand column (like a table of contents for the site), and is sometimes repeated
as a menu at the top of the page. It is made up of single words or groups of words or icons, all of which are clickable and lead to another page as (hopefully) indicated by the words or icons.

The areas that are not particularly useful, and which students should be taught to identify and ignore, are advertisements, self-promotions and the browser overhead. With a clearer idea of where the most and least useful information is likely to be found on the screen, students can save themselves time and frustration and enter an information-finding task with increased confidence.

**Technical characteristics of webpages**

Webpages have technical features that students need to be able to use. Most obvious among these are windows, scrolling and links. Teachers would also benefit from an awareness of the way websites are organised so they can choose websites that use manageable organisational structures.

**Windows** are an important characteristic of webpages. They are described by Abbey (2000) as areas of the screen that provide a particular view, and include palettes, icons, buttons and tools. They are used to represent or display information, and can expand the amount of information available. Multiple pages and frames can exist within a single page, creating visual complexity and the potential for information overload due to a combination of browsing and complex visualisation – particularly if the user has not developed an adequate mental map of the knowledge being presented.

**Scrolling** (in English language websites marked by the arrows column on the right side of a window and also sometimes below it) is the system used on a webpage (or a window within a webpage) to allow the user to see what is beyond the limits of the screen. Although the metaphor for scrolling originates from even further back in history than the development of print (Corbel 1997), it has been considered a problem in screen reading. Users reported that they became lost and felt uncomfortable when scrolling through long documents (Thurstun 2000), and that, according to Spool, Scanlon, Schroeder, Snyder and DeAngelo (1999), they disliked horizontal scrolling even more than vertical scrolling.

**Links** are extremely important. If they appear in text, they are usually presented in blue text that is underlined. They establish pathways of movement within the Web space and permit access to specific information. However, they can cause frustration and waste time unless they have been thoughtfully signalled, work properly and lead to useful destinations. The one graphic element of links that Spool et al (1999) found to be helpful is the change in colour of the blue underline link to purple to show where the user has already gone. Nielsen (2000), however, while agreeing that this
design element is useful considers having the colour blue for hypertext links as ‘the mother of bad web design conventions’, since blue is a colour known to reduce readability.

The organisation of the material that the webpage leads into is also important. The *Style manual for authors, editors and printers* (Commonwealth of Australia 2002) lists the following organisational structures:

- sequential for small electronic documents;
- hierarchical for more detailed material; and
- Web-like structures arranged in associative patterns, trying to offer access to everywhere from everywhere within a site.

Ideally, websites chosen for use by students should not involve more than three clicks in order to find information (Whitbread 2001), so the teacher is well advised to avoid exposing students to websites with complex, confusing structures.

There is good reason for ESL teachers to try to direct students to websites that follow the conventions and advice about design given by researchers like Nielsen. As Abbey (2000) observes, cognitive load refers to the demands on the learner’s working memory during instruction, and a complex or unconventionally designed screen which uses different fonts, objects, navigation tools and layout patterns will generally have a high procedural or functional cognitive load because each component will need to be perceived and interpreted by the learner. A screen which uses standard conventions in text, graphics, navigation and layout will be more easily interpreted and consequently have a much lower cognitive load. (Abbey 2000: 51)

The user experience reported here also emerges in the aforementioned usability studies and design recommendations, and these have led to the consolidation of the homepage into what can be considered a text type.

**Linguistic and textual characteristics of homepages affecting reading**

Like the language of any other text type, the language of homepages and of cyberspace in general has its own characteristics. Text is presented and formatted in a particular way to facilitate screen reading and the finding of information linked to the homepage, and the vocabulary of the new technology reflects its inventors and its purposes. According to Gibbs (2000: 23),

[a] disadvantage of cyberlanguage is that it may be perceived to offer a running and superficially attractive undermining of educational attempts to foster generally accepted decorums of linguistic convention.
The ESL teacher is quite likely to find that conventional advice about reading print text does not always apply to the reading of electronic text and webpages. So it could well be a matter of teaching new skills to enable students to feel comfortable in the vast ocean of language and information made available to them via the Web.

Characteristics of the language used in homepages (and often internal pages of a site) include the ‘chunking’ of information, special conventions for indicating stress, increased lexical density and nominalisation. Page layout, structure for accessing information and vocabulary are also influenced by electronic delivery, and icons frequently replace text altogether.

Text needs to be presented differently on screen from the way it is presented in print form, since the reading skills required for online reading are different. Recognising the difficulties of online reading (Thurstun 2000), writers for the screen tend to minimise text. The attitude of multimedia designers to text can be summed up in Bolter’s claim that

“[n]ullmedia designers consider it an admission of failure to clutter the screen with blocks of text. The worst criticism one can make of a multimedia system is to call it a mere ‘page turner’, a set of texts that the user examines one after another – in other words, the electronic equivalent of a book”

(Bolter 1996:106).

Nielsen (1997a) considers that readability is increased by about 58 per cent by halving the word count of an online document. A study carried out by Morkes and Nielsen (1998) claims that for online texts, halving the word count improves readability by 58 per cent, and a scannable layout with bulleted lists makes them 47 per cent more readable.

So text becomes denser online and information is often ‘chunked’, for example in dot point form. Dot (bullet) points appear more frequently and nominalisation can be expected to be used more heavily than in an equivalent print text. Surprisingly, Spool et al (1999) found that text considered difficult to read in print is more useful to onscreen readers than text normally considered less difficult, perhaps because high lexical density increases scannability and screen readers tend to scan rather than read in-depth (Nielsen 1997a).

For this reason there is strong agreement among designers and usability researchers on the need to break print content down into smaller, logical units when organising material for on-screen use. This ‘chunked’ text is often very spare and lexically dense, since writers of electronic text take pride in presenting meaning with as few words as possible. One of the most important skills of the Internet writer is considered to be the skill of reducing text while retaining meaning (A’Herran 1997), and chunking facilitates this.
The arrangement of text-based material on the screen is important for the ease with which the user can deal with it, and teachers and students should be aware of the preferred conventions for directing attention to information. The *Style manual for authors, editors and printers* (Commonwealth of Australia 2002) recommends signposting relevant material by the use of bridging words between sections of texts, cross-references or hyperlinks. It also recommends visually highlighting material by the use of: headings (which map the document’s structure, show the reader where to find information, group information into readily comprehensible chunks and indicate what follows); lists (because bullets draw attention to elements of a discussion but they need to be prefaced with a lead-in phrase); the use of bold or coloured keywords to improve scannability (in preference to italics, which are difficult to read onscreen, or underlining that indicates links); and the placement of text in boxes. All these are considered desirable organisational structures aimed at making the scanning of text easier.

Stress in text is normally signalled through the use of italics, but on-screen italics are not very readable. Whitbread (2001) recommends the use of colour variation and underlining (though this is problematic if it is confused with indication of linking), and suggests that bold should be reserved for headings.

The page layout on homepages is also worthy of attention. Apart from the typical two- or three-column format mentioned earlier, there are other elements of page layouts that can make scanning and searching easier. Nielsen (1999) makes the following recommendations for the enhancement of readability, including navigation:

- Name and logo placed on every page of a site, with the logo linking to the homepage.
- Straightforward, simple headlines and page titles.
- Groupings and subheadings to break long lists into smaller units (to facilitate scanning).
- Avoidance of clutter and cramming.
- Use of link titles so that users know where each link will take them. (Adapted from website, Nielsen 1999).

Learning to anticipate these conventions for presenting text and facilitating navigation should help students.

Awareness of the organisational structure of a particular website, which is likely to be reflected in the homepage, is also helpful both to teachers and
students. There are two different types of design for accessing information: exact access schemes, where information is alphabetical, chronological, sequential or geographical – as in dictionaries, radio guides, instruction manuals, atlases; and ambiguous access schemes, where information is divided into topics and the reader is free to search in a different way to that in which the author has structured the material (Commonwealth of Australia 2002: 444). Designed to suit the presentation of different types of information, students can be helped by learning about the two different structures so they can predict the one they are likely to be dealing with for a particular topic.

The vocabulary of cyberspace is a growing area of research. Morkes and Nielsen (1998) found that the use of objective rather than subjective language makes a text 27 per cent more readable, and Gibbs (2000) draws attention to the ‘maleness’ of cyberlanguage, much of which has been coined by males with a scientific orientation. Terms with an undercurrent of violence or maleness – such as *abort, execute, crash, firewall, zap, hit* – also frequently appear on the screen.

Another area of vocabulary that may need clarification for L2 students is the likelihood of needing to deal with the language of youth culture (not necessarily found in dictionaries), which tends to find its way into cyberspace. It may well be that youth culture is influencing the language of webpages, just as it has influenced (or almost created) the language of electronic communication. Wong (2000) points out that cyberspace is an ‘adolescent’ space – hence the likelihood of encountering cartoon culture and adolescent language. She suggests that a more phonetic way of writing may develop. Text messaging may also be influencing the language of the Web, along with conventions used in email and chat rooms. These conventions evolve and change rapidly, particularly with many abbreviated forms including acronyms (IMHO), abbreviations as in SMS messaging (*where r u from?*), emoticons, deliberately informal spellings and lack of punctuation (*dunno*) (Gibbs 2000). It is possible that this is influencing the language of webpages.

There are also issues of transparency of vocabulary and presumed meanings of graphics (McLean 2000): buttons like *Show, Go get it, Find it* are transparent, but what of *Wired, HotMetal* and *Electric Wizard* (McLean 2000).

Icons frequently replace text on homepages and webpages in general, particularly to indicate functions and actions that the user can perform. Kress (1995: 1) suggests that technology is ‘taking us both backwards and forwards into a new era of iconic forms of communication, backwards and forwards into hieroglyphics’. He considers that this move to icons is
related to an increasing awareness of the problems of using language alone as a form of communication in heterogeneous and multicultural societies. It also seems to be a development imposed by the nature of the technology itself. However, icons that are easily identifiable in one culture might not be so transparent in another, and ESL students may need clarification. Ganderton (1998), for example, studied Australian students using French websites in a French language classroom and found that icons were not helpful in the learning of vocabulary. One participant stated that they were ‘hard to understand if you didn’t know the French words. We couldn’t really pick what it was from the picture.’

**How readers read websites**

There seems to be conflicting opinion and evidence about the way readers read websites, perhaps because it is such a new area of research. This research is on native speakers of English and has ramifications for design (which tends to be transferred from English to other language websites), but we can extrapolate from it in the hope that it may help teachers advise their students about approaches to the reading of homepages.

It has been assumed (Whitbread 2001; Thurstun 2002) that the eye starts ‘reading’ a webpage by going first to any moving image, then to any still images or graphics on the screen. Only then, according to this theory, does the eye go to text, starting around the top centre of the screen, and then scanning with a left to right pattern as the user judges the likelihood of each of the three areas (left, right and central columns) containing useful information. It has been thought that the area at the bottom of the frame is often neglected. However, doubts are being cast on some aspects of this approach.

Although designers have sought to attract the attention of the user through graphics, plus all the bells and whistles made possible by computer technology, evidence is coming to light that graphics do not actually increase the usability of the site, and that users direct considerably less attention to them than to text. The Eyetrack Study carried out by the Poynter Institute and Stanford University in 2000 tracked the eye movements of subjects as they viewed their favourite news sites, and found that graphics actually attracted very little attention from readers. The following shows the attention given to different aspects of a page (see http://www.poynterextra.org/et/i.htm and click on to text, most likely):

- **Text articles**: 92% were looked at
- **Briefs**: 82%
Photos 64%
Banner ads 45%
Graphics 22%

It should be kept in mind that the Eyetrack Study (which is ongoing) was investigating news sites, and subjects were logging on to an accustomed site.

There is general agreement that readers scan rather than read in-depth, particularly on a homepage (Nielsen 1997a). Abbey (2000) describes the browsing experience as one in which the users may not read the text at all, but skim and scan, looking for comprehensible key words and headings. Once they have found the information they are looking for they are more likely to read both in-depth and carefully from beginning to end. However, screen reading is considered 25 per cent slower than print reading (Nielsen 1997b). The degree of domain knowledge the user brings to a site is thought to be critical to the amount of textual and contextual detail preferred: those with high domain knowledge prefer higher density information screens, while those with little domain knowledge would rather have less information and more explanations. The Web supports browsing, which involves skimming rather than reading text, rapid visual search and selection of buttons or hyperlinks, and impatience to move on to the next page.

Spool et al (1999) show that, for L1 users at least, greater lexical density and text difficulty make it easier to find on-screen information. Using readability scales – for example Flesch Reading Ease, the Gunning Fog Index and the Flesch-Kincaid Grade Level – they tested how successful users considered sites to be and found the opposite results for print and screen (reading). The Gunning-Fog Index measures the average number of words, sentences and words of more than three syllables in a text. The authors found that the higher (ie more unreadable) the index, the more successful the site proved to be. Users found sites with a high index to be more authoritative, clear, complete, satisfying, useful and less overwhelming, while they found the more ‘readable’ sites to be too wordy. The authors consider these scales to be measuring something that assists in information finding, possibly by making materials easier to skim and scan.

Hard-to-read sites have fewer conjunctions and standard grammatical structures. Since all the elements on a screen are competing for the user’s attention, it makes sense that the extra words we normally use to complete a sentence may get in the way of the meaningful words in the text. Removing them, and thus making the text harder to read by traditional measures, may increase scanability. (Spool et al 1999: 74)
It is the function words that are often omitted in chunked, online text, and these are part of the redundancy built into language. This redundancy allows for interference in processing (e.g., over the telephone) and safeguards communication of the intended message. When these are left out, language becomes more dense. This densification of text, however, is a particular challenge to ESL learners. They therefore need to be able to adjust to text that is quite different to the print texts they are expected to be able to read. In homepages, they are likely to find lexically dense text and condensed sentences. For the teacher, this suggests that special attention to scanning skills may be needed.

The lack of predictable patterns or schema is considered a major challenge for website users, and lack of familiarity with English can be assumed to exacerbate these difficulties for learners of English. Hypertext allows the user to change the structure of most texts by clicking on links, so any structure or schema that the writer may try to establish is likely to be changed, leaving the user without the structural patterns that might identify its genre and predict its development. Charney (1994) claims that the greater control offered by hypertext actually places pressure on the user. This is particularly the case when there are too many links and the structure is confusing (Tuman 1992). The fact that the text lacks fixed boundaries is probably the greatest challenge for the reader trying to get a sense of the text (Thurstun 2000). Castelli, Colazzo and Molinari (1998) argue that cognitive abilities are related to efficient use of hypertext, but even the best cognitive performers in their study (that is, those with well-integrated analytical and synthetic skills) could only handle a limited number of links and clearly connected pathways. Castelli et al (1998) also suggest there are gender differences between users. Females gave more time and attention to texts, concentrated more on what they were searching for and showed less distraction. Males, however, were more able to experiment with and control spatial exploration and performed better on navigation.

Finding information on the Web can be difficult regardless of gender, language background and even age. However, if designers provide useful content in a format that works in the way people expect it to, difficulties can be minimised. Advice from Spool et al (1999: 33, adapted below) on the basis of a detailed usability study of nine sites includes the following points, which teachers might do well to keep in mind when selecting websites for use by students:

- When users can predict where a link will lead they are more successful in finding information.
• The success of a link depends on how well users can differentiate one link from another (and this depends on spacing and whether the aim of the link is unambiguous). Users need to be able to eliminate links that seem obviously wrong for their search.

• Problems with navigation occur when users do not have the domain (content, business area) knowledge needed to navigate a site.

• Ambiguous terms used in links (e.g. global vs international) means that users do not know they are in the wrong place.

• There appears to be a negative correlation between the number of links and successful navigation.

• Links embedded in text do not work so well, and wrapped links are confusing.

• Links outside to other sites can be perplexing if users cannot get back home.

• Making comparisons is difficult because of the need to jump between sites and keep several in memory.

**Implications for the teacher and learner**

Usability studies and research on the way users read websites have influenced website designers to produce homepages with structures and characteristics suited to the medium, the purpose and the user. This, combined with user reluctance to change the way they ‘read’ from one website to another, is producing enough predictability to allow homepages to be considered (with some flexibility) a text type. Homepages have come to have their own specific characteristics: their own structure, technical, linguistic and textual features. This paper has outlined the need for particular strategies for the teaching of this text type.

The basic message to teachers is that students can be helped to understand this text type and thus anticipate its structure and certain characteristics, and that this will facilitate their reading and use of the Web. It is possible for the informed teacher involved in Web-based teaching to select sites that follow the recommended conventions and which may, therefore, be easier to use. The teacher should also be aware of poorly designed websites that may place undue frustration or cognitive load on the student and are best avoided. Learners can be encouraged to anticipate the characteristics they are likely to encounter in the presentation and language of the homepage, and to adopt strategies that can reduce the frustrations and increase the value of using this global resource, the World Wide Web.
Advice to teachers and students that has emerged from this research can be summarised as follows.

**LOOK FOR**
- Preferred format/webpage structure
- Clear, functional links (not too many)
- Familiar content
- Chunked information
- Ease of skimming (signposting, highlighting)
- Uncluttered appearance
- Simple organisational structure
- Light cognitive load.
(And set suitable tasks)

**AVOID**
- Horizontal scrolling
- Complexity and clutter
- Too much freedom.

**STRATEGIES FOR LEARNERS**
- Work out what’s useful (recognise irrelevant ‘fluff’, content and navigation areas).
- Anticipate characteristics of website (structure, scrolling, windows, links).
- Anticipate search styles (exact or ambiguous access schemes).
- Understand how text is presented. Use skimming skills (for chunked, dense, signposted text).
- Avoid information overload.
REFERENCES


