THE LANGUAGE OF WEBSITES

Mark Boardman

Inter text

The Language of Websites

© Mark Boardman

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2005
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Simultaneously published in the USA and Canada
by Routledge
270 Madison Ave, New York, NY 10016

Routledge is an imprint of the Taylor & Francis Group

Typeset in Stone Sans/Stone Serif by
Florence Production Ltd, Stoodleigh, Devon
Printed and bound in Great Britain by
TJ International Ltd, Padstow, Cornwall

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the
British Library

Library of Congress Cataloging in Publication Data
Boardman, Mark, 1961–
The language of websites/Mark Boardman.
p. cm. – (The intertext series)
Includes bibliographical references and indexes.
1. Web sites. 2. World Wide Web. 3. English language – Usage.
I. Title. II. Series: Intertext (London, England)
TK5105.888.B62 2004
025.04'014 – dc 22 2004007127

ISBN 0–415–32853–5 (hbk)
ISBN 0–415–32854–3 (pbk)

acknowledgements

My thanks to Angela Goddard for making the initial suggestion that I have a go at writing this book – thereby implying that I might be up to the task – and for her support in the early stages of writing. I am similarly grateful to Tim Shortis for help with secondary sources and his positive comments. Thanks also to Christy Kirkpatrick for her support and patience in the early stages of writing. Kate Parker's support, patience and understanding have been very much appreciated throughout the writing process. I am highly indebted to Adrian Beard for his enthusiastic and insightful editorial guidance.

I also appreciated the encouraging comments made by staff and students at South Craven School while I was trying to balance writing with teaching.

I am very grateful to the following people and organisations for their kind permission to let me use screenshots from their websites in this book: Nick Reynolds (joint lead vocalist of Foruta), Anthony Coxar, Matt Wells (creator of Gigablast), Chuck Olsen, Mon Athakravi, Hank Lake Coghlan (of Lake Coghlan – IT & Marketing for Business & Law), Salford University, AltaVista Operations Ltd, Footprint Films (Text 9 designed by Franki & Jonny) and Telegraph.co.uk, the website for the *Daily Telegraph* and the *Sunday Telegraph*.

conventions

Throughout this book, the convention of quoting a website address with the http:// prefix has been used. Modern web browsers no longer require this. It is preserved here because it gives a better understanding of the kind of document you are loading when you look at a web page. You also still need to include the prefix if you are making links as a writer of web pages – because there is more than one kind of link.

The Index of terms gives a short explanation of each special term, with the page number where it is first used. Each term is **emboldened** on its first mention only.

There's a lot to be said for being able to sit at a terminal and just dream.

(Tim Berners-Lee)

Bedrooms lack the potentially global bandwidth of homepages on the Web.

(Daniel Chandler and Dilwyn Roberts-Young)

contents

introduction

Some background to language and the Web

Analysing a **text** is not an end in itself, although it may often seem that way if you have to do it under exam conditions. By separating out acts of communication and calling them texts, we hope to gain a better understanding of human communication. Additionally, in the electronic age, there is a problem with defining texts as either written or spoken. A **website** is a particular kind of electronic text that is technologically and culturally related to some aspects of written communication, but websites also have a relationship with spoken interaction and with other forms of electronic text. The technological and cultural factors that led to the appearance of websites are complex, but it would be wrong to attempt a linguistic analysis of websites in a vacuum.

How the **personal computer** (PC) came to be, and how this development dovetailed with the appearance of the Web, are crucial contextual factors in understanding how websites work as cultural artefacts. The context of websites is technologically determined. Much of the context of spoken and written interaction relies on a common sense awareness of the world around you. This is not so with websites: you cannot intuitively know how a computer displays web pages. For this reason, the introduction will provide some historical and technical background, which will be continued in Unit one, with a more specific focus on using language and media/communication studies **frameworks** for analysing web text.

Text analysis is a description of the process of reading – and in the case of websites the texts can only be read in an essentially fragile, technologically defined environment. That environment, or more specifically, the reader's awareness of its fragility, is a key factor in the reading process. This book sets out partly to show you the nature of the technologically defined space that is the context of web reading, so you will have to read through some technical background before you get to the analysis of texts. Technical explanations have generally been kept as simple as possible – but some technical material is needed in order to see how language interfaces with technology. Technology *is* the physical context of the Web, so it needs to be described. To leave it out would be like analysing a conversation without knowing the situation in which the conversation took place.

Any terminology that is given a full explanation in the glossary at the back of the book will appear in bold print the first time it is used. Because of the nature of the Web, the terminology is necessarily a mixture of technological, linguistic and media/communication studies terminology. Also, because this is an interdisciplinary book attempting an approach based on a combination of linguistic, ICT, media/communication studies and, occasionally, literary frameworks, it is a big glossary – larger than the glossaries in most other Intertext titles. The only real assumption the book makes is that the reader is aware of basic word classes – noun, verb, adjective, adverb, determiner, preposition and conjunction. For an approachable handbook that unlocks most of the grammatical mysteries an advanced or undergraduate student would ever encounter, you cannot do much better than David Crystal's (1998) *Rediscover Grammar.*

WHAT THE BOOK IS NOT

The activities in the book are essentially jumping-off points for any number of possible paths that could be taken when looking into the complex ways in which the interface between language and technology works. The commentaries highlight *some* of the possible avenues of investigation suggested by the activities, and also attempt to exemplify a workable linguistic methodology – but the commentaries are definitely not 'answers' to the questions in any kind of closed-off, complete sense. There is no *single* answer to a text analysis question, so if your response

to the activity is different from the commentary this should not be seen as a problem. The Intertext series is all about raising your game as an analyser of texts on language-, media- or communication studies-based courses at advanced and first-year undergraduate level. If you can imitate the style and approach of the commentaries in this book, your game will be raised. The commentaries work by exemplifying an approach, a methodology – *not* by providing 'tick-box', neatly packaged 'answers'.

Similarly, the book suggests a few areas that could be developed by further academic research at higher education level, but it is not making claims to be definitive or exhaustive academically. For example, some of the analysis in this book hinges on the idea of **narrative** as the driving force behind the Web and how it works as a form of communication. This book is too small to pursue all the issues raised by the connection between narrative and the Web – but if you are interested in knowing more about this area (a well-established one academically), there is plenty of further reading in the bibliography.

HUMAN COMMUNICATION

We need to represent ourselves. We need to record our lives. Whatever you decide was the original purpose of cave paintings, it is at least clear that they show a human desire to make a long-term record of human activities. No other species on the planet seems to do this.

Another capability that separates us from other species is language. Chimpanzees can be taught some lexical knowledge, but they are incapable of handling the grammatical, **semantic** and **pragmatic** relations that are the true foundations of human language – and humans are alone in possessing a sophisticated vocal tract. Once we had language, we could go beyond pictures and begin to connect events with a system that has come to be known as narrative. We could now learn not only from our own experience, but also by listening to the experiences of others. The device used to store these experiences and narratives is the human brain – far from infallible, and even now very far from being perfectly understood. But each individual has a brain that is unreliable in slightly *different* ways, unique to them. So, if you send ten people to the same meeting or the same football match or the same party, you will get ten different spoken accounts of that event afterwards.

The storage and retrieval systems that the brain uses are imperfect, and we subconsciously fill in invented details when we cannot remember the actual details, because we wish to sound fluent when recounting an event. Add to this the infinite flexibility of language – choosing a word, choosing the order of the words, choosing features like **pitch**, **intonation** and pragmatic meaning – and you have a planet resonating with millions of unique spoken narratives every second of every day. This has been so since human language began. A good storyteller is someone who has trained their storage and retrieval system to hold more information than average, can retrieve information quickly and can then construct narratives that satisfy and entertain. Cultures are built upon how we respond to the interplay of these narratives – old ones and constantly changing new ones.

As our lives became more complex, we felt the need to fix these narratives in a form that could be replayed without variation and without error – to fix and record our culture and later to make the recordings themselves into cultural artefacts with **conventions** of their own. The only evidence of fixed narratives from the hunter-gatherer phase in our evolution comes from cave paintings – the precise purpose of which remains unknown. There may have been a predictable, culturally defined way of decoding them into spoken language, but this too will never be known.

It is because of the change from hunter-gatherer to farmer that we see social groups becoming less geographically mobile and beginning to keep records of property. Thus, the earliest attempts at fixing spoken language in **graphical**, **symbolic** form may have begun with Neolithic counting tokens about nine thousand years ago. As these systems for recording spoken language became more complex, we eventually had something that could indeed fix our narratives, so that we could store them away, making them independent of anyone's individual memory. The creator of the narrative can die and be forgotten – but, if the fixed (written) version of the narrative is preserved, it can be read and replayed hundreds or thousands of years later. There is the inevitable problem that spoken language changes very quickly so that, even if a document still exists, the **code** used to construct it may no longer be in use – a problem that any student who has struggled with Shakespeare (or whoever it was who wrote those plays) will recognise.

If you want to make your narrative known to large numbers of people, you have to find a way of giving other people a copy of it. Before writing systems were invented, the only way of doing this was to arrange for someone else to memorise your story. This is subject to the inherent limitations of the human memory that we have outlined

above, and because of this the story will undergo changes as it is passed from one member of a society to another. This will also happen through time, as the story is passed down from one generation to the next. Even after the invention of writing, though, literacy was not widespread, so that the transmission of knowledge and cultural values would still take place orally while literacy remained the preserve of the educated few. Even now, many languages and cultures have no writing system. But in the (post-industrial revolution) Western world we are accustomed to the idea that writing is one way of preserving spoken language.

Written language has now, of course, developed into a **variety** in its own right, but there is little doubt that its original purpose was to transcribe and fix spoken language. One example of this is the fragmentary evidence of Old English poetry in written form – a partial preservation of some artefacts from a purely oral tradition of poetic composition and recital. Having no means, at that time, of mass-producing written text, the educated elite of Anglo-Saxon England could only disseminate copies of poems by writing them out manually and distributing them, very slowly, to a very restricted audience. With the growth of literacy over the next five hundred years, written language began to develop its own features and was no longer simply a brass rubbing of spoken language. Now, in the early twenty-first century, written language has many **dialectal** and **discourse** features distinct from spoken language.

Spoken language is natural. Literacy is artificial. This is why opinion varies about how to teach children to read and write in their early years, and it is why the teaching of literacy is the subject of legislation in many countries. You cannot legislate for the acquisition of spoken language: it happens by itself, given normal social, cognitive and physical development in the child. There is no need for any kind of technology to spread the spoken **forms** and texts of a language: this was given a considerable helping hand by the invention of **analogue** recording equipment in 1877, but be assured it had been happening for many thousands of years prior to that, using the natural storage and retrieval capacity of the human brain. However, because *written* texts are artificial, they need technology to disseminate them. In the middle of the fifteenth century, Johann Gutenberg invented a system for doing this, which remained essentially unchanged for five hundred years. Mass production of written texts, stored and distributed on paper, gave the written forms of language a public profile. The mechanism for retrieval is the scanning of the eye over the page and the cognition of the scanned stream of letters as meaningful language. This is an unnatural process that needs to be taught. Fast forward to the 1930s, and we

see the first breakthrough in storing text on media other than paper. The M. Shultz Company's repetitive typewriter used a system of punched cards which operated the keys of the typewriter automatically – in much the same way as a pianola plays a piece of music. Herein was the beginning of document templates being stored for repeated production – the birth of **word processing**. The first machine to be marketed using the term 'word processing' was IBM's Magnetic Tape/Selectric Typewriter or MT/ST in 1964. This used analogue tape as the first reusable medium for storing text – to be replaced by magnetic cards in 1969. In the early 1970s, word-processing machines began to incorporate basic video screens for manipulating text prior to printing, and floppy disk drives for storing text.

Sidestepping a little, and rewinding to 1822, Charles Babbage invented a machine for performing mathematical calculations. Although this machine was destined not to be manufactured until 1991 (when the Science Museum in London followed Babbage's plans to prove that it could be done), the history of computing followed a mechanical path until 1935, when Konrad Zuse began work on an electronic version of his own mechanical calculating machine. Precisely who invented the first **digital** computer remains a matter of debate, but the concept of 'digital switching' to accomplish calculations was famously deployed during the Second World War to break the German Enigma Code, using a computer called Colossus Mark I. These early computers were powered by 'relays' or electrically operated mechanical switches. It was while one such machine was stopped for repair that someone noticed a dead moth in one of the relays (the possible cause of the problem), and this has led to the all too familiar metaphorical term '**bug**' in modern computer troubleshooting. In 1947, work was completed on the **transistor** – a type of **electronic** switch – and this paved the way for the electronics of the personal computer that we are familiar with today. And in 1961, integrated circuits appeared, effectively highly miniaturised transistor circuits, which developed into the microprocessor chips that form the basis of personal computers today.

Unit one

Getting really wired: the physical context of websites

AIMS OF THIS UNIT

The Introduction provided the historical and technical background needed to begin to explore the immediate context of websites. We will look at that context in more detail in this unit, and we will begin to explore ways of using that contextual awareness to apply analytical frameworks to websites in a meaningful way.

A TERRIBLE BEAUTY IS BORN: HOW PERSONAL COMPUTERS BEGAN

The first personal computer was marketed by Apple in 1977. Dedicated word-processing machines continued to exist (and do to this day), but if you wanted more power to manipulate text, and also more electronic space to store your documents, you could load some word-processing software into your personal computer. **Software** is simply a set of instructions for the computer, prepared by a computer **programmer**, so that a non-specialist user can 'load' it into the computer. Loading different software can give you many other functions like a searchable

'filing cabinet', a bookkeeping ledger or an arcade game. A dedicated word-processing machine has its software **hard-wired** so that if you want to upgrade to more powerful functions you have to replace the machine. Personal computers marked the separation of machines from software, so that you could keep your old machine but buy new software for it.

A problem for early developers of the personal computer was how to make the user feel that they were working in a familiar environment. To cut a rather complex and technical story very short, the solution to this was the now famous **desktop** metaphor – a working environment that makes the computer **environment** *look like* a desk with a set of filing cabinet draws built into it. To enable the user to control the desk and the filing cabinet, the architects of the personal computer borrowed the **QWERTY keyboard** from the familiar office typewriter, and added a virtual hand that the user could use to delve into the electronic world of the computer and move things around – what we now know as a mouse. Both the desktop and the mouse were invented by Xerox at their Palo Alto Research Center in the 1970s. They still form the basis of the software environment we now know as the **operating system**.

We will always need metaphors to interact with computers. Computers only understand sequences of 1 and 0 (digits or digital information), telling them to open or close millions of tiny switches, but no person can calculate quickly enough to enter the raw numbers (modern computers are too fast and too complicated), so we need a fake world where intuitive manual actions are fed to the computer as raw numbers. This is what software does: it is called the 'operating system' or user interface.

The eternal problem is how to make the computer anticipate how humans think and behave, and how to make people think and behave in a way that the computer can anticipate – a concept that Brenda Laurel (Bardini 1997) has referred to as 'horrible recursion', and one that lies at the heart of user interface design. Not surprisingly, many science fiction writers have imagined that this will conclude with the metaphor becoming a reality – where the computers become so complex that they are capable of changing the conditions of the interface themselves: they become 'conscious'.

EVERYTHING WITHIN: THE ORIGINS OF THE WEB

Web pages are just another software environment, another refinement of the user interface – albeit an influential one, which is why most aspects of the desktop environment on modern personal computers can be made to behave like a web page. But where did web pages come from?

If you want to retrieve a document from a filing cabinet drawer in your own desk, you do not have to leave your seat. If you want to retrieve a document from someone else's desk drawer, you have to get out of your seat and walk or wait for someone to post it to you. Exactly the same was true of the personal computer in the late 1970s and early 1980s. Documents could be made digital and then stored or retrieved from a magnetic disk in their original form – free from error or deterioration. You could even alter the document and store it in a revised form alongside of or instead of the old one. It seemed that printing and writing had come of age and we had finally and completely overcome the error-prone storage and retrieval system of the human brain.

We wanted more though. By the late 1980s, **networking** of large **mainframe** computers (**Wide Area Networks**) had been a reality for many years. Early incarnations of the **Internet** consisted of linked computers that could exchange **packets** of information over long distances – mainly between universities in the United States. This same principle was implemented in the late 1980s for linking personal computers in a business environment – the concept of the **Local Area Network**. It arose from the desire to move and copy documents or **files** from one personal computer to another in an office, thus saving time and money, but it left in place the pathways that, by the mid-1990s, would make possible the information explosion known as the World Wide Web.

Tim Berners-Lee, a physicist frustrated by the fact that his many computing projects at the European Laboratory for Particle Physics were stored on different machines, decided to develop a system that would link these projects together and allow information (or **data**) to travel between them. He wanted it to be a system that would allow a group of physicists to collaborate on several projects simultaneously, without having to move backwards and forwards between several different machines and switch between several different, incompatible, computer environments. He based his new system on **hypertext**, a concept that was influential in the design of computer environments long before the Web. Hypertext stems largely from the ideas of Ted Nelson and Douglas Engelbart.