

Reading Spatial Hypertext

Catherine C. Marshall
Microsoft Research, Silicon Valley

Introduction

Critics of hypertext have long expressed skepticism about whether people could or indeed ever would sit down at a screen and read hypertexts. They may cite the more general problems of reading from the computer screen (see, for example, [25]) or they may place the blame partially at the feet of the authors (see, for example, Birkerts' indictment of hypertext fiction [8][2]). Some have even put forth usability arguments, that it's simply more efficient to *use* paper texts (see, for example, [24]), thereby avoiding the problematic idea of reading altogether.

In spite of their lack of consensus on why hypertext is not a good vehicle for reading, critics of hypertext usually agree on one key point: hypertext does not represent how *they themselves* want to read anything, be it a novel, a reference manual, or a magazine. On this point, they're emphatic, and they are quick to brush aside the exhortations of the hypertext theorists as the theorists wax enthusiastic in their postmodern embrace of the new medium. It seems unlikely that the two sides of the well-worn page versus pixel debate will ever resolve their differences.

Yet as time passes, reading is changing. It is difficult to ignore the transitions that are taking place in reading practice across the board: novels are written and read on cell phones; blogs command huge online audiences (and may even form the basis of writers' day jobs); the energy in newspaper publishing is channeled toward electronic distribution and reader interactivity; and several generations of e-book hardware and software have come and gone. It seems that people are reading hypertexts on the screen whether critics on either side like it or not. The revolution has been quiet and altogether bloodless.

The fact that people are reading on the screen without a second thought doesn't mean there's less to know. On the contrary, there seems to be even more to consider. But studies of how hypertext affects reading have been sparse. Since reading is regarded as a cognitive process, generally these evaluations have relied on controlled lab studies that use metrics related to efficiency and effectiveness (speed, comprehension, and the like) [11]; furthermore, they have tended to rely on physically obtrusive technologies like eye-trackers and, at the most extreme, functional Magnetic Resonance Imaging (fMRI) [13]. Because reading is so difficult to observe in practice (it's considered almost creepy to watch someone read), few detailed field studies of reading texts, let alone hypertexts, have been performed. Most field studies have sought to describe the territory *around* reading (see, for example, [1]), rather than focusing on reading itself. In short, most field studies of reading have done everything but looking at reading itself: they have constructed taxonomies of types of reading; examined specific reading-related interactions (e.g. annotation and clipping); and have taken reading as an invisible part of other activities like writing or collaborative work.

What might we see if we looked in greater detail at real people reading real hypertexts as an unselfconscious part of their everyday practice? I'm focusing this chapter more specifically on spatial hypertext—taking up spatial hypertext from a reader's perspective—but it is likely that the two topics I will discuss in the chapter apply in some measure to reading hypertexts in

general. Understanding how people read hypertexts and spatial hypertexts in the wild, given the inherent invisibility of reading and how complex these everyday activities are, may yield some important new conceptual and technological directions for those of us who are developing systems or writing hypertexts.

Recently a series of lab studies have evaluated various aspects of spatial hypertext as it is used in information triage, the rapid assessment of the salience of topically relevant documents in the context of performing a task. These lab studies have necessarily involved reading, since triage consists of a mosaic of variations on the general theme of reading—scanning, skimming, short periods of careful reading—interleaved with interactions we can measure (e.g. scrolling, mouse clicks, object creation and manipulation, and so on). For example, a study reported on in [30] examines the trade-offs among different kinds of structuring functionality; another triage study reported in [3] explores how the triage task devolves into window management as the reader shifts between spatial hypertext overviews of a document collection and the documents themselves). Yet another series of studies edges ever closer to reading territory by describing patterns of reading and structuring during triage [6], especially as they reflect the reader's interests [4]. The results of these studies, as well as the data we gathered in performing them, have influenced my understanding of the topics I discuss here, although I intend to maintain a primary focus on activities I have observed outside the lab.

Thus in this chapter, I'm going to discuss intuitions gathered from three sources: several diverse ethnographic projects; my own informal and undocumented observations of people reading, both online and on paper; and from the recordings made during the lab studies I referred to above that were designed to reveal particular phenomena associated with reading spatial hypertext.

I have two main topics I want to explore. One arises from watching people read and navigate; the other comes from watching people *not* read and *not* navigate.

Navigation, anticipation, and compulsion

I never quite understood what I saw when I watched people read hypertexts. They'd click on a link before they were quite ready to leave where they were and they'd go merrily on their way without completely reading the lexia they'd just left. Writers may have consciously constructed fine rhetorical bridges between the departure and arrival nodes [16], but it seemed as if there was something less conscious (and less self-conscious) going on for the reader. What was it?

Readers complained about fragmentation, yet it seemed to be their own damned fault. They'd click reflexively, not bothering to gather the context from the text that followed the link. Some even characterized this fragmentation (and the attendant anxiety) as a symptom of our interruption-prone lives [17]. We're moving too fast, they seemed to say. We just need to slow down and take it all in.

In fact, when we started our work on spatial hypertext, it was partly as a reaction to the all-or-nothing nature of link-following. Relationships between lexia—and the action of moving among them—seemed too binary, too unambiguous. Surely the relationships writers and readers perceived were not that cut-and-dried. So spatial hypertext gave us the ability to play with juxtaposition, emergence, and simultaneity [26]. It gave us a forum for arguing about the

meaning of proximity and the semantics of space [15]. And it didn't demand that readers hop away from one place to go to another; as Rosenberg observed, OR was replaced by AND.

This was all fine in theory, but it still didn't explain what I saw when I watched people read hypertexts (both as a surreptitious onlooker and as an authorized observer with signed consent forms in hand). Readers would click before they were ready to go. Jane Yellowlees Douglas had a compelling explanation of readers' relationships with hypertexts in terms of immersion, engagement, and flow [12][11], but what I saw still didn't fit any existing explanation and it seemed to have much to do with readers' perceptions of fragmentation.

Then I watched two other seemingly unrelated kinds of interaction as part of ethnographic projects. One project centered on observations of gamblers playing video poker (which is now a fully electronic experience) and the second on observations of people reading magazines (specifically, the *New Yorker*), both on paper and on a tablet computer [18]. These were detailed observations of human interaction with the two different forms of engagement.

What did they have to do with each other? Moreover, what did they have to do with reading hypertexts?

When people read works on paper, they engage in a variety of anticipatory gestures. The act of turning a page is more subtle than moving paper through air (as the simulation presented by [9] would suggest). Readers have many ways of getting ready to turn a page without actually doing so, and will often pause midway through the action in an unconscious bid for continuity. Or they can be momentarily distracted (say, in the case of the *New Yorker*, by a cartoon in the upper left corner or an ad along the side of the page), but seem to be able to resume reading without even noticing that they had been distracted.

In a traditional node-link hypertext, the complex gesture is compressed into a single action, clicking. We might also predict that spatial hypertext systems that use reduced document representations or metadata surrogates (like VIKI [22], Tinderbox [7], or VKB [29]) would suffer the same fate, but spatial hypertext with gestural interfaces (like Intergrams [26]) would not.

Now let's examine a second navigational phenomenon that is apparent if you really watch a person reading a magazine. A magazine layout can be thought of as hypertextual, even if a single article is linear. Insets, sidebars, photos, interleaved ads, cartoons, and other page design elements make it unlikely that a reader will approach linear text in a strictly linear manner. In fact, from the videotapes in the *New Yorker* study, it was readily evident that readers don't read linearly, even when they seem fairly immersed in the text. They jump ahead or look back in regular running text (and here I'm referring to actually looking backward or forward, not just the normal saccades reported in the psychology literature), so that their reading is seldom completely linear. The reader's engagement with the text goes beyond either the idealized immersive or active experience suggested by [23] or [2]. It is readily evident from the position of the reader's body that his or her attention shifts and returns; that attention is by turns deeper and shallower; and that readers jump over spans of text or return to text they've already read. Interviews with these readers later make us realize that these hypertextual shifts are not notable to (or sometimes even noticed by) the reader [18].

The reason these hops, skips, and jumps are so seamless to the reader seems connected to the reader's ability to physically focus his or her attention with little interruption. For example, to tighten focus onto a single running column of text, a reader might fold the magazine to render competing text and graphics invisible; conversely, opening up a magazine to a full two-page spread allows the reader to jump back and forth among different points in the text while still maintaining global context. By using a range of strategies, the reader is able to avoid the potential fragmentation inherent in both the binary jump between lexia (as is part and parcel of the everyday hypertext reading experience) and the discrete transitions between overview and lexia (as is part of the normal spatial hypertext reading experience).

So that's my first intuition: if we want to reduce the reader's sense of discontinuity, not only do we need to replace ordinary hypertext's OR with spatial hypertext's AND; we also need to minimize the disjunction between view and overview. And we need to more closely examine the role of anticipatory gestures in navigation.

Which brings us to video poker. By contrast to reading complexly organized material on paper, when people play video poker, the action may turn physically compulsive. If you spend time in an establishment with video poker machines¹, you can readily observe some of the players using stylized card-playing gestures when they interact with the machines. But unlike the real card games, the intervals between gestures are very short indeed. It's clear that the cards aren't on the screen long enough for the player to "read" the hand that's been dealt. Yet the player has the sensation of playing poker.²

In this case, the gesture has taken precedence over the results. Anticipatory motions are not encouraged (or indeed possible) when a player is using a video poker machine. Where else have I seen anticipatory motions subject to compulsive action? The page-turning buttons (as one would see in e-book devices) or link-clicking screen presentations are subject to all-or-nothing, compulsion-prone actions. In other words, a reader can lift the corner of a magazine page without really turning it, but if the same reader's hand is poised over the page-turning button or hovering over his mouse, it's likely that the entire gesture will trigger the actual event before the reader is ready.

Thus my first intuition can be amplified by reaching a better understanding of anticipatory gestures and how they can play a role in partial transitions.

Not reading and not navigating

Spatial hypertext goes a long way to remedying the tyranny of the click. But it still might cause inadvertent fragmentation as readers seek to minimize interruptive context switching and transitions. Here's how.

¹ In some states, video poker is part of the state lottery system, and it is fairly easy to encounter this behavior. Of course, most Indian reservations and Las Vegas-style casinos also have such machines. See [nyt article] for a fairly complete description of this sort of gambling.

² I say this because the players themselves distinguish video poker from so-called "line games," which are modeled after traditional slot machines. As far as I can tell, most of the game play is identical, although theoretically the player can pause long enough to contemplate a strategy; in practice, very few of them do.

As I mentioned earlier, Frank Shipman's group at Texas A&M University has built on our earlier information triage study [21] by performing a series of controlled studies in which participants complete several new hypothetical information triage tasks (see [3, 4, 5, 6, 30]). In each of these studies, there is at least one condition in which the participants use the VKB spatial hypertext system [29], so although they are lab studies, they are a source of data that may be mined to confirm (or call into question) intuitions about reading spatial hypertext.

Two behaviors that I have observed informally in a variety of situations were readily evident in the screen-capture data. These are behaviors significant for reading spatial hypertext in general, beyond the performance of the triage task. First, as I mentioned in the previous section, readers expend much time and effort on managing the transitions between VKB's spatial overview and the presentation of the space's constituent Web pages in browser windows and, in fact, seem to tire of doing so. This intuition formed the basis of the study described in [3]; the lab study data confirmed this informal observation.

Second, and of great salience to the point I'm about to make, readers are willing (or gradually become willing) to work from metadata-only reduced document representations (the surrogates we display in the spatial hypertext overview), even if the metadata seems less than adequate to meet the needs of the task at hand. In other words, if they think they can get away with it, people don't bother reading the actual lexia; rather, they only scan the surrogate.

The first observation, that readers grow frustrated with the management of screen real-estate, seems hardly surprising and to some extent is solvable. Bigger higher-resolution displays, multiple monitors, and secondary displays all address a reader's frustration by simply offering her more pixels. Furthermore, people who don't have the capital equipment budgets to spend on more display horsepower can also resort to that old-fashioned solution, a print copy of what they can't see on the screen. When they're not in the controlled study situation, it seems very likely that they'd find a workaround.

Yet I'm proposing – as Levy [17] and others before him (most notably Thorngate [31] and James [14]) have suggested – that it's an attention problem. As we noted in the previous section, attention is really what both the reader and the hypertext system are trying to manage.

And what of the second observation that people resort to metadata-only approaches to doing work? Is this necessarily a problem? After all, most of us read a large portion of our email exactly this way: We warily examine the metadata that our email application displays for us, and then we may delete multiple messages without even looking at the content. While many of us may recall inadvertently deleting important messages from our friends and colleagues, most of us hasten to add that the sender then alerted us to our mistake.

But good metadata is unusual. In fact, our wary examination of email metadata is tuned to look for bad metadata, metadata that signals deception, rather than judging the content by using the metadata at face value. We can usually tell when people have been driven by metadata-only approaches. Upon reflection, most of us can name an instance in which one of our papers has been cited based on its title alone; the paper itself clearly hasn't been read.

So spatial hypertext may not be the whole solution to the perils of clicking.

**Finding a better solution:
Reading surfaces and new modes of interacting with them**

The hardware and software requirements of a computer for reading have been noted by many researchers (for example, the XLibris project focused on satisfying these requirements [28]). But a focus on page-based reading isn't enough; reading – especially reading hypertexts – is more complicated than that. At least four other factors come into play:

1. the anticipatory gestures, partial actions, and focusing techniques we observe when we watch people read complicated paper forms like magazines;
2. the compulsive nature of simple physical actions like clicking or button-pushing;
3. the observation that human attention is difficult to manage, especially in the absence of spatial fixity and visually overwhelming situations; and
4. the growing realization that metadata often isn't rich enough, right enough, or sufficiently trustworthy to act upon.

We need to begin to look at a range of spatial hypertext solutions that incorporate both the basic physicality of reading and the realities of human attention. One way may be to provide the reader with a reading surface, a stable display that is dedicated to presenting the material that corresponds to the metadata shown on another display surface; such a reading-oriented display would use the principles that drove the creation of systems like XLibris (e.g. mobility and paper-like modes of interaction) and Microsoft Reader (in particular, readability) and would incorporate new modes of interacting with these reading surfaces that would acknowledge anticipatory gestures and the potential breakdown of binary actions. The lab study I referred to earlier ([3]) revealed that this type of multi-display configuration is by no means an automatic improvement; displays at different focal distances annoyed some participants. Yet there is something compelling about developing hardware specifically suited to reading and storing large personal collections of documents that can be synchronized with a spatial hypertext vehicle—look at the changes the iPod brought to how people listen to music.

There is no guarantee, however, that in every situation we will turn someone from a user into a reader. Can the metadata that non-reading users rely on be improved? We watch search engines develop better and better document surrogates; search engine developers know that a false click is apt to be a final click. It is wise for spatial hypertext developers to pay close attention to the surrogates they use to represent documents; a recent release of VKB demonstrates that lesson by using a richer representation of individual lexia (see, for example, Figure 2 in [5]), but there are other summarization techniques that may also be appropriate when the explicit metadata is either incorrect or untrustworthy.

Reading spatial hypertext need not continue to be invisible and secondary to structuring and manipulating spatial hypertext.

References

- [1] Adler, A., Gujar, A., Harrison, B., O'Hara K. and Sellen, A. A diary study of work-related reading: design implications for digital reading devices. In *Proceedings of CHI98* (Los Angeles, CA, April 1998), ACM Press, 241-248.
- [2] Adler, M. and van Doren, C. *How to Read a Book*. New York: Simon & Schuster, 1972.
- [3] Badi, R., Bae, S., Moore, J. M., Meintanis, K., Zacchi, A., Hsieh, H., Shipman, F., Marshall, C.C. Recognizing User Interest and Document Value from Reading and Organizing Activities in Document Triage. *Proceedings of the ACM Conference on Intelligent User Interfaces (IUI 2006)*. (Sydney, Australia, Jan. 29 – Feb. 1, 2006), pp. 219-225.
- [4] Bae, S., Badi, R., Meintanis, K., Moore, J.M., Zacchi, A., Hsieh, H., Marshall, C.C., Shipman, F.M. Effects of Display Configurations on Document Triage. *Lecture Notes in Computer Science*, Volume 3585, Sep 2005, Pages 130-143.
- [5] Bae, S., Kim, D., Meintanis, K., Moore, J.M., Zacchi, A., Shipman, F.M., Hsieh, H., and Marshall, C.C. Supporting Document Triage via Annotation-Based Multi-Application Visualizations. *Proceedings of ASIS&T 2008*. (Columbus, OH, October 24-29, 2008).
- [6] Bae, S., Marshall, C.C., Meintanis, K., Zacchi, A., Moore, J. M., Hsieh, H., Shipman, F. Patterns of Reading and Organizing Information in Document Triage. *Proceedings of ASIS&T 2006*. (Austin, TX, 3-8 November 2006).
- [7] Bernstein, M. *The Tinderbox Way*. Watertown, MA: Eastgate, 2007.
- [8] Birkerts, S. *The Gutenberg Elegies: The Fate of Reading in an Electronic Age*. Boston: Faber and Faber, 1994.
- [9] Chu, Y-C., Bainbridge, D., Jones, M. and Witten, I. Realistic books: a bizarre homage to an obsolete medium? In *Proceedings of JCDL 2004* (Tuscon, AZ, USA, June 7 - 11, 2004), ACM Press, pp. 78 – 86.
- [10] Crawford, W. Paper Persists: Why Physical Library Collections Still Matter. *Online Magazine*. January/February 1998, 42-48.
- [11] Dillon, A., Kleinman, L., Choi, G. O., and Bias, R. Visual search and reading tasks using ClearType and regular displays: Two experiments. *Proc. CHI'06*, ACM Press, New York, 2006, 503–511.
- [12] Douglas, J. Y. and Hargadon, A. The pleasure principle: immersion, engagement, flow. In *Proceedings of Hypertext 2000* (San Antonio, Texas, USA, May 30 - June 3, 2000), New York: ACM Press, pp. 153 – 160.
- [13] Hoeft, F., Ueno, T., Reiss, A. L., Meyler, A., Whitfield-Gabrieli, S., Glover, G., Keller, T. A., Kobayashi, N., Mazaika, P., Jo, B., Just, M. A., & Gabrieli, J. D. E. (2007) Prediction of

children's reading skills using behavioral, functional, and structural neuroimaging measures. *Behavioral Neuroscience*, 121, 602-613.

- [14] James, W. *The Principles of Psychology*. New York: Dover Publications, 1950.
- [15] Kaplan, N. and Moulthrop, S. Where no mind has gone before: ontological design for virtual spaces. In *Proceedings of ECHT'94* (Edinburgh, Scotland, June, 1994). New York: ACM Press, pp. 206 – 216.
- [16] Landow, G. Relationally encoded links and the rhetoric of hypertext. In *Proceedings of Hypertext '87* (Chapel Hill, North Carolina, United States 1987), New York: ACM Press, pp. 331 – 343.
- [17] Levy, D.M. *Scrolling Forward: Making Sense of Documents in the Digital Age*. New York: Arcade Publishing, 2001.
- [18] Marshall, C.C. and Bly, S. Turning the Page on Navigation. *Proceedings of JCDL'05*, Denver, CO, June 7-11, 2005, pp. 225-234.
- [19] Marshall, C.C. and Ruotolo, C. Reading-in-the-Small: a study of reading on small form factor devices. In *Proceedings of JCDL'02* (Portland, OR, July 14-18, 2002). ACM Press, New York, NY, 2002, 56-64.
- [20] Marshall, C.C. and Shipman F.M. III. Searching for the Missing Link: Discovering Implicit Structure in Spatial Hypertext. In *Proceedings of Hypertext '93* (Seattle, WA, November 14-17, 1993), New York: ACM Press, pp.217-230.
- [21] Marshall C. and Shipman, F. Effects of Hypertext Technology on the Practice of Information Triage. In *Proceedings of Hypertext '97* (Southampton UK, April, 1997), New York: ACM Press, pp. 124-133.
- [22] Marshall, C.C., Shipman, F.M. III, and Coombs, J.H. VIKI: Spatial Hypertext Supporting Emergent Structure. In *Proceedings of the European Conference on Hypermedia Technologies* (Sept. 18-23, Edinburgh, Scotland), New York: ACM Press, pp. 13-23.
- [23] Nell, V. *Lost in a Book: The Psychology of Reading for Pleasure*. New Haven: Yale University Press, 1988.
- [24] Nielson, J. The matters that really matter for hypertext usability. In *Proceedings of Hypertext '89* (Pittsburgh, Pennsylvania, USA, 1989), New York: ACM Press, pp. 239 – 248.
- [25] O'Hara, K. and Sellen, A. A comparison of reading paper and on-line documents. In *Proceedings of CHI '97* (Atlanta, Georgia, USA, March 22 - 27, 1997) New York: ACM Press, pp. 335 – 342.
- [26] Rosenberg, J. *Intergrams*. Watertown, MA: Eastgate, 1995.

- [27] Rosenberg, J. The Structure of Hypertext Activity. In *Proceedings of Hypertext '96*, New York: ACM Press, pp. 22-30.
- [28] Schilit, B. N., Price, M.N., Golovchinsky, G., Tanaka, K., and Marshall, C.C. As We May Read: The Reading Appliance Revolution. *IEEE Computer*, Vol. 32, No. 1, January 1999, pp. 65-73.
- [29] Shipman, F., Hsieh, H., Airhart, R., Maloor, P., and Moore, J.M. The Visual Knowledge Builder: A Second Generation Spatial Hypertext. In *Proceedings of Hypertext '01*, New York: ACM Press, 113-122.
- [30] Shipman, F.M., Hsieh, H. Moore, J.M., and Zacchi, A. Supporting personal collections across digital libraries in spatial hypertext. In *Proceedings of JCDL 2004* (Tuscon, AZ, USA, June 7 - 11, 2004), ACM Press, pp. 358 – 367.
- [31] Thorngate, W. On Paying Attention. In *Recent Trends in Psychology*. W J. Baker, et al., Eds. 1988, Springer-Verlag: New York. p. 247-263.