"I have been here before"

An investigation into spatial verbalizations in hypertext navigation

Charlotte van Hooijdonk

Tilburg University

Faculty of Arts

PO Box 90153

5000 LE Tilburg

The Netherlands

Tel: +31 13 466 2736

Fax: + 31 13 466 3113

E-mail: c.m.j.vanhooijdonk@uvt.nl

Alfons Maes

Tilburg University

Faculty of Arts

PO Box 90153

5000 LE Tilburg

The Netherlands

E-mail: Maes@uvt.nl

Nicole Ummelen

Tilburg University

Faculty of Arts

PO Box 90153

5000 LE Tilburg

The Netherlands

E-mail: Ummelen@uvt.nl

Biosketch

Charlotte van Hooijdonk is a PhD student at Tilburg University, the Netherlands. A part of her PhD thesis deals with spatial cognition in digital environments. Another part of her thesis, deals with information presentation and evaluation of multimodal information in the medical domain.

Alfons Maes holds a chair at Tilburg University, the Netherlands, and is head of the department Communication Science. His research interests include multimodal and digital communication, and document design. He published articles on discourse reference, document design, instructive discourse.

Nicole Ummelen is currently affiliated with the Communication Science department of the Faculty of Arts at Tilburg University. Her research focuses on information selection processes, especially in instructional documents and digital environments.

Abstract

We conducted an explorative study to investigate whether hypertext users use spatial expressions to conceptualize cognitive actions they are involved in, and how these expressions relate to the type of actions (executions versus evaluations) and the level of actions (syntactic vs. semantic vs. pragmatic). As a method, we used ten thinking aloud protocols of hypertext users who were navigating a website. The results of the protocol analysis indicate that spatial expressions were most frequent when users describe executions on the syntactic action level. The exploration allows us to critically assess the value of the thinking aloud method to shed light on the cognitive actions and processing involved in using hypertext.

Keywords

Human-computer interaction, spatial metaphor, navigation aids, thinking aloud protocols, hypertext use.

1 Spatial navigation in hypertext

1.1 *Effective navigation in hypertext: navigation maps*

An important goal in usability research is to investigate which factors influence effectiveness and efficiency in navigating hypertext. Studies focus for instance on individual differences in hypertext use (e.g., Campagnoni & Ehrlich, 1989; Vincente & Williges, 1988)and differences in users' tasks (e.g., Marchionini, 1989; Marchionini & Shneiderman, 1988). More recently, effects of the design of navigation aids were studied (e.g., Chen & Rada, 1996; Danielson, 2002; Dias & Sousa, 1997; Gupta & Gramopadhye, 1995; Maes, Van Geel, & Cozijn, 2005; McDonald & Stevenson, 1999). Hypertext users encounter at least two types of problems when they try to find information. The first problem is *disorientation* or *lostness*, embodied in the following three questions: Where am I?, Where do I have to go next?, and How do I get there? (Edwards & Hardman, 1989; Elm & Woods, 1985; Otter & Johnson, 2000). The second is *cognitive overhead*, which refers to the amount of cognitive resources necessary to successfully complete a task in hypertext (Conklin, 1987). A large number of navigation aids have been developed to prevent users from getting lost and to reduce the cognitive overhead: hierarchical navigation bars (bread crumbs), paging buttons, alphabetical content lists, history tools, (expandable) menus, back buttons, spatial maps, etc. These tools typically present information structure in a schematic or spatial way. For example, a contents list presents the topics in a plain or indented list, based on theme or alphabet; bread crumbs show the depth of the information in a left-to-right order on the screen (Lida, Hull, & Pilcher, 2003), and site maps offer many designs to present information structure spatially: top-bottom or left-right trees, spider structures etc [1].

This spatial character of navigation aids suggests that the concept of space is important for users who try to conceptualize hypertext structure and tasks. Hypertext research indeed suggests that spatial site maps are beneficial to users. For example, McDonald & Stevenson (1999) concluded that users navigated more efficiently with a spatial map than with a contents list, although they did not find any differences in terms of effectiveness. Dee-Lucas & Larkin (1995) did find a difference in effectiveness: participants recalled more nodes in the spatial map condition than in the alphabetical index condition. These differences were smaller when participants were asked to read the content of the hypertext in order to summarize it, which suggests that spatial maps facilitate *finding* information rather than *studying* it.

Although the results of these and other experiments suggest that spatial navigation help works better than other types of navigation instruments, there are reasons to question such a conclusion (Maes et al., 2005). First, the navigation maps as they are used in usability experiments differ substantially in the way they support the content, structure, task, and information access. They represent the hypertext's content and structure either globally or partially, they contain either 'labels only' or labels plus additional information, they either show the navigation history, or not. These differences themselves may result in substantial differences in effect. Second, all maps that were studied contain semantic labels. The presence of these labels does not allow us to determine the exclusive contribution of the spatial element to the usability of the maps, as the labels (or other content oriented design variables) are likely to also affect usability. Third, even if a spatial map proves beneficial, it is not

[1] For a survey of special site map examples, see www.cybergeography.org/atlas/web_sites.html

yet clear why, when and how it helps hypertext users. Does spatial design support the comprehension of the information structure? Does it enable users to set, monitor, and reach their goals more efficiently? Or does it mainly support the spatial-perceptual processes involved in hypertext use, like locating information, mentally replaying the navigating path, or transforming information into a spatial arrangement? These questions require more information about whether and how users conceptualize information environments and computer tasks spatially.

1.2 *The role of space in conceptualizing hypertext (tasks)*

Space is one of the most powerful tools for humans to conceptualize abstract thought (e.g., Gibbs, 2005; Lakoff & Johnson, 1980). We 'translate' time into space ("*we are entering a new age*"), connect good things with *up* (*"she is top!"),* bad things with *down* ("*I am feeling down*"), important things with near ("*things you carry in your heart*") and less important with *far* ("*a far-from-my bed show*"), etc. It is therefore not surprising that hypertext also created a large number of spatial metaphors, like 'lostness', 'hyperspace', or 'navigation'. The pervasive conceptual force of space does not explain, however, when and how users represent information and tasks spatially.

The premise of the spatial metaphor is that navigation in hypertext is conceptualized and understood on the basis of navigation in a physical environment. But what does this exactly mean? Does it enable us to conclude that the distance of two website pages is exactly 3,44 meters? Or does the spatial metaphor merely facilitate talking about hypertext and tasks in an intelligible way? Or is it something in between? In a lucid analysis, Boechler (2001) makes clear that space in hypertext can never be conceived of in purely literal or 'Euclidian' terms. Navigating from one page to another is not literally going 'deeper' in the site, going to the homepage is only metaphorically going up and the distance between pages cannot be expressed in metrical terms. This is not peculiar, as humans often conceive space in non-literal terms. But in her survey, Boechler makes clear that we have hardly any evidence on the working of spatial notions and metaphorical extensions in the minds of computer users.

There have been several attempts to apply spatial notions to hypertext use. For example, Shum (1990) applies *distance* and *direction* to hypertext, two elements which are known to be crucial in the study of how users conceive physical space, as it is clear from geography and psychology (Downs & Stea, 1973; Golledge, 1999; Taylor & Tversky, 1992a, 1992b; Tversky, 2003). The *whereness* of an object basically consists of a distance and a direction (Downs & Stea, 1977). According to Shum, each hypertext node has a certain *distance*, which can be quantified in absolute and relative terms, such as the number of nodes users have to visit, system response time, ease of returning to the previous node, or number of link traversals. *Direction* is defined as going forward and backward in the hypertext document. Although Shum tried to conceptually apply these definitions to hypertext, he did not investigate whether or not users really make use of these spatial concepts to mentally represent a hypertext environment. Similarly, Kim (1999) demonstrated the advantages of the familiar spatial metaphor of a shopping mall in accessing and using hypertext. Other researchers however contest the validity of the spatial metaphor. For example, Farris, Elgin & Jones (2002) concluded that the user's representation of a hypertext is non-spatial. They conducted an experiment in which participants had to explore a web site. The information on the web site was held constant, but the number of levels within the information structure varied. After exploring the web site, participants were asked to draw the web site's information structure. The analysis of these drawings indicated that the participants did not draw the spatial information structure of the web sites, but they drew conceptual relations between the information items instead. Therefore, Farris et al. (2002) concluded that the users' representation of a hypertext is non-spatial.

In sum, these studies seem to contradict each other at first glance. But this contradiction should be interpreted with great care, as the conclusions do not always seem to be reliable. Farris et al., for instance, offered their participants chunks with clear semantic relationships in a web site without any global spatial navigation aid, such as a site map, which makes it likely that participants are more guided by their prefixed semantic knowledge than the somewhat ad hoc and unsupported division of the chunks in different information levels.

1.3 *How spatial conceptualization of hypertext is investigated?*

Most usability studies draw conclusions about users' mental representations on the basis of performance results: the number of clicks, the recall of links or the quality of a drawing is assumed to reflect the adequacy of the representation. However, the relation between these dependent variables and the mental conceptualization of users is weak and always requires some type of subjective interpretation.

Users' representations can also be investigated by other methods. Thus, Maglio & Matlock (2003) asked experienced and inexperienced web users to judge the sensibility of metaphorical sentences on a seven point Likert scale. Sentences containing active motion (e.g.: *John went to a new web site today*) were rated as significantly more adequate than those describing a more passive user. Apart from that, they also asked hypertext users to verbalize what they do and think during their hypertext task. First, they asked users to execute free search tasks on the web. Afterwards, the participants were asked to tell what they just did. The transcripts of these interviews were coded to mark seven types of web actions. The results indicated that both novices and experts talked about their experiences in terms of physical motion and actions.

In this study, we will elaborate on this elicitation method in an attempt to get a more fine-grained view of how users conceptualize their task and use spatial conceptualizations. Unlike Maglio & Matlock, we are not only interested in whether spatial metaphorical expressions are used, but also in the proportion of these expressions in their verbal production and in the relationship between spatial conceptualizations and the type of cognitive action of the user: do spatial conceptualizations mainly show up in verbalizing low level actions (such as clicking or typing), or also in planning and monitoring the task?

Unlike Maglio and Matlock, we ask users to verbalize their actions online, while they are executing their task. We realize that we cannot assume that the resulting thinking aloud protocols directly tap cognitive processes. Furthermore, other drawbacks of this instrument may also apply here, for example, a user's conceptualisation of a hypertext may well be non-verbal, which would require a mental translation into a verbal form and thus additional cognitive processing. Yet, protocol analysis is a well-known tool for finding metaphorical language in interaction research (Kuhn, 1996) and we consider it to be a valuable tool for explorative work in this field, provided that the data are interpreted critically and carefully.

Before describing the set up and the results of this exploration, we will discuss different ways of categorizing actions of users who are navigating in a hypertext.

1.4 *Categorizing users' actions in hypertext*

A generally accepted overall model of hypertext use is not readily available (Chen & Rada, 1996). How theoreticians model hypertext use depends on the type of computer task (e.g. solving open or closed information problems), and the perspective (e.g. a learning or usability perspective). Yet, several researchers have attempted to classify user's actions while navigating through hypertext. In this section, we will discuss some ways of classifying hypertext actions. These classifications should enable us to reliably determine action levels in thinking aloud protocols.

Using hypertext can be seen as an interaction between two actors: the user and the hypertext system. The user initiates an action, the computer responds, the user evaluates the computer's response, etc. That way, all users' actions can be categorized in either *executions* or *evaluations*. This distinction corresponds to a distinction in Norman's (1998)'Action Cycle' of human-computer interaction (see figure 1). Users execute actions and evaluate the result by comparing the computer's reaction with their goal.

Figure 1. The Action Cycle (Norman, 1998, p. 47)

[INSERT FIGURE 1]

More fine-grained models of hypertext use are based on the idea that users have to execute these cognitive actions on different levels, ranging from physical (pushing buttons, waiting) to conceptual (anticipating on new information behind a link, comparing computer response to their real world tasks). These different levels may be compared to three levels involved in language processing: readers are assumed to build a surface, a propositional and a mental representation when reading a text (e.g., Fletcher & Chrysler, 1990; Johnson-Laird, 1983; Kintsch & Dijk, 1978). Hypertext users can be said to be mentally engaged in surface (i.e., executing physical, motional, perceptual actions), propositional (e.g. understanding the content and structure of hypertext) and mental/situational (e.g. planning and monitoring) actions.

This analogy has also been used in other HCI models. Shneiderman & Wald (1987) distinguish (1) *syntactic* programme specific knowledge (e.g., *what is the function of F1 in this programme*?), (2) *semantic* knowledge about the meaning of programme functions (e.g. *files can be opened, saved and copied*) and (3) *task* knowledge referring to the intended use of the programme in the real world (e.g. *write a letter*). In linguistic research, the latter knowledge type is usually referred to as '*pragmatic*' knowledge.

Figure 2. The syntactic/semantic model (Shneiderman, 1987, p. 43)

[INSERT FIGURE 2]

Other frameworks use different concepts, but reflect similar differences. Thus, Zapf, Brodbeck, Frese, Peters, and Pümper (1992) develop a taxonomy of errors in human-computer interaction, based on three task levels, first the level of planning, problem solving, and decision-making, second the level of flexible action patterns in which actions are regulated, third, the sensorimotor level of organizing automatic movement sequences without conscious attention. Norman's Action Cycle also reflects the syntactic and pragmatic task levels. Executions start with higher-level 'pragmatic' actions (planning and intending actions) and result in the low level ('syntactic') execution of actions.

Finally, Dillon (2004) has developed a model of hypertext use in which the three levels are represented. His TIME framework consists of four interactive elements, i.e., **T**ask, **I**nformation model, **M**anipulations skills, and visual **E**rgonomics (see figure 3).

Figure 3. The TIME framework (Dillon 2004, p.140)

[INSERT FIGURE 3]

Dillon's Task level implies the users' goal in the real world. The Information Model refers to the user's representation of the hypertext structure and content (semantic). Manipulations skills and visual ergonomics refer to motional and perceptual activities.

In our explorative study, we chose to depart from this tripartition as it can be conceptualized on the basis of these HCI-studies and linguistic research: *syntactic* (surface, perceptual, physical), *semantic* (propositional) and *pragmatic* (mental/situational) actions. The tripartition does not intend to make direct claims about the representations involved in using hypertext, but it should serve as a tool for analyzing actions involved in hypertext use. The analysis will explore whether and to what extent these three levels can serve as a useful analytical tool in distinguishing different action levels in hypertext.

2 Research method

We conducted an explorative thinking aloud study to investigate which type of actions (executions versus evaluations) and which levels of actions (syntactic vs. semantic vs. pragmatic) are expressed in spatial terms.

2.1 *Materials*

We collected ten thinking aloud protocols in two different usability studies. One study was set up to investigate the usability of a web site about the European Committee, the other study to investigate the usability of a medical web site. The web sites in the two studies were conventional web sites with many textual hyperlinks and several standard search facilities (a sitemap and a search function), see figure 3. In both studies, users were asked to perform simple search tasks in a hypertext (looking up the answers to factual questions), and to think aloud while executing these tasks. Participants' actions and verbalizations were recorded with Camtasia[2] camcorder software.

Figure 3. Screenshots from the homepages of the European Committee web site (left) and the medical web site (right)

[INSERT FIGURE 3]

[2] http://www.techsmith.com/products/studio/default.asp

2.1.1 The Medical web site study

Seven participants, four women and three men (ages between 27 and 57) participated in this study. Four of them were expert and three were novice Internet users. First they were asked to solve a digital version of the tower of Hanoi puzzle while thinking aloud. This common practice task was used to familiarize the participants with the thinking aloud method. Then they were presented with a Dutch medical website[3]. The participants were asked to answer two fact-finding questions and two explorative questions ("*What is the meaning of the word melatonin?*" and "*Which vaccinations do you need when traveling to Swaziland in Africa?*"). While executing these tasks they were asked to verbalize their thoughts. The experimenter kept silent. She only reminded participants to keep thinking aloud by saying 'keep talking' after a period of silence (Ericsson & Simon, 1993). The participants were allowed to take as much time as they needed to complete a task. If they could not find the answer to a question they were allowed to move on to the next task. After finishing a particular task, participants were instructed to go to the homepage of the web site.

The verbal protocols were elicited in three different ways: the thinking aloud method, the co-discovery method, and the instructing method. In the thinking aloud method, a novice and an expert user were asked to perform simple search tasks in a hypertext and to think aloud while executing these tasks. In the co-discovery method, two participants, either two novices or two experts, were asked to work together and to think aloud while executing the simple search tasks. In the instructing method, an expert user was asked to instruct a virtual novice (i.e., this participant was not present in the room) to find the answer on simple search tasks in a hypertext[4].

---

[3] www.medicinfo.nl

4 These different elicitation methods were used for another explorative goal of this same study: to see which differences the elicitation types would cause in the protocols. However, the differences were not expected to interfere with the research questions of the study described here, see the results.

2.1.2 The European Committee study

The objective of the European Committee study was to investigate how expert users conceptualize their actions in hypertext by using the thinking aloud method. Five subjects, four women and one man, between 20 and 25 years old participated in this study. First, the participants received the same practice task as the participants in the medical web site study. Next, they were presented with the homepage of the Dutch web site of the European Committee[5]. The participants had to find the answers of six fact-finding tasks, like "*Who is the current Dutch commissioner in the European Committee*?" The procedure was the same as in the medical web site study, but only individual thinking aloud protocols were collected here.

2.3 *Coding procedure*

The ten verbal protocols were coded with the program MAXQDA[6]. One analyst[7] coded three tasks of the five verbal protocols of the medical web site study: "*What does the word melatonin mean?*", "*What vaccinations do you need when traveling to Swaziland in Africa?*" and "*Go back to the home page*". This analyst also coded eight randomly chosen tasks of the five verbal protocols of the European Committee study. The total corpus consisted of 694 coded segments.

To determine the reliability of the analysis, a second analyst[8] also independently coded parts of the corpus on the basis of the same coding scheme that was defined first (see section 2.4). Differences between the two analysts were discussed, which resulted in some adjustments of the coding system. This procedure took place two times. The second analyst coded 128 utterances during the final

---

5 http://europa.eu.int/comm/index_nl.htm
6 http://www.maxqda.com/
7 The first author
8 The third author

analysis. Following standard practice, we qualify Cohen's κ as adequate if its value was higher than .70. The results indicated that both analysts highly corresponded in judging the utterances as executions or evaluations (Cohen's κ = .80; n = 128). The two analysts also highly corresponded in judging the utterances as spatial or non-spatial (Cohen's κ = .72; n = 578). Finally, both analysts corresponded in judging the utterances as syntactic, semantic, or pragmatic (Cohen's κ = .78; n = 100[9]).

2.4 *Coding system*

Each utterance was coded as belonging to a category of the following variables: spatial representation (yes or no), type of action (execution or evaluation), and level of action (syntactic, semantic, or pragmatic). In the following subsections, we will describe our criteria for coding the protocols.

2.4.1 Spatial or non-spatial verbalizations

Like Maglio & Matlock (2003), we were especially interested in the non-literal use of spatial expressions. It is evident that spatial expressions are used to describe visible actions on the screen ("*I am scrolling down on the screen*"), as there is hardly any other choice. So we only coded segments as spatial when they conceptualized non-literal and non-visible actions in terms of space. We distinguished three types of non-literal spatial expressions:

[9] In the final analysis, the analysts disagreed on several occasions on the categorization of the semantic vs. syntactic action level. These utterances were classified in a separate category, which may account for the high Kappa-score. See Results and Discussion sections.

- the user describes the next action as moving to or arriving at another place, by using expressions like ‘gaan naar’ (go to), ‘komen bij’ (arriving at), ‘zoeken bij’ (search at). ‘kijken bij[10]’ (look at).

  I am going back to the homepage

- the user describes his location as being on a particular place, by using expressions like ‘zijn in/bij/terug’ (be in/at/back), ‘zitten in/bij’ (sit in/at).

  This is not where I have to be

  I am in the main menu

- the user describes information as being somewhere in a physical location

  There is more information behind this hyperlink

2.4.2 Verbalizations of executions and evaluations

We distinguished two action types: execution and evaluation.

Verbalizations are considered to be executions if:

- the user verbalizes the action he performs, e.g.:

  I am clicking on GO

- the user verbalizes his intention to act, e.g.:

  I will go back to this item

Linguistic characteristics that indicate this type of utterances are verbs reflecting actions, such as click, go, scroll, read, type, or move.

Verbalizations are considered to be evaluations in the following cases:

- the expression of a user’s perception of elements in the environment, e.g.:

  A pop-up appears

- the expression of an evaluation of a user’s action, e.g.:

[10] Unlike in English, the Dutch preposition ‘bij’ has a clear locative interpretation.

I cannot click on this item

- the expression of a user’s evaluation of his task, e.g.:

  I think I have found the answer

- the expression of a user’s speculation on where information could be found, e.g.:

  Maybe at the hyperlink called organization

2.4.3 Verbalizations of syntactic, semantic, and pragmatic action levels

We defined three action levels: a syntactic level, a semantic level, and a pragmatic level. Utterances at the *syntactic level* concern users’ verbalizations of technical actions and perceptions of elements in the hypertext environment. We distinguished the following four types of syntactic verbalizations:

- Verbalizations of a user’s perception of hypertext elements on the computer screen, e.g.:

  I see three hyperlinks

- Verbalizations of a user’s coordination of actions with mouse and keyboard, e.g.:

  I am double clicking on this object

- Verbalizations of user’s assumptions on or questions about technical aspects of the hypertext, e.g.:

  Is this element clickable?

- Verbalizations of user’s technical actions, e.g.:

  I type in the word movement

Utterances at the *semantic level* concern a user’s understanding of the meaning of the hypertext’s content. We distinguished two types of verbalizations at this level:

- Verbalizations concerning the comprehension of the content on the screen, e.g.:

  This is an interview about the books he likes

- Verbalizations concerning semantic inferences made during reading and interpreting, e.g.:

  This is about the nations who are united in the European Union

Utterances at the *pragmatic level* concern users’ reflections on their real world goals. We distinguished four types of verbalizations on this level:

- Verbalizations reflecting questions about or relations with the search task, e.g.:

  What is the name of the book I am looking for?

- Verbalizations reflecting and evaluating screen results in terms of the search task goal, e.g.:

  I think I have found the answer

- Verbalizations reviewing the searching process, e.g.:

  Maybe if I search on a new version of Publication Magazine, I will find the answer

- Verbalizations of users' strategies concerning the search task, e.g.:

  I am going to search on seats

2.4.4 Not coded / not classifiable

Utterances that were not related to the task or took the form of fillers were not coded, e.g.: Ehh; Well; Wait. These utterances were left out of the analyses. Of a total of 694 utterances, 116 items were not related to the task (17 %). Furthermore, some cases for which no appropriate code or more than one code could be found were classified as in a separate category of unclear and mixed cases (see table 1 and discussion section).

3 Results

3.1 *Overall results*

Table 1 shows frequencies of action types, action levels, and spatial verbalizations in the complete set of coded utterances. The table shows that overall, evaluations occur more frequently than executions (χ^2 (1) = 39.91, p< .001), that the syntactic action level occurs more frequently than the semantic and pragmatic levels (χ^2 (3) = 369,34, p< .001), and that non-spatial utterances occur more frequently than spatial utterances (χ^2 (1) = 249.83, p< .001).

Table 1. Frequencies of occurrence of action types, action levels, and spatial verbalizations in 578 coded utterances from 10 verbal protocols (Scores are percentages of utterances; n = 578)

Action type	Execution	36.9
	Evaluation	63.1
Action level	Syntactic	56.9
	Semantic	4.8
	Pragmatic	26.3
	Unclear	11.9
Spatial verbalization	Spatial	14.8
	Non-spatial	68.4

Not all utterances could be coded unambiguously in one of the three action level, as the unclear cases show. A major case here was whether an utterance reflected a syntactic action level or a semantic action level. Utterances such as: "*I am reading the* *headings*", can be interpreted as a semantic action level because it refers to the semantic processing of the information on the screen. It can also be interpreted as a syntactic action level, if it is expressing the technical, low-level activity of reading from the screen. Given the relatively large number of these syntactic-semantic cases, we included them separately in Table 3 below.

3.2 *Spatial verbalizations related to action type and action level*

Is there a reliable relation between spatial expressions and action level or action type? A multinomial logistic regression analysis with participant, experience of the participant (novice vs. expert) and the participant's role (thinking aloud, co-discovery, and instructor) as independent variables and spatial verbalizations as dependent variable was used to test this question. Before discussing this main question, we will go into some effects of user and task characteristics that were intentionally or unintentionally varied in this study. The analysis showed significant effects for the user characteristics but not for the web sites: speaker (χ^2 (11) = 24.59, p< .05), experience of the user (χ^2 (1) = 4.23, p <. 05), and participant's role (χ^2 (2) = 7.08, p< .05). The web site (European Committee web site vs. medical web site) did not have an effect on the amount of spatial verbalizations (χ^2 (1) = 3.51, p = .61).

Experts tend to use more spatial expressions than novices (19% versus 11%), and co-discovery participants tend to use fewer spatial expressions (7%) than both thinking aloud participants (19%) and instructors (16%). Furthermore, individual users appear to differ somewhat in their tendency to use spatial expressions. The average percentage of spatial expressions per speaker is 17 per cent, and some individual speakers use fewer or very occasionally more spatial expressions. The following subsection does not go further into the reasons for these individual differences, but departs from the set of utterances that was collected and coded and tries to relate the occurrence of spatial expressions to action types and action levels.

3.2.1 *Action type*

Table 2 shows the frequencies of spatial verbalizations within executions or evaluations. Spatial verbalizations were most frequent when users were verbalizing executions

(χ^2 (1) = 34.10, p < .001).

Table 2. Spatial verbalizations related to executions and evaluations (Scores are percentages of utterances; n = 578)

	Executions (n = 213)	Evaluations (n = 365)
Spatial verbalizations (n = 99)	29.1	10.1
Non-spatial verbalizations (n = 479)	70.9	89.9

3.2.2 *Action level*

Table 3 shows the frequencies of spatial verbalizations at different action levels. Spatial verbalizations were found more frequently on the syntactic action level (χ^2 (3) = 25.98, p< .001) than on the other two levels.

Table 3. Frequencies of spatial verbalizations related to the action level (scores are percentages of utterances; n = 578)

	Perceptual/ syntactic level (n = 329)	Syntactic/ semantic level (n = 69)	Semantic level (n = 28)	Pragmatic level (n = 152)
Spatial verbalizations (n = 99)	22.2	0.0	0.0	17.1
Non-spatial verbalizations (n = 479)	77.8	100.0	100.0	82.9

3.3 *Spatial verbalizations in relation to other performance data*

In additional analyses we looked for relations between spatial verbalizations and other performance data, such as the type of search task, the correctness of the task outcomes and search times.

In order to see whether or not the number of spatial verbalizations depends on the type of search task, we divided the search tasks in both web sites in subtasks and tested whether the amount of spatial verbalizations was related to the specific subtask. In both web sites the spatial verbalizations differed depending on the subtask: European Committee web site (χ^2 (5) = 13.42, $p < .025$), medical web site (χ^2 (2) = 19.60, $p < .001$). In the European Committee web site, most spatial verbalizations occurred during a subtask that required participants to search for the chairman's name of the EU. In this instance, subjects typically had trouble finding the answer and had to search a large part of the site. In the medical web site, most spatial verbalizations occurred in a subtask where participants were instructed to return to the homepage of the web site. This -simple- subtask differed from other subtasks because there was no reference to a real world task of answering medical questions, only to the web site structure.

The amount of spatial verbalizations was not related to completing the task in a successful way (χ^2 (1) = 1.43, $p = .23$). The number of spatial verbalizations appeared to be related to the time it took participants to complete the fact-finding tasks (χ^2 (13)= 31.08, $p < .005$). Spatial verbalizations occurred most frequently within fact-finding tasks that were completed within 20 seconds.

4 Discussion

The exploration executed in this study served different purposes, which all merit to be discussed shortly. The main purpose was to investigate whether and how hypertext users spatially conceptualize cognitive actions they are involved in. Second, we wanted to discover whether thinking aloud protocols are a suitable method to shed light on the types of cognitive actions at work while using hypertext. Furthermore, we wanted to know whether the tripartition syntactic, semantic and pragmatic can be regarded as a suitable mould for the classification of these cognitive actions of hypertext users. Finally, we wanted to find legitimization for the widespread design decision to represent digital information spaces as spatial constructs, i.e. sitemaps, instead of verbal summaries.

The exploration suggests a clear-cut result: users predominantly use spatial expressions to conceptualize executions and low-level syntactic actions. A logical result as it sounds. Syntactic actions are more 'down to earth' and therefore more directly related to perceptual space than higher order actions. Also, executions are more goal oriented than evaluations, and therefore more suited to be conceptualized by the well known GOAL AS DIRECTION metaphor. Still, it is strange that there are so little spatial conceptualizations on the semantic and pragmatic level, although space is perfectly suited for conceptualizing information structures or plans and goals of users. It should be noted that not only the number of spatial semantic and pragmatic expressions is low, but also the overall proportion of semantic and pragmatic (as opposed to syntactic) expressions. Apparently, hypertext users are much more involved in shallow cognitive tasks (clicking, typing, reading, etc) than in deep processing (understanding content and structure, monitoring plans etc.). This is too premature a conclusion, however. The uneven distribution of syntactic, semantic and pragmatic expressions may be caused by the online character of the thinking aloud method. Thinking aloud users have to conceptualize their thoughts immediately on the fly, which may incite them to verbalize the here and now of each and every screen, instead of stepping back and talk about global structure or task progress.

For the low number of semantic (spatial) expressions, there may be an additional explanation, i.e. the narrow definition of semantic (as opposed to syntactic) expressions. At the outset we decided to only code segments as semantic when the understanding of the content of elements was verbalized. But much more expressions can be said to entail an awareness of a larger information structure on the part of the user. For example, when the user verbalizes the syntactic action "*I am going back to the link on commissioners*", there is a clear awareness of some structural organisation in the hypertext, which may be seen as a semantic attribute.

In sum, it is unclear to what extent the online character of the thinking aloud method overstressed the attention for low-level actions. Furthermore, the analysis showed that one expression can express different levels of cognitive actions, which again can be seen as a shortcoming of this method to answer questions like we were interested in. The explorative analysis showed that users indeed use space to talk about their task, but the thinking aloud method is apparently not the suitable instrument to shed light on the relation between spatial conceptualizations and levels of cognitive action. In order to determine and explain the beneficial nature of spatially organized navigation help, we will need more sophisticated techniques to further our understanding of the use of space in hypertext.

Acknowledgements

This research was partly conducted within the IMIX project Interactive Multimodal Output Generation (Imogen), sponsored by The Netherlands Organization of Scientific Research (NWO).

Notes

1. For a survey of special site map examples, see ww.cybergeography.org/atlas/web_sites.html
2. http://www.techsmith.com/products/studio/default.asp
3. www.medicinfo.nl
4. These different elicitation methods were used for another explorative goal of this same study: to see which differences the elicitation types would cause in the protocols. However, the differences were not expected to interfere with the research questions of the study described here, see the results.
5. http://europa.eu.int/comm/index_nl.htm
6. http://www.maxqda.com/
7. The first author
8. The third author
9. In the final analysis, the analysts disagreed on several occasions on the categorization of the semantic vs. syntactic action level. These utterances were classified in a separate category, which may account for the high Kappa-score. See Results and Discussion sections.
10. Unlike in English, the Dutch preposition 'bij' has a clear locative interpretation.

References

Boechler, P. M. (2001). How spatial is hyperspace? Interacting with hypertext documents: cognitive processes and concepts. *CyberPsychology & Behavior, 4*(1), 23-46.

Campagnoni, F. R., & Ehrlich, K. (1989). Information retrieval using a hypertext-based help system. *ACM transactions on information systems, 7*, 271-291.

Chen, C., & Rada, R. (1996). Interacting with Hypertext: A Meta-Analysis of Experimental Studies. *Human-Computer Interaction, 11*(2), 125-156.

Conklin, J. (1987). Hypertext: an introduction and survey. *IEEE Computer, 20*(7), 17-41.

Danielson, D. R. (2002). Web navigation and the behavioral effects of constantly visible site maps. *Interacting with computers, 14*(5), 601-618.

Dee-Lucas, D., & Larkin, J. H. (1995). Learning form electronic texts: effects of interactive overviews for information acces. *Cognition and Instruction, 13*(3), 431-468.

Dias, P., & Sousa, A. P. (1997). Understanding Navigation and Disorientation in Hypermedia Learning Environments. *Journal of Educational Multimedia and Hypermedia, 6*(2), 173-185.

Dillon, A. (2004). TIME: a framework for the design and evaluations of digital documents, *Desiging usable electronic text* (pp. 134 - 150). London: CRC Press.

Downs, R. M., & Stea, D. (1973). Cognitive maps and spatial behavior: process and products. In R. M. Downs & D. Stea (Eds.), *Image and Environment* (pp. 8-27). Chicago: Aldine Publishing.

Downs, R. M., & Stea, D. (1977). *Maps in minds: reflections on cognitive mapping*. London: Harper & Row.

Edwards, D. M., & Hardman, L. (1989). "Lost in hyperspace": cognitive mapping and navigation in an hypertext environment. In R. MacAleese (Ed.), *Hypertext: theory into practice* (pp. 90-105). Exeter: Intellect.

Elm, E. C., & Woods, D. D. (1985). *Getting lost: a case study in interface design.* Paper presented at the Human Factors Society, Santa Moncia, California.

Ericsson, K., & Simon, H. (1993). *Protocol analysis: verbal reports as data*. Cambridge: MITT Press.

Farris, J. S., Jones, K. S., & Elgin, P. D. (2002). Users' schemata of hypermedia: what is so "spatial" about a website? *Interacting with computers, 14*, 487-502.

Fletcher, C. R., & Chrysler, S. T. (1990). Surface forms, textbases, and situation models: recognition memory for three types of textual information. *Discourse Processes, 13*, 175-190.

Gibbs, R. W. J. (2005). Embodiment in metaphorical imagination. In D. Pecher & R. A. Zwaan (Eds.), *Grounding Cognition: The role of perception and action in memeory, language and thinking* (pp. 65-92). Cambridge: Cambridge University Press.

Golledge, R. G. (1999). Human wayfinding and cognitive maps. In R. G. Golledge (Ed.), *Wayfinding behavior: cognitive mapoing and other spatial processes* (pp. 5-45). Baltimore: John Hopkins University Press.

Gupta, M., & Gramopadhye, A. K. (1995). An evaluation of different navigational tools in using hypertext. *Computers and Industriual Engineering, 29*(1-4), 437-441.

Johnson-Laird, P. N. (1983). *Mental models*. Cambridge: Cambridge University Press.

Kim, J. (1999). An empirical study of navigation aids in customer interfaces. *Behaviour & Information Technology, 18*(3), 213-224.
Kintsch, W., & Dijk, T. A. v. (1978). Toward a Model of Text Comprehension and Production. *Psychological Review, 85*(5), 363-394.
Kuhn, W. (1996). *Handeling data spatially: spatializing user interfaces.* Paper presented at the International symposium on sptail data handling, Delft, The Netherlands.
Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago [etc.]: University of Chicago Press.
Lida, B., Hull, S., & Pilcher, K. (2003). *Breadcrumb navigation: a exploratory study of usage* [WWW]. Usability News. Retrieved 13 May, 2004, from the World Wide Web:
Maes, A., Van Geel, A., & Cozijn, R. (2005). Signposts on the digital highway: The effect of semantic and pragmatic hyperlink previews. *Accepted for Interacting with computers*.
Maglio, P. P., & Matlock, T. (2003). The conceptual structure of information space. In K. Höök & D. Benyon & A. J. Munro (Eds.), *Designing information spaces: the social navigation approach* (pp. 385-404). London: Springer.
Marchionini, G. (1989). Information seeking strategies of novices using a full text electronic encyclopedia. *Journal of the American Society for Information Science, 40*, 54—66.
Marchionini, G., & Shneiderman, B. (1988). Finding Facts vs. Browsing Knowledge in Hypertext Systems. *IEEE Computer, 21*, 70-80.
McDonald, S., & Stevenson, R. J. (1999). Spatial versus conceptual maps as learning tools in hypertext. *Journal of Educational Multimedia and Hypermedia, 8*(1), 43-64.
Norman, D. A. (1998). *The Design of everyday things*. London: The MITT Press.
Otter, M., & Johnson, H. (2000). Lost in Hyperspace: metrics and mental models. *Interacting with computers*(13), 1-40.
Shneiderman, B., & Wald, C. (1987). *Designing the user interface : strategies for effective human-computer interaction*. Reading, Mass. [etc.]: Addison-Wesley.
Shum, S. (1990). Real and virtual spaces: mapping from spatial cognition to hypertext. *Hypermedia, 2*(2), 133-158.
Taylor, H. A., & Tversky, B. (1992a). Descriptions and depictions of environments. *Memory & Cognition, 20*(5), 483-496.
Taylor, H. A., & Tversky, B. (1992b). Spatial Mental Models Derived from Survey and Route Descriptions. *Journal of Memory and Language, 31*, 261-292.
Tversky, B. (2003). Structures of mental spaces: how people think about space. *Environment and Behavior, 35*(1), 66-80.
Vincente, K. J., & Williges, R. C. (1988). Accommodating Individual Differences in Searching a Hierarchical File System. *International Journal Man Machine Studies, 29*, 647-668.
Zapf, D., Brodbeck, F. C., Frese, M., Peters, H., & Prümper, J. (1992). Errors in Working with Office computers: a first validation of Taxonomy of observed errors in a field setting. *International Journal Human-Computer Interaction, 4*(4), 311-339.