The Plain Web

Chuck Letourneau
Starling Access Services, Ottawa, Ontario

Presented to the Fourth Biennial Conference of the PLAIN Language Association International, September 27, 2002

www.starlingweb.com/

Table of Contents

I actually began my career in Web accessibility because of plain language. Way back in 1995 I was working for the Federal Government of Canada and I wanted to make a Web site to advertise services for my clients who were employees who had disabilities of one sort or another. I wanted to make sure that anybody could visit my site and get the information they needed. Because I had some doubts about whether people with certain disabilities could use the Web I did some research and discovered a marvelous set of guidelines developed by Dr. Gregg Vanderheiden and his crew at the University of Wisconsin's TRACE Center. However, being a self-taught and still very amateur Web programmer, I could barely make sense of that highly technical document. But, I needed to understand the concepts, so I stuck to it and finally managed to figure it out. At that point, and with some rather uncharacteristic foresight, I realized that Web accessibility for persons with disabilities was the sort of thing that would, sooner or later, become an issue, especially for governments. I also felt that few enough people would make the effort to decipher the excellent, but complex and technical TRACE document. Don't forget that in 1995 most Web developers were, like me, entirely self-taught and had got whatever knowledge we had mostly by copying the work of others who had managed to develop working Web sites. I felt there was a need for a "plain language" version of the Web accessibility guidelines to serve the average Web professional.

It turned out I was right on both counts: in the last few years governments around the world have made it law or policy to ensure their Web sites are accessible to all their citizens. And almost every Web master or content developer I have met demands less-technical but still-usable documentation on the whole subject of Web accessibility. I feel privileged to have had some role in causing the former and creating the latter. My efforts in the young field of Web accessibility caused me to get involved with the World Wide Web Consortium - the W3C - an industry consortium of over 400 companies whose business involves the Internet either as technical innovators, content providers or both. The consortium includes companies like Microsoft, IBM, Adobe, Compaq, America On-Line, Macromedia, and so on.

In 1997, recognizing the importance of keeping the Web accessible to everyone, the W3C started the Web Accessibility Initiative. A bunch of working groups were formed to address various issues including: accessibility of content, authoring tools, browsers, Web languages, etc. Because of my prior involvement in the field, I was asked to co-chair (with Dr. Vanderheiden) the Web Content Accessibility Guidelines working group. Our task was to produce, essentially, a standard that Web developers could follow to ensure that Web sites were accessible. We did that, and published them in May of 1999. The document outlines 14 major guidelines that developers must consider, and one of those deals with plain language.

What's wrong with the Web?

Now I'll tell you why the Web causes problems for some people. It's not having a disability that's the problem. Absolutely not! The biggest problem is that many Web sites and the information on them are in a form that simply can't be used by a lot of people.

In my experience most people who have disabilities are quite capable of reading and understanding. The difference is that they may read or comprehend differently from people without "print-handicaps". A person who is blind may not be able to see my Web page, but that's not a problem if the information is converted to spoken form or perhaps Braille. A person who is deaf or hard of hearing may get nothing from an audio clip of a politician's speech, but that barrier disappears if I provide a text-transcript of the speech, or visual captioning.

What if my visitor can't read... perhaps because of a reading disability, perhaps illiteracy, or simply doesn't read the language of my site? Well, the text-to-speech technology used by someone who is blind will work just as well here. Similarly, someone who can read the language but has difficulty comprehending the spoken word could benefit from having equivalent textual content.

But a far more common problem is that a lot of the content published on the Web is simply incomprehensible: and not just to people who have cognitive disabilities, or educational or literacy limitations. I am most familiar with government Web sites and, well... do I really have to say anything to you folks about the quality of the majority of public sector content? But, given the nature and background of most Web content providers, I suppose this is not surprising: few are professional writers, and fewer still are plain language professionals.

PLAIN Language guidelines and checkpoints

Most of the WCAG guidelines are intended to overcome technological barriers faced by people with certain functional limitations. How can we make text available to someone who can't see? How can we make audio content available to someone who can't hear? How can we make it possible to surf the Web for someone who can't use a mouse? Those are all pretty easy in one sense. But the question of how to make the content or topic of a document understandable by someone who for whatever reason doesn't understand it is not, in my experience, a technological question.

From the beginning, we realized that we had two problems: we had to identify and overcome the technological barriers to using the Web (and we have, by and large, done so) and we had to address the barriers to understanding. Even the very first guidelines made some mention of the need for "plain language". Granted, we were most concerned with ensuring that Web content was understandable by people with learning or cognitive disabilities, but, in my circles, we have a saying:

"Anything that makes the Web easier for someone with a disability makes it easier for everyone."

So, what did the guideline say? It said:

Ensure that documents are clear and simple so they may be more easily understood.

And we explained that by saying:

Consistent page layout, recognizable graphics, and easy to understand language benefit all users. In particular, they help people with cognitive disabilities or who have difficulty reading. (However, ensure that images have text equivalents for people who are blind, have low vision, or for any user who cannot or has chosen not to view graphics.)

Using clear and simple language promotes effective communication. Access to written information can be difficult to impossible for people who have cognitive disabilities, learning disabilities, or who are deaf. This consideration also applies to the many people whose first language differs from your own.

So that told us why to use plain language, but didn't tell us what to do or how to do it. In the Web Content Accessibility Guidelines, the what-to-dos are described in a series of checkpoints that accompany each guideline. The checkpoint in this case states:

Use the clearest and simplest language appropriate for a site's content.

I should point out to you that this is a "Priority 1" checkpoint. That means that:

An author must follow this guideline, or one or more groups of users will find it impossible to access information in the document. Implementing this guideline is a basic requirement for some groups to be able to use Web documents.

All governments or organizations that have adopted the W3C's guidelines as their own law or policy at least require compliance with all the Priority 1 checkpoints. The Government of Canada's "Common Look and Feel Policy", for instance, requires all Federal Web sites to be compliant with both Priority 1 and Priority 2 checkpoints by December 31, 2002. The same is true, I believe, for the Government of Ontario.

Unfortunately, during the period we were developing the guidelines - from 1997 through 1999 - we didn't have any plain language professionals involved. We were obviously bright enough to know that plain language was important, but we didn't really know what to tell people. So the best we could say was:

Follow these writing suggestions:

- Strive for clear and accurate headings and link descriptions. Review every heading, outline, and menu to see if the crucial words mean exactly what is intended, and if there are more common words that would convey the same meaning.
- State the topic of the sentence or paragraph at the beginning of the sentence or paragraph.
- Limit each paragraph to one main idea.
- Avoid idiomatic language, technical jargon, and other unfamiliar vocabulary and expressions.

- Avoid specialized meanings of familiar vocabulary, unless explanations are provided.
- Avoid the passive voice.
- Avoid complex sentence structures.
- Make link phrases terse yet meaningful enough so they make sense when read out of context, alone or as part of a series of links.

A revised version of guidelines included the following advice:

5.1 Writing style

The following writing style suggestions should help make the content of your site easier to read for everyone, especially people with reading and/or cognitive disabilities. Several guides (including [HACKER]) discuss these and other writing style issues in more detail.

- Strive for clear and accurate headings and link descriptions. This includes using link phrases that are terse and that make sense when read out of context or as part of a series of links (Some users browse by jumping from link to link and listening only to link text.) Use informative headings so that users can scan a page quickly for information rather than reading it in detail.
- State the topic of the sentence or paragraph at the beginning of the sentence or paragraph (this is called "front-loading"). This will help both people who are skimming visually, but also people who use speech synthesizers. "Skimming" with speech currently means that the user jumps from heading to heading, or paragraph to paragraph and listens to just enough words to determine whether the current chunk of information (heading, paragraph, link, etc.) interests them. If the main idea of the paragraph is in the middle or at the end, speech users may have to listen to most of the document before finding what they want. Depending on what the user is looking for and how much they know about the topic, search features may also help users locate content more quickly.
- Limit each paragraph to one main idea.
- Avoid slang, jargon, and specialized meanings of familiar words, unless defined within your document.
- Favor words that are commonly used. For example, use "begin" rather than "commence" or use "try" rather than "endeavor."
- Use active rather than passive verbs.
- Avoid complex sentence structures.

To help determine whether your document is easy to read, consider using the Gunning-Fog reading measure (described in [SPOOL] with examples and the algorithm online at [TECHHEAD]). This algorithm generally produces a lower score when content is easier to read. As example results, the Bible, Shakespeare, Mark Twain, and TV Guide all have Fog indexes of about 6. Time, Newsweek, and the Wall St. Journal an average Fog index of about 11.

By the way, I've noticed that with each passing year I find it harder and harder to detect and correct jargon in my chosen field. As I become more knowledgeable I become less critical: I can understand my colleagues even if they speak pure jargon, so I have to actually think harder about how to translate into plain language.

HTML Techniques

There are a number of other features built in to the most common Web page development language - HTML - that can help with plain language. I am going to have to use a bit of jargon here.

Alternative text

If you have a Web site, you probably have some pictures, drawings or other graphics on some pages, possibly for marketing purposes, possibly for visual appeal, possibly as helpful visual aids (e.g. icons, link-buttons, etc.). There is a feature in the language that allows you to briefly describe, in text, the meaning or intent of each image. [Example] People who can't see, or who can't understand an image can read this description instead. Of course the textual description only show up if you ask the browser to display it. In fact the most recent international standards for Web languages all require this alternative text to be supplied, all the time.

Long descriptions

Sometimes, an image may be too complex to describe in just a few words... think of a picture of 3-dimensional bar graph in a report describing your company's sales of ice-cube trays [Example]. It may benefit more than just a blind person to have the intent of this image described in some detail. A feature of HTML allows you to link to a separate document that contains the long description. This means you don't have to clutter up your main page with explanations that not everyone might want. Again, the display of this extra page is optional.

Abbreviations and Acronyms

One of my biggest gripes, especially when looking at Government sites is their aggressive and annoying use of acronyms and abbreviations. Many of my government colleagues can talk at length without actually using real words. Unfortunately, this habit often appears in their writing as well. If I know what "ADIO" or "ComFlanSync" means, then anyone listening to me must know it as well. HTML gives you the ability to use acronyms and abbreviations and provide the expanded versions as well. The expansion is only seen if you ask the browser to show it to you.

Summaries and Captions

Many people insist on putting data on the Web in tabular form. Tables are sometimes incredibly difficult to comprehend. HTML recommends, and provides a method for including textual summaries and captions for tabular data.

Language tags

Every HTML document is required, by the standard, to provide an indication of the main language of the page: English, or French, or German, or whatever. There are special codes for every language. This element can be used either by the browser to automatically get you a copy of a multilingual document in your language of choice, or by certain special applications to do an automatic translation of a unilingual document into another language. But beyond that, if you have just a snippet French in an otherwise English page, you can identify the language of just that snippet. Some future application might be able to translate snippets on the fly.

Metadata - Information about information

Another way to aid comprehension of a Web document or Web site is to briefly summarize its content, and provide a standard set of keywords to better identify its subjects. This kind of information can be stored in something called metadata, which is made available on request to things like search engines or specialty browsers. By the way, I heard someone talking about meta-metadata a little while ago, and I still have the headache it gave me.

The Future

So, what challenges face us? Well, the most common complaints we hear from Web developers who have been told to implement the W3C guidelines is that they don't know what plain language means, that in any case their subject couldn't possibly be made more understandable, that the audience they care about does understand and to hell with the rest of the world, and really who decides what is plain enough language to meet the standard.

I am pleased to be able to tell you that the working group currently developing the second version of the Web Content Accessibility Guidelines is trying to address these issues in a more detailed fashion. They even have a plain language professional involved this time: Avi Arditti (who is attending this conference), with the Voice of America, has been quite active in this regard. You can search on the W3C's Web site for more information about the most recent work in this area. Look at the list-serve archives of the Web Content Accessibility Guidelines Working Group for message subjects that contain the subject: 4.1, or plain language, or natural language. The URL is www.w3.org/WAI/GL. There you can also find links to the latest drafts of the new guidelines and see their progress on the new plain language checkpoints.

Chuck Letourneau

www.starlingweb.com/

Chuck Letourneau was co-chair of the World Wide Web Consortium (W3C) working group that developed the Web Content Accessibility Guidelines. He is not a plain language professional (not too obviously, he hopes), though he cares deeply about the subject of plain language.