

Emacspeak At Twenty: Looking Back, Looking Forward

T. V. Raman*

<2014-09-12 Fri>

Contents

1	Introduction	2
2	Using UNIX With Speech Output — 1994	2
3	Key Enabler — Emacs And Lisp Advice	3
4	Key Component — Text To Speech (TTS)	3
5	Emacspeak And Software Development	4
5.1	Programming Defensively	5
6	Emacspeak And Authoring Documents	6
7	Emacspeak And The Early Days Of The Web	6
8	Audio Formatting — Generalizing Aural CSS	8
9	Conversational Gestures For The Audio Desktop	9
9.1	Speech-Enabling Interactive Games	10
10	Accessing Media Streams	10
11	EBooks— Ubiquitous Access To Books	11

*tv.raman.tv@gmail.com

12 Leveraging Computational Tools — From SQL And R To IPython Notebooks	12
13 Social Web — EMail, Instant Messaging, Blogging And Tweeting Using Open Protocols	12
14 The RESTful Web — Web Wizards And URL Templates For Faster Access	14
15 Mashing It Up — Leveraging Evolving Web APIs	15
16 Conclusion	16
17 References	17

1 Introduction

One afternoon in the third week of September 1994, I started writing myself a small Emacs extension using Lisp Advice to make Emacs speak to me so I could use a Linux laptop. As Emacspeak turns twenty, this article is both a quick look back over the twenty years of lessons learned, as well as a glimpse into what might be possible as we evolve to a world of connected, ubiquitous computing. This article draws on Learning To Program In 10 Years by Peter Norvig for some of its inspiration.

2 Using UNIX With Speech Output — 1994

As a graduate student at Cornell, I accessed my Unix workstation (SunOS) from an Intel 486 PC running IBM Screen-Reader. There was no means of directly using a UNIX box at the time; after graduating, I continued doing the same for about six months at Digital Research in Cambridge — the only difference being that my desktop workstation was now a DEC-Alpha. Throughout this time, Emacs was my environment of choice for everything from software development and Internet access to writing documents.

In fall of 1994, I wanted to start using a laptop running Linux; a colleague (Dave Wecker) was retiring his 386mhz laptop that already had Linux on it and I decided to inherit it. But there was only one problem — until then I had always accessed a UNIX machine from a secondary PC running a screen-reader — something that would clearly make no sense with a laptop!

Another colleague, Win Treese, had pointed out the interesting possibilities presented by package **advice** in Emacs 19.23 — a few weeks earlier, he had sent around a small snippet of code that magically modified Emacs’ version-control primitive to first create an *RCS* directory if none existed before adding a file to version control. When I speculated about using the Linux laptop, Dave remarked — you live in Emacs anyway — why dont you just make it talk!

Connecting the dots, I decided to write myself a tool that augmented Emacs’ default behavior to *speak* — within about 4 hours, version 0.01 of Emacspeak was up and running.

3 Key Enabler — Emacs And Lisp Advice

It took me a couple of weeks to fully recognize the potential of what I had built with Emacs Lisp Advice. Until then, I had used screen-readers to listen to the contents of the visual display — but Lisp Advice let me do a lot more — it enabled Emacspeak to generate highly context-specific spoken feedback, augmented by a set of auditory icons. I later formalized this design under the name speech-enabled applications. For a detailed overview of the architecture of Emacspeak, see the chapter on Emacspeak in the book Beautiful Code from O’Reilly.

4 Key Component — Text To Speech (TTS)

Emacspeak is a speech-subsystem for Emacs; it depends on an external Text-To-Speech (TTS) engine to produce speech. In 1994, Digital Equipment released what would turn out to be the last in the line of hardware DECTalk synthesizers, the DECTalk Express. This was essentially an Intel 386 with 1mb of flash memory that ran a version of the DECTalk TTS software — to date, it still remains my favorite Text-To-Speech engine. At the time, I also had a software version of the same engine running on my DEC-Alpha workstation; the desire to use either a software or hardware solution to produce speech output defined the Emacspeak speech-server architecture.

I went to IBM Research in 1999; this coincided with IBM releasing a version of the Eloquent TTS engine on Linux under the name *ViaVoice Outloud*. My colleague Jeffrey Sorenson implemented an early version of the Emacspeak speech-server for this engine using the OSS API; I later updated it to use the ALSA library while on a flight back to SFO from Boston in

2001. That is still the TTS engine that is speaking as I type this article on my laptop.

20 years on, TTS continues to be the weakest link on Linux; the best available solution in terms of quality continues to be the Linux port of Eloquence TTS available from Voxin in Europe for a small price. Looking back across 20 years, the state of TTS on Linux in particular and across all platforms in general continues to be a disappointment; most of today's newer TTS engines are geared toward mainstream use-cases where *naturalness* of the voice tends to supersede intelligibility at higher speech-rates. Ironically, modern TTS engines also give applications far less control over the generated output — as a case in point, I implemented Audio System For Technical Readings (AsTeR) in 1994 using the DECTalk; 20 years later, we implemented MathML support in ChromeVox using Google TTS. In 2013, it turned out to be difficult or impossible to implement the type of audio renderings that were possible with the admittedly less-natural sounding DECTalk!

5 Emacspeak And Software Development

Version 0.01 of Emacspeak was written using IBM Screen-Reader on a PC with a terminal emulator accessing a UNIX workstation. But in about 2 weeks, Emacspeak was already a better environment for developing Emacspeak in particular and software development in general. Here are a few highlights in 1994 that made Emacspeak a good software development environment, present-day users of Emacspeak will see that that was just scratching the surface.

- Audio formatting using voice-lock to provide aural syntax highlighting.
- Succinct auditory icons to provide efficient feedback.
- Emacs' ability to navigate code structurally —

as opposed to moving around by plain-text units such as characters, lines and words. S-Expressions are a major win!

- Emacs' ability to specialize behavior based on major and minor modes.
- Ability to browse program code using tags, and getting fluent spoken feedback.
- Completion **everywhere**.

- Everything is searchable — this is a huge win when you cannot see the screen.
- Interactive spell-checking using ISpell with continuous spoken feedback augmented by aural highlights.
- Running code compilation and being able to jump to errors with spoken feedback.
- Ability to move through diff chunks when working with source code and source control systems; refined diffs as provided by the **ediff** package when speech-enabled is a major productivity win.
- Ability to easily move between email, document authoring and programming — though this may appear trivial, it continues to be one of Emacs' biggest wins.

Long-term Emacs users will recognize all of the above as being among the reasons why they do most things inside Emacs — there is little that is Emacspeak specific in the above list — except that Emacspeak was able to provide fluent, well-integrated contextual feedback for all of these tasks. And that was a game-changer given what I had had before Emacspeak. As a case in point, I did not dare program in Python before I speech-enabled Emacs' Python-Mode; the fact that white space is significant in Python made it difficult to program using a plain screen-reader that was unaware of the semantics of the underlying content being accessed.

5.1 Programming Defensively

As an aside, note that all of Emacspeak has been developed over the last 20 years with Emacspeak being the only adaptive technology on my system. This has led to some interesting design consequences, primary among them being a strong education in **programming defensively**. Here are some other key features of the Emacspeak code-base:

1. The code-base is extremely **bushy** rather than deeply hierarchical — this means that when a module breaks, it does not affect the rest of the system.
2. Separation of concerns with respect to the various layers, a tightly knit core speech library interfaces with any one of many speech servers running as an external process.

3. Audio formatting is abstracted by using the formalism defined in Aural CSS.
4. Emacspeak integrates with Emacs' user interface conventions by taking over a single prefix key **C-e** with *all* Emacspeak commands accessed through that single keymap. This helps embedding Emacspeak functionality into a large variety of third party modules without any loss of functionality.

6 Emacspeak And Authoring Documents

In 1994, my preferred environment for authoring **all** documents was **L^AT_EX** using the Auctex package. Later I started writing either **L^AT_EX** or HTML using the appropriate support modes; today I use **org-mode** to do most of my content authoring. Personally, I have never been a fan of What You See Is What You Get (WYSIWYG) authoring tools — in my experience that places an undue burden on the author by drawing attention away from the content to focus on the final appearance. An added benefit of creating content in Emacs in the form of light-weight markup is that the content is long-lived — I can still usefully process and re-use things I have written 25 years ago.

Emacs, with Emacspeak providing audio formatting and context-specific feedback remains my environment of choice for writing all forms of content ranging from simple email messages to polished documents for print publishing. And it is worth repeating that I **never** need to focus on what the content is going to look like — that job is best left to the computer.

As an example of producing high-fidelity visual content, see this write-up on Polyhedral Geometry that I published in 2002; all of the content, including the drawings were created by me using Emacs.

7 Emacspeak And The Early Days Of The Web

Right around the time that I was writing version 0.01 of emacspeak, a far more significant software movement was under way — the World Wide Web was moving from the realms of academia to the mainstream world with the launch of NCSA Mosaic — and in late 1994 by the first commercial Web browser in Netscape Navigator. Emacs had always enabled integrated access to FTP archives via package *ange-ftp*; in late 1993, William Perry released Emacs-W3, a Web browser for Emacs written entirely in Emacs Lisp. W3

was one of the first large packages to be speech-enabled by Emacspeak — later it was the browser on which I implemented the first draft of the Aural CSS specification. Emacs-W3 enabled many early innovations in the context of providing non-visual access to Web content, including audio formatting and structured content navigation; in summer of 1995, Dave Raggett and I outlined a few extensions to HTML Forms, including the **label** element as a means of associating metadata with interactive form controls in HTML, and many of these ideas were prototyped in Emacs-W3 at the time. Over the years, Emacs-W3 fell behind the times — especially as the Web moved away from cleanly structured HTML to a massive soup of unmatched tags. This made parsing and error-correcting badly-formed HTML markup expensive to do in Emacs-Lisp — and performance suffered. To add to this, mainstream users moved away because Emacs’ rendering engine at the time was not rich enough to provide the type of visual renderings that users had come to expect. The advent of DHTML, and JavaScript based Web Applications finally killed off Emacs-W3 as far as most Emacs users were concerned.

But Emacs-W3 went through a revival on the emacspeak audio desktop in late 1999 with the arrival of XSLT, and Daniel Veillard’s excellent implementation via the **libxml2** and **libxslt** packages. With these in hand, Emacspeak was able to hand-off the bulk of HTML error correction to the **xsltproc** tool. The lack of visual fidelity didn’t matter much for an eyes-free environment; so Emacs-W3 continued to be a useful tool for consuming large amounts of Web content that did not require JavaScript support.

During the last 24 months, **libxml2** has been built into Emacs; this means that you can now parse arbitrary HTML as found in the wild without incurring a performance hit. This functionality was leveraged first by package **shr** (Simple HTML Renderer) within the **gnus** package for rendering HTML email. Later, the author of **gnus** and **shr** created a new light-weight HTML viewer called **eww** that is now part of Emacs 24. With improved support for variable pitch fonts and image embedding, Emacs is once again able to provide visual renderings for a large proportion of text-heavy Web content where it becomes useful for mainstream Emacs users to view at least some Web content within Emacs; during the last year, I have added support within emacspeak to extend package **eww** with support for DOM filtering and quick content navigation.

8 Audio Formatting — Generalizing Aural CSS

A key idea in Audio System For Technical Readings (AsTeR) was the use of various voice properties in combination with non-speech auditory icons to create rich aural renderings. When I implemented Emacspeak, I brought over the notion of audio formatting to all buffers in Emacs by creating a **voice-lock** module that paralleled Emacs' **font-lock** module. The visual medium is far richer in terms of available fonts and colors as compared to voice parameters available on TTS engines — consequently, it did not make sense to directly map Emacs' **face** properties to voice parameters. To aid in projecting visual formatting onto auditory space, I created property **personality** analogous to Emacs' **face** property that could be applied to content displayed in Emacs; module **voice-lock** applied that property appropriately, and the Emacspeak core handled the details of mapping personality values to the underlying TTS engine.

The values used in property **personality** were abstract, i.e., they were independent of any given speech engine. Later in the fall of 1995, I re-expressed these set of abstract voice properties in terms of Aural CSS; the work was published as a first draft toward the end of 1995, and implemented in Emacs-W3 in early 1996. Aural CSS was an appendix in the CSS-1.0 specification; later, it graduated to being its own module within CSS-2.0.

Later in 1996, all of Emacs' **voice-lock** functionality was re-implemented in terms of Aural CSS; the implementation has stood the test of time in that as I added support for more TTS engines, I was able to implement engine-specific mappings of Aural-CSS values. This meant that the rest of Emacspeak could define various types of voices for use in specific contexts without having to worry about individual TTS engines. Conceptually, property **personality** can be thought of as holding an **aural display list** — various parts of the system can annotate pieces of text with relevant properties that finally get rendered in the aggregate. This model also works well with the notion of Emacs overlays where a moving overlay is used to temporarily highlight text that has other context-specific properties applied to it.

Audio formatting as implemented in Emacspeak is extremely effective when working with all types of content ranging from richly structured markup documents (\LaTeX , org-mode) and formatted Web pages to program source code. Perceptually, switching to audio formatted output feels like switching from a black-and-white monitor to a rich color display. Today, Emacspeak's audio formatted output is the only way I can correctly write **else if** vs **elsif** in various programming languages!

9 Conversational Gestures For The Audio Desktop

By 1996, Emacspeak was the only piece of adaptive technology I used; in fall of 1995, I had moved to Adobe Systems from DEC Research to focus on enhancing the Portable Document Format (PDF) to make PDF content repurposable. Between 1996 and 1998, I was primarily focused on electronic document formats — I took this opportunity to step back and evaluate what I had built as an auditory interface within Emacspeak. This retrospect proved extremely useful in gaining a sense of perspective and led to formalizing the high-level concept of *Conversational Gestures* and structured browsing/searching as a means of thinking about user interfaces.

By now, Emacspeak was a complete environment — I formalized what it provided under the moniker *Complete Audio Desktop*. The fully integrated user experience allowed me to move forward with respect to defining interaction models that were highly optimized to eyes-free interaction — as an example, see how Emacspeak interfaces with modes like **dired** (Directory Editor) for browsing and manipulating the filesystem, or **proced** (Process Editor) for browsing and manipulating running processes. Emacs' integration with **ispell** for spell checking, as well as its various completion facilities ranging from minibuffer completion to other forms of dynamic completion while typing text provided more opportunities for creating innovative forms of eyes-free interaction. With respect to what had gone before (and is still par for the course as far as traditional screen-readers are concerned), these types of highly dynamic interfaces present a challenge. For example, consider handling a completion interface using a screen-reader that is speaking the visual display. There is a significant challenge in deciding *what to speak* e.g., when presented with a list of completions, the currently typed text, and the default completion, which of these should you speak, and in what order? The problem gets harder when you consider that the underlying semantics of these items is generally not available from examining the visual presentation in a consistent manner. By having direct access to the underlying information being presented, Emacspeak had a leg up with respect to addressing the higher-level question — when you do have access to this information, how do you present it effectively in an eyes-free environment? For this and many other cases of dynamic interaction, a combination of audio formatting, auditory icons, and the ability to synthesize succinct messages from a combination of information items — rather than having to forcibly speak each item as it is rendered visually provided for highly efficient eyes-free interaction.

This was also when I stepped back to build out Emacspeak's table browsing facilities — see the online Emacspeak documentation for details on Emac-

peak’s table browsing functionality which continues to remain one of the richest collection of end-user affordances for working with two-dimensional data.

9.1 Speech-Enabling Interactive Games

So in 1997, I went the next step in asking — given access to the underlying information, is it possible to build effective eyes-free interaction to highly interactive tasks? I picked **Tetris** as a means of exploring this space, the result was an Emacspeak extension to speech-enable module **tetris.el**. The details of what was learned were published as a paper in Assets 98, and expanded as a chapter on Conversational Gestures in my book on Auditory Interfaces; that book was in a sense a culmination of stepping back and gaining a sense of perspective of what I had build during this period. The work on Conversational Gestures also helped in formalizing the abstract user interface layer that formed part of the XForms work at the W3C.

Speech-enabling games for effective eyes-free interaction has proven highly educational. Interactive games are typically built to challenge the user, and if the eyes-free interface is inefficient, you just won’t play the game — contrast this with a task that you **must** perform, where you’re likely to make do with a sub-optimal interface. Over the years, Emacspeak has come to include eyes-free interfaces to several games including Tetris, Sudoku, and of late the popular 2048 game. Each of these have in turn contributed to enhancing the interaction model in Emacspeak, and those innovations typically make their way to the rest of the environment.

10 Accessing Media Streams

Streaming real-time audio on the Internet became a reality with the advent of RealAudio in 1995; soon there were a large number of media streams available on the Internet ranging from music streams to live radio stations. But there was an interesting twist — for the most part, all of these media streams expected one to look at the screen, even though the primary content was purely audio (streaming video hadn’t arrived yet!). Starting in 1996, Emacspeak started including a variety of eyes-free front-ends for accessing media streams. Initially, this was achieved by building a wrapper around **trplayer** — a headless version of RealPlayer; later I built Emacspeak module **emacspeak-m-player** for interfacing with package **mplayer**. A key aspect of streaming media integration in emacspeak is that one can launch and control streams without ever switching away from one’s primary task;

thus, you can continue to type email or edit code while seamlessly launching and controlling media streams. Over the years, Emacspeak has come to integrate with Emacs packages like **emms** as well as providing wrappers for **mplayer** and **alsaplayer** — collectively, these let you efficiently launch all types of media streams, including streaming video, without having to explicitly switch context.

In the mid-90's, Emacspeak started including a directory of media links to some of the more popular radio stations — primarily as a means of helping users getting started — Emacs' ability to rapidly complete directory and file-names turned out to be the most effective means of quickly launching everything from streaming radio stations to audio books. And even better — as the Emacs community develops better and smarter ways of navigating the filesystem using completions, e.g., package **ido**, these types of actions become even more efficient!

11 EBooks— Ubiquitous Access To Books

AsTeR — was motivated by the increasing availability of technical material as online electronic documents. While AsTeR processed the \TeX family of markup languages, more general ebooks came in a wide range of formats, ranging from plain text generated from various underlying file formats to structured EBooks, with Project Gutenberg leading the way. During the mid-90's, I had access to a wide range of electronic materials from sources such as O'Reilly Publishing and various electronic journals — The Perl Journal (TPJ) is one that I still remember fondly.

Emacspeak provided fairly light-weight but efficient access to all of the electronic books I had on my local disk — Emacs' strengths with respect to browsing textual documents meant that I needed to build little that was specific to Emacspeak. The late 90's saw the arrival of Daisy as an XML-based format for accessible electronic books. The last decade has seen the rapid convergence to **epub** as a distribution format of choice for electronic books. Emacspeak provides interaction modes that make organizing, searching and reading these materials on the Emacspeak Audio Desktop a pleasant experience. Emacspeak also provides an OCR-Mode — this enables one to call out to an external OCR program and read the content efficiently.

The somewhat informal process used by publishers like O'Reilly to make technical material available to users with print impairments was later formalized by BookShare — today, qualified users can obtain a large number of books and periodicals initially as Daisy-3 and increasingly as **EPub**. Book-

Share provides a RESTful API for searching and downloading books; Emacspeak module **emacspeak-bookshare** implements this API to create a client for browsing the BookShare library, downloading and organizing books locally, and an integrated ebook reading mode to round off the experience.

A useful complement to this suite of tools is the Calibre package for organizing ones ebook collection; Emacspeak now implements an **EPub Interaction** mode that leverages Calibre (actually sqlite3) to search and browse books, along with an integrated **EPub mode** for reading books.

12 Leveraging Computational Tools — From SQL And R To IPython Notebooks

The ability to invoke external processes and interface with them via a simple read-eval-loop (REPL) is perhaps one of Emacs' strongest extension points. This means that a wide variety of computational tools become immediately available for embedding within the Emacs environment — a facility that has been widely exploited by the Emacs community. Over the years, Emacspeak has leveraged many of these facilities to provide a well-integrated auditory interface.

Starting from a tight code, eval, test form of iterative programming as encouraged by Lisp. Applied to languages like Python and Ruby to explorative computational tools such as R for data analysis and SQL for database interaction, the Emacspeak Audio Desktop has come to encompass a collection of rich computational tools that provide an efficient eyes-free experience.

In this context, module **ein** — Emacs IPython Notebooks — provides another excellent example of an Emacs tool that helps interface seamlessly with others in the technical domain. IPython Notebooks provide an easy means of reaching a large audience when publishing technical material with interactive computational content; module **ein** brings the power and convenience of Emacs' editing facilities when developing the content. Speech-enabling package **ein** is a major win since editing program source code in an eyes-free environment is far smoother in Emacs than in a browser-based editor.

13 Social Web — EMail, Instant Messaging, Blogging And Tweeting Using Open Protocols

The ability to process large amounts of email and electronic news has always been a feature of Emacs. I started using package **vm** for email in 1990, along

with **gnus** for Usenet access many years before developing Emacspeak. So these were the first major packages that Emacspeak speech-enabled. Being able to access the underlying data structures used to visually render email messages and Usenet articles enabled Emacspeak to produce rich, succinct auditory output — this vastly increased my ability to consume and organize large amounts of information. Toward the turn of the century, instant messaging arrived in the mainstream — package **tnt** provided an Emacs implementation of a chat client that could communicate with users on the then popular AOL Instant Messenger platform. At the time, I worked at IBM Research, and inspired by package **tnt**, I created an Emacs client called **ChatterBox** using the Lotus Sametime API — this enabled me to communicate with colleagues at work from the comfort of Emacs. Packages like **vm**, **gnus**, **tnt** and **ChatterBox** provide an interesting example of how availability of a clean underlying API to a specific service or content stream can encourage the creation of efficient (and different) user interfaces. The touchstone of such successful implementations is a simple test — can the user of a specific interface tell if the person whom he is communicating with is also using the same interface? In each of the examples enumerated above, a user at one end of the communication chain cannot tell, and in fact shouldn't be able to tell what client the user at the other end is using. Contrast this with closed services that have an inherent *lock-in* model e.g., proprietary word processors that use undocumented serialization formats — for a fun read, see this write-up on Universe Of Fancy Colored Paper.

Today, my personal choice for instant messaging is the open Jabber platform. I connect to Jabber via Emacs package **emacs-jabber** and with Emacspeak providing a light-weight wrapper for generating the eyes-free interface, I can communicate seamlessly with colleagues and friends around the world.

As the Web evolved to encompass ever-increasing swathes of communication functionality that had already been available on the Internet, we saw the world move from Usenet groups to **Blogs** — I remember initially dismissing the blogging phenomenon as just a re-invention of Usenet in the early days. However, mainstream users flocked to Blogging, and I later realized that blogging as a publishing platform brought along interesting features that made communicating and publishing information **much** easier. In 2005, I joined Google; during the winter holidays that year, I implemented a light-weight client for Blogger that became the start of Emacs package **g-client** — this package provides Emacs wrappers for Google services that provide a RESTful API.

14 The RESTful Web — Web Wizards And URL Templates For Faster Access

Today, the Web, based on URLs and HTTP-style protocols is widely recognized as a platform in its own right. This platform emerged over time — to me, Web APIs arrived in the late 90's when I observed the following with respect to my own behavior on many popular sites:

1. I opened a Web page that took a while to load (remember, I was still using Emacs-W3),
2. I then searched through the page to find a form-field that I filled out, e.g., start and end destinations on Yahoo Maps,
3. I hit **submit**, and once again waited for a heavy-weight HTML page to load,
4. And finally, I hunted through the rendered content to find what I was looking for.

This pattern repeated across a wide-range of interactive Web sites ranging from AltaVista for search (this was pre-Google), Yahoo Maps for directions, and Amazon for product searches to name but a few. So I decided to automate away the pain by creating Emacspeak module **emacspeak-websearch** that did the following:

1. Prompt via the minibuffer for the requisite fields,
2. Consed up an HTTP GET URL,
3. Retrieved this URL,
4. And filtered out the specific portion of the HTML DOM that held the generated response.

Notice that the above implementation hard-wires the CGI parameter names used by a given Web application into the code implemented in module **emacspeak-websearch**. REST as a design pattern had not yet been recognized, leave alone formalized, and module **emacspeak-websearch** was initially decryd as being fragile.

However, over time, the CGI parameter names remained fixed — the only things that have required updating in the Emacspeak code-base are the

content filtering rules that extract the response — for popular services, this has averaged about one to two times a year.

I later codified these filtering rules in terms of XPath, and also integrated XSLT-based pre-processing of incoming HTML content before it got handed off to Emacs-W3 — and yes, Emacs/Advice once again came in handy with respect to injecting XSLT pre-processing into Emacs-W3!

Later, in early 2000, I created companion module **emacspeak-url-templates** — partially inspired by Emacs' **webjump** module. URL templates in Emacspeak leveraged the recognized REST interaction pattern to provide a large collection of Web widgets that could be quickly invoked to provide rapid access to the right pieces of information on the Web.

The final icing on the cake was the arrival of RSS and Atom feeds and the consequent deep-linking into content-rich sites — this meant that Emacspeak could provide audio renderings of useful content without having to deal with complex visual navigation! While Google Reader existed, Emacspeak provided a light-weight **greader** client for managing ones feed subscriptions; with the demise of Google Reader, I implemented module **emacspeak-feeds** for organizing feeds on the Emacspeak desktop. A companion package **emacspeak-webspace** implements additional goodies including a continuously updating ticker of headlines taken from the user's collection of subscribed feeds.

15 Mashing It Up — Leveraging Evolving Web APIs

The next step in this evolution came with the arrival of richer Web APIs — especially ones that defined a clean client/server separation. In this respect, the world of Web APIs is a somewhat mixed bag in that many Web sites equate a Web API with a JS-based API that can be exclusively invoked from within a Web-Browser run-time. The issue with that type of API binding is that the only runtime that is supported is a full-blown Web browser; but the arrival of native mobile apps has actually proven a net positive in encouraging sites to create a cleaner separation. Emacspeak has leveraged these APIs to create Emacspeak front-ends to many useful services, here are a few:

1. Minibuffer completion for Google Search using Google Suggest to provide completions.
2. Librivox for browsing and playing free audio books.
3. NPR for browsing and playing NPR archived programs.

4. BBC for playing a wide variety of streaming content available from the BBC.
5. A Google Maps front-end that provides instantaneous access to directions and Places search.
6. Access to Twitter via package **twittering-mode**.

And a lot more than will fit this margin! This is an example of generalizing the concept of a mashup as seen on the Web with respect to creating hybrid applications by bringing together a collection of different Web APIs. Another way to think of such separation is to view an application as a **head** and a **body** — where the **head** is a specific user interface, with the **body** implementing the application logic. A cleanly defined separation between the **head** and **body** allows one to attach *different* user interfaces i.e., **heads** to the given **body** without any loss of functionality, or the need to re-implement the entire application. Modern platforms like Android enable such separation via an Intent mechanism. The Web platform as originally defined around URLs is actually well-suited to this type of separation — though the full potential of this design pattern remains to be fully realized given today’s tight association of the Web to the Web Browser.

16 Conclusion

In 1996, I wrote an article entitled User Interface — A Means To An End pointing out that the size and shape of computers were determined by the keyboard and display. This is even more true in today’s world of tablets, phablets and large-sized phones — with the only difference that the keyboard has been replaced by a touch screen. The next generation in the evolution of **personal** devices is that they will become truly personal by being wearables — this once again forces a separation of the user interface peripherals from the underlying compute engine. Imagine a variety of wearables that collectively connect to one’s cell phone, which itself connects to the cloud for all its computational and information needs. Such an environment is rich in possibilities for creating a wide variety of user experiences to a single underlying body of information; Eyes-Free interfaces as pioneered by systems like Emacspeak will come to play an increasingly vital role alongside visual interaction when this comes to pass.

—T.V. Raman, San Jose, CA, September 12, 2014

17 References

- Auditory User Interfaces Klewer Publishing, 1997.
- Advice An Emacs Lisp package by Hans Chalupsky that became part of Emacs 19.23.
- Beautiful Code An overview of the Emacspeak architecture.
- Speech-Enabled Applications Emacspeak at CHI 1996.
- EWW Emacspeak extends EWW .
- In The Beginning Was The Command Line By Neal Stephenson

Emacs 24.4.50.1 (Org mode 8.2.6)