

How adolescents navigate Wikipedia to answer questions

Ladislao Salmerón and Raquel Cerdán

Department of Developmental and Educational Psychology & ERI Lectura

University of Valencia, Spain

Johannes Naumann

Goethe University Frankfurt am Main, Germany

Article in press: 06/11/2014. Please check the journal version:

Salmerón, L., Cerdán, R., & Naumann, J. (in press). How adolescents navigate Wikipedia to answer questions. *Infancia y Aprendizaje*.

Acknowledgements

This research was funded by the Spanish Minister of Economy and Innovation (Secretaría General de Universidades, EDU2011-25885). We thank Pilar García-Carrión and Gema Tavares for their assistance in data collection.

Correspondence concerning this article should be addressed to Ladislao Salmerón, ERI Lectura & Department of Developmental and Educational Psychology, University of Valencia, Avda. Blasco Ibáñez, 21. 46010 Valencia, Spain. E-mail: ladislao.salmeron@valencia.edu

Abstract

In one experiment, we explored how high-school students' use of hyperlink relevance cues while they navigate to answer questions from hypertexts. Current evidence has shown that students may navigate by either performing a deep semantic analysis of the relationship between the question and the existing hyperlinks, or by matching words in the question to words in the hyperlink labels. We focused on how students combine both cues during navigation, and how comprehension skills relate to the use of such cues. Our study revealed that 14 year old students ($N = 53$) selected hyperlinks by relying to a similar degree on both word matching and semantic overlap. Furthermore, when there was a conflict between an irrelevant link cued via word matching and a relevant link only cued through semantic overlap, students' comprehension skills facilitated their initial selection of an informative relevant link. To conclude, we discuss the implications of these results for current models of hypertext navigation.

Keywords: hypertext navigation, text comprehension, high-school students, comprehension skills

Many current instructional activities in high schools require students to use web documents to learn about particular topics. For example, in Webquests students are provided with a set of questions and they are instructed to use a set of Web resources to answer them. In that situation, a critical issue is to understand how the students select a hyperlink to find information useful to answer the question, because an incorrect decision might lead the students to irrelevant documents, which might impair their comprehension (Naumann, Richter, Christmann, & Groeben, 2008). The main goal of this paper is to shed light on which cues young students use to select relevant hyperlinks in web pages, and how the use of those cues is related to students' skills to comprehend printed text.

Relevance cues of hyperlinks

Which criteria students use to select hyperlinks that are relevant to fulfil their assignments? To determine the relevance of a hyperlink students can rely on different cues, that could be classified into visual and hyperlink label cues. On the one hand, visual cues are perceptual features that by social convention signal the relevance of a link, such as typography or position in a list of hyperlinks. Those cues may be particularly salient when hyperlinks are presented in a search engine results page (SERP). For example, Rouet, Ros, Goumi, Macedo-Rouet & Dinet (2011) found that the probability that middle school students would chose a link from a list increased when the link included a word from the search query in capitals, as compared to when the same link was displayed in lower case letters. In the same scenario, Pan et al. (2007) reported that graduate students mostly selected the top three pages from a SERP when performing simple search tasks (see also Salmerón, Kammerer & García, 2013). An analysis of the students' eye-movements revealed that most of them did not look at links below the third link. In sum, there is ample evidence that students can rely on perceptual signs as relevance cues. However, these cues are highly dependent on the interface used in the navigation system (e.g. a SERP can be displayed as a list or as a matrix of results, Kammerer & Gerjets, 2012), and they are usually not available while students navigate hypertext documents such as Wikipedia.

On the other hand, hyperlink label cues refer to the extent that link labels can be matched to the task demands. Two particular textual features are regularly used by students to assess text

relevance: semantic or word overlap with the question to be answered (Lewis & Mensink, 2012). The use of semantic overlap to select hyperlinks has been particularly stressed by semantic models of hypertext navigation. These models aim to explain and predict people's navigation across webpages while performing tasks such as buying a product in an on-line store or searching for information about a particular disease (CoLiDeS+ by Juvina & van Oostendorp, 2008; SNIF-ACT, by Fu & Pirolli, 2007). Although these models propose different mechanisms to predict people's navigation, they all share the notion that users assess the semantic relationship between their goals and some of the existing hyperlinks in a webpage, and that they base their decision about which hyperlink to follow on the degree of semantic overlap between the goal and the hyperlink. For example, in one of the most acknowledged models, the SNIF-ACT (Fu & Pirolli, 2007), users compare their goal with the semantic information conveyed in a link, by means of a semantically based spreading activation mechanism. Links that semantically match the goal demands would receive a high information scent. Thus, the perceived relationship, or information scent, is compared against the scent of the already assessed links. Once the users reach a point in which the expected gain from the rest of available links won't pay for the extra effort needed to assess them, they might click on the link with the highest information scent. In sum, from the point of view of these models users perform a quite systematic assessment of the available hyperlinks' relevance before selecting the most appropriate one.

A second relevance cue based on the hyperlink label is word matching, i.e. to decide that a link is relevant just because it includes a word from the search query. In this line, Wiley et al. (2009; Exp. 1) observed that most of the undergraduate students assessed selected links based on the presence of a keyword ('Mt. St. Helens') from their query ('What caused the eruption of Mt. St. Helens?') (see also Cerdán, Gilabert, & Vidal-Abarca, 2011).

In sum, prior evidence suggests that students use two different hyperlink label content cues to determine hyperlink's relevance. The extent to which word matching and semantic cues interact during question-answering tasks on hypertext is still an open question, because this issue remains largely unexplored in the empirical literature. An exception is the study by Keil and Kominsky (2013), that asked students from middle school to university to what extent a list of web pages were relevant

to solve a question. They found that students rated as most relevant web pages whose title semantically overlapped with the question. In addition, students rated as low relevant the pages that were only related by means of word matching, but not semantically. Participants in Keil and Kominsky (2013) study were not able to access the pages, and therefore their results may be difficult to generalize to web navigation.

Comprehension skills and hyperlink selection

By correctly balancing the use word and semantic cues students would navigate efficiently across links with relevant information, which in turn would improve their comprehension of hypertext (Salmerón, Cañas, Kintsch & Fajardo, 2005). Therefore, it is important to identify which individual characteristics help students to combine word and semantic cues to select links relevant for their learning goal. A potential explanation may focus on their skills to comprehend printed texts. Skilled comprehenders can implement important comprehension procedures that had been practiced extensively, such as identifying relevant ideas in the text and inferring the idea conveyed in a passage independently of the particular words used by the author (Perfetti, 2007).

In this line, Rouet et al. (2011) studied how high and low skilled students combine different relevance cues to select hyperlinks. Specifically, they compared label content cues (i.e. the link label referred to information relevant for the search query) and visual cues (i.e. the inclusion in the link of a keyword from the search query in capitals). The authors found that skilled comprehenders prioritized the content of the hyperlink label as a relevance cue over other competing visual cue (i.e. keyword in capitals). Low skilled students, on the contrary, prioritized the visual cue over the hyperlink label content. They interpreted this effect based on the fact that low skilled students need to focus most of their attention on word decoding at the expense of inferential processes. Rouet et al.'s (2011) study provides strong evidence that students use different relevance cues (visual versus hyperlink label) based on their comprehension skills. However, the study did not isolate the two different hyperlink label cues discussed above, i.e. word matching and semantic overlap. Indeed, predictions regarding the role of comprehension skills and the selection of a relevant link might vary depending if students prioritize semantic overlap or word match as main relevance cues. In the first case, we might expect

that comprehension skills have a direct impact on navigation, because they are essential to uncover the potential semantic relationships between links and students' tasks (Coiro & Dobler, 2007). Prior research evidenced that comprehension skills facilitate the selection of links relevant for the learning goals of young and undergraduate students (Naumann, Richter, Flender, Christmann, & Groeben, 2007; Naumann et al., 2008), and the selection of a cohesive navigation path by young students (Salmerón & García, 2011). Note that these studies did not control for the potential confounding of word matching between the students' goals and the hyperlink labels.

On the contrary, if students emphasize word matching we might predict that comprehension skills may not be critical for navigation, because they are not essential to match words in the link label to keywords in the students' task. Supporting this view, Cerdán et al. (2011) found that both low and highly skilled students initially accessed sections from an electronic text that contained a word match with some of the words in the to-be-solved question, even though those sections were not informative for the question. For example, students were presented with the question "*Which **adverse reactions** can the vaccine provoke*" and they tended to visit the following distracting location of the text that had not a semantic relation but contained a word overlap: "*Ask your doctor if you are under medical treatment or have had **adverse reactions** to the Flu vaccine. If these are very intense it may be dangerous to the Fetus in case of pregnancy.*"

In sum, the predicted impact of comprehension skills on navigation may vary according to the importance given by students to word and semantic cues during navigation.

Research questions

The aim of the present research was to test to what extent students prioritize word matching or semantic overlap cues to select relevant hyperlinks in a complex learning hypertext. To this end we ran an experiment in which high school students navigated a Wikipedia document to answer a set of questions about the French Revolution. Questions mentioned both a relevant link (which included information necessary to answer the question) and an irrelevant link (which did not include relevant information). The wording of the irrelevant link always occurred literally in the questions. Half of the

questions in addition literally contained the wording of the relevant link (literal-link questions). The other half, in contrast, contained only a paraphrase of the relevant link's wording, and therefore the question had a literal match only with the irrelevant link (paraphrased-link questions).

Predictions concerning students' navigation to select particular links vary according to the emphasis they give to semantic and word matching cues. Specifically, we will test predictions from three different views that either: a) emphasize semantic overlap cues over word matching for navigation (semantic view), b) stresses word matching as main relevance cue (word matching view), or c) assume that both cues might participate in hyperlink selections to a similar degree (semantic and word matching view). The first two views predict that students' will initially select links that are respectively semantically related or that share a word with the question (see Table 1). Only the semantic and word matching view predict an interaction between question wording (relevant link paraphrased or literally mentioned). In paraphrased-link questions both the relevant and irrelevant link would receive certain activation in regard to their relevance, either via a positive word match with the question wording (irrelevant link) or via a deep semantic assessment of its relationship with the question (relevant link). In literal-link questions, on the contrary, only the relevant link would receive additional activation by a deeper semantic analysis, and thus would be selected more often. In addition, the selection of those links would occur much more frequently than that of other available un-cued links (i.e. links not mentioned in the question).

-Insert Table 1 about here-

The second research question relates to the influence of comprehension skills in students' decisions to select particular links in a document (see Table 2). Again, only the semantic and word matching view predict an interaction with question wording. While comprehension skills may not be necessary when the relevant link is already cued to a higher degree than the irrelevant one (literal link questions), they may be required to assess relevance in paraphrase questions, where both links receive certain activation by means of two different relevance cues: word matching (irrelevant link) and a deep semantic process (relevant link).

-Insert Table 2 about here-

Method

Participants.

Fifty-three students from 2° ESO (Spanish equivalent of eight-grade) from three classes of a private school from a major Spanish city participated in the experiment (mean age 14.2 years, 51% female, mostly Caucasian)¹. On average, they had been using the Internet for 6 years ($SD = 2.13$). In their school curriculum, the topic of the French Revolution had not been covered yet, and thus most of them had low background knowledge on the topic (mean self-ratings = 1.02, $SD = 1.20$, on a scale from 0 –minimum- to 10 –maximum).

Apparatus.

We presented the questions and recorded participants' on-line behavior using an ad-hoc program written in Microsoft Visual Basic ® simulating a well-known web browser. The program recorded participants' clicks on hyperlinks and browser buttons, as well as visit times.

Materials.

Hypermedia.

We adapted a real hypermedia on the topic 'The French Revolution' addressed to high school students and distributed by the Spanish Minister of Education through its on-line educational repository (Tapia, 2004). The hypermedia mimicked the interface of Wikipedia ®. The main page included a navigable table of contents at the beginning of the page, as done in many Wikipedia pages (see Figure 1). The main page was composed of 1,559 words, and was organized in three main sections (causes, development and consequences of the Revolution). 8 out of the 10 subsections of the main page included 3 hyperlinks to additional Wikipedia-like documents (see section below). The linked documents were adapted from publicly available sources on the Internet, such as Wikipedia and other on-line repositories, and were shortened to have a similar length ($M = 288.96$ words, $SD = 43.59$).

-Insert Figure 1 about here-

Hyperlinks.

Eight subsections of the main page included 3 embedded hyperlinks. For each question students were told to visit a particular subsection where they could find relevant information (although they were free to navigate any of the subsections at any time). For each question only one of subsections' link lead to a page that included relevant and necessary information for the question (i.e. relevant link). A second link lead to a page that included irrelevant information for the question, and the hyperlink label was literally mentioned in the question (i.e. irrelevant cued link). A third link pointed to a page with irrelevant information for the question, and the hyperlink label was not mentioned in the question (i.e. irrelevant uncued link). The uncued link was included as a control measure to discard a potential random selection of links. In each subsection the order of the appearance of the different types of hyperlinks was varied, so that across questions the link mentioned first was relevant (25%), irrelevant-cued (41.7%) or irrelevant-uncued (33.3%), whereas the link mentioned at the third (and last) position was relevant (50%), irrelevant-cued (25%) or irrelevant-uncued (25%)².

Questions.

We constructed six retrieve and six integrate multiple choice questions, in an effort to include the most typical kind of questions students are presented to answer from documents (OECD, 2009). Retrieve questions required participants to select and extract specific pieces of information from a linked page. Integrate questions demanded students to connect several pieces of information through inferences within the main hypermedia and a linked page. For each question only one subsection of the main page included a relevant link to answer the question. Each of the eight subsections of the main page included a relevant link in at least one question, while four subsections contained a relevant link for two questions.

Half of the questions literally mentioned the label corresponding to the relevant link of the target subsection (i.e. literal-link questions). For example, in the question "Which significant

intellectual figure inspired the organization of the new political system in the French estate?” the students were pointed to the following section of the main Wikipedia page:

“New political system

On August 1789, the Assembly worked towards a new constitution that ultimately would transform the French estate into a constitutional monarchy. The Constitution of 1789 organized the country in three separate administrative powers: the executive, the legislative, and the judicial. This distinction was based on the theories of significant intellectual figures from the Enlightenment.

Two years after the beginning of the Revolution France celebrated elections that elected a new constituent assembly called the National Convention. Its main task was to develop a new Constitution. The members of the Convention decided to suppress the Monarchy and to establish a Republican system.”

For this particular question, the relevant link in the literal condition was: “significant intellectual figures”. The other half of the questions mentioned the label corresponding to the relevant link through a paraphrase (e.g. “important philosopher”). We wrote all questions in a way that the relevant link could be easily identified as the main clause, and the irrelevant link as a dispensable part. The relevant link could adopt several syntactic functions in the question (e.g. subject, direct object), but it never included secondary complements, verbs or functional words, because they would have been too vague in the context of the question to be considered as the main focus. In addition, all questions literally mentioned the label corresponding to an irrelevant cued link of the target subsection (e.g. “French state”). The irrelevant link mentioned in the question was apparently semantically relevant in the text, but unnecessary in the context of the question. The irrelevant link was mentioned always at the end of the question, and was part of a dispensable complement. Finally, an irrelevant uncued link was not mentioned in the question, but it was included in the target section (e.g. “Republican system”).

Comprehension skills test.

We used a standardized test of comprehension skills (TPC; Vidal-Abarca, et al., 2007), which is composed of two expository texts and ten multiple-choice questions per text. Questions target different comprehension processes as proposed by Kintsch (1998).

Procedure.

In a first session (approximately 45'), students completed the comprehension skills test in their own classroom. In a second session that took place after two days (approximately 40'), participants performed the experiment in the school computer room in small groups. Participants practiced with example materials, specifically a Wikipedia page on the Solar System with a similar structure to the experimental materials (i.e. the main page included a table of contents, each subsection only included three links...). They practiced with three questions, and during this phase, they were able to ask any question regarding the procedure. The experiment only started once all students confirmed that they felt confident with the task. They were told that their goal was to navigate a hypermedia document with the aim to answer some questions. They were instructed that in most cases the relevant information was included in a hyperlinked page, and they were encouraged to click on the links. They were free to access as many links as needed. There was no time limit to answer the questions, and all participants completed the task in the assigned time. On average, students spend 75.7 seconds to complete each question ($SD = 52.9$).

Results

In order to analyze students' navigation we computed two different measures. On the one hand, initial link selection represented if a link was selected or not as the student initial choice during the question-answering phase for each question. It represented initial link choices for each question, which would mostly rely on students' assessment of the potential utility of the link to answer the question. As this, they constituted our main variable to test the predictions of the semantic and word match models. On the other hand, subsequent link selections indicated if the relevant link was accessed after students had initially accessed an irrelevant link. It captured students' ability to identify the relevant link after an initially unfruitful access to an informative page.

Relevance cues and students' hyperlink selection.

First, we ran a repeated measures ANCOVA with hyperlink type (relevant, irrelevant cued, and irrelevant uncued) and matching (literal and paraphrase) as independent variables, comprehension skills as covariate, and initial link selection as dependent variable. Main effects resulted in significant differences for hyperlink type, $F(2, 100) = 37.65, p < .001, \eta^2 = .44$, but not for matching ($F(1, 50) = 0.08, p = .78$). Bonferroni post-hoc analyses showed that, after controlling for the level of comprehension skill, students tended to select the relevant link ($M = 24.80\%, SD = 16.58$) more often than the irrelevant cued link ($M = 19.60\%, SD = 15.14$), $p = .07$, and they initially selected both types more often than the irrelevant uncued link ($M = 4.80\%, SD = 07.21$) (in both cases $p < .001$).

A significant two-way interaction between hyperlink type and matching, $F(1, 50) = 13.21, p = .001, \eta^2 = .21$, qualified the main effect (see Figure 1). This interaction allowed us to test the predictions regarding our first research question on the role of content-based relevance cues in students' hyperlink selection. Bonferroni contrasts indicated that in literal-link questions the relevant link ($M = 31.10\%, SD = 24.52$) was selected initially more often than the irrelevant cued link ($M = 15.10\%, SD = 15.86$) ($p < .001$), and both links were selected initially more often than the irrelevant uncued link ($M = 3.50\%, SD = 7.21$) (in both cases $p < .001$). This pattern of results met the predictions of both a pure semantic view and a combined semantic and word view, as the two assume a deep semantic analysis of the relationship between the question goal and the available links to be operational. Contrasts for paraphrased-link questions provided further information. Supporting a semantic and word matching view, but inconsistent with a pure semantic view, in paraphrased-link questions the relevant ($M = 18.60\%, SD = 17.31$) and irrelevant cued link ($M = 24.00\%, SD = 20.91$) were chosen initially to a similar degree ($p = .46$), and both were selected initially more often than the irrelevant uncued link ($M = 6.10\%, SD = 12.25$) (in both cases $p < .001$).

We further analyzed to what extent students selected the relevant link after having accessed an irrelevant link –cued or uncued- initially. Indeed, most students' navigation decisions after an initial selection corresponded to a move from an irrelevant to the relevant link (57.29 %). Thus, we computed the percentage of questions in which students subsequently selected the relevant link from all possible cases (i.e. number of questions in which they had initially selected an irrelevant –cued or

uncued-link) by type of matching. Students eventually accessed the relevant link 87.31% ($SD = 47.05$) of possible cases in literal-link questions, compared to 57.85% ($SD = 43.81$) in paraphrased-link questions. The big difference between each condition in regard the initial probabilities of subsequent link selections prevented us from running a systematic comparison of means.

Nevertheless, an exploratory inspection of data suggested that after an initial access to an irrelevant node students found it easier to rely on a word matching cue than in a more demanding deep semantic processing to relate the relevant link to the question demands, and subsequently accessing it.

Comprehension skills and students' hyperlink selection.

In this section we outline results from the previous ANCOVAs involving effects of the covariate comprehension skills. Results are reported as far as they are relevant to the research question two, related to the influence of comprehension skills in students' decisions to select particular links in a document. Specifically, we report interaction effects of comprehension skills on initial link selection, and correlations of comprehension skills with initial link selection in the different experimental conditions that drive these interactions, where they are found.

Overall, there was a positive relation of comprehension skills with initial link selection, $F(1, 50) = 6.38, p < .05, \eta^2 = .11$, indicating that proficient readers responded to the question after accessing some link more often than less proficient readers³. This main effect however was qualified by two two-way interactions. First, there was a significant interaction of comprehension skills with matching, $F(1, 50) = 4.71, p < .05, \eta^2 = .09$. This interaction was due to a significant correlation of initial link selections with comprehension skills in the paraphrased-link conditions, $r(51) = .42, p < .05$, that did not occur in the literal-link conditions, $r(51) = .20, p = .17$. The interaction of comprehension skills with hyperlink type turned out significant as well, $F(2, 100) = 3.04, p = .05, \eta^2 = .06$. While there was a positive correlation of initial link selections in the relevant link conditions, $r(51) = .35, p < .05$, this correlation was only marginally significant in the irrelevant cued link conditions, $r(51) = .23, p = .10$, and insignificant in the irrelevant uncued link conditions, $r(51) = .05, p = .72$. This pattern of results is consistent only with a pure semantic view, which proposes an overall facilitation of comprehension skills on the initial selection of a relevant link.

However, those results were qualified by the three-way interaction of comprehension skills, matching and link type, that turned out marginally significant, $F(2, 100) = 2.45, p < .10, \eta^2 = .03$. This three-way interaction was, as predicted by the semantic and word matching view, due to the fact that within the paraphrase conditions, there was a strong interaction of hyperlink type and reading skills, $F(2, 100) = 7.35, p < .01, \eta^2 = .23$. This interaction emerged from a strong correlation of comprehension skills with the initial selection of relevant links, $r(51) = .51, p < .001$, that did not occur for both irrelevant cued links, $r(51) = .22, p = .12$, and irrelevant uncued links, $r(51) = -.08, p = .59$ (see Figure 2, lower panel). Within the literal conditions, in contrast, no interaction of comprehension skills and hyperlink type occurred whatsoever, $F(2, 100) = 0.03, p = .97$ (see Figure 2, upper panel). Within these conditions, there was no association of comprehension skills with initial link selection, independently of hyperlink type. For relevant hyperlinks, the correlation amounted to $r(51) = .09, p = .51$, for irrelevant cued links it was $r = .16, p = .27$, and for irrelevant uncued links it was $r(51) = .22, p = .13$. A deep involvement of comprehension skills to initially select a relevant link was only necessary when both the relevant and irrelevant links were cued to a rather similar degree (paraphrased-link questions), but not when the relevant link was already cued to a substantially higher degree (literal-link questions).

-Insert Figure 2 about here-

To help interpreting the three-way interaction we depicted it in Figure 2. To do so we rescaled the continuous variable 'comprehension skill' to a dichotomous variable, only for representation purposes. Note that the regression analyses with this variable respected the continuous variable, and therefore didn't involve any grouping of the students. Figure 2 displayed two theoretical vectors, one for poor comprehenders (1 *SD* below the mean on the comprehension skills test) and another one for good comprehenders (1 *SD* above the mean). Those vectors represented a prototypical poor/good comprehender based on a projection of the whole data set. Interestingly, as predicted by the semantic and word matching view, in literal link questions, poor comprehenders as well as good comprehenders were in most instances able to discard the irrelevant link (note that in these questions, poor

comprehenders could identify the relevant link through word matching as well, see Figure 2, upper panel). In paraphrase questions, however, poor comprehenders were overtaxed with the task of discarding the irrelevant, but cued link (see Figure 2, left-hand side of the lower panel). And even good comprehenders, i.e. student scoring one standard deviation above average in comprehension skill, chose the relevant (but uncued) and the irrelevant (but cued) to an identical degree (see Figure 2, right-hand side of the lower panel). It is only students who score more than one standard deviation above average that will favour the relevant (but uncued) link over the irrelevant (but cued) link in a majority of instances.

Finally, in an attempt to replicate the positive effect of navigation on task performance (Naumann et al., 2007, 2008), we run a set of partial correlations between students' initial selection of the relevant link and success on the literal-link and paraphrased-link questions, controlling for comprehension skill level. Significant correlations indicated a positive relationship between initially selecting the relevant link (i.e. percentage of questions in which the relevant link was clicked first) and success in both literal-link and paraphrased-link questions ($r(49) = .33, p = .02$, and $r(49) = .40, p < .01$, respectively).

Discussion

Results from our study shed light on how high-school students combine hyperlink label content cues of relevance (word matching and semantic overlap) to navigate in complex hypermedia scenarios such as Wikipedia. Prior research on hypertext navigation report the isolated effects of either semantic (e.g. Fu & Pirolli, 2007) or word matching cues (Wiley et al., 2009). The present research adds to this literature in showing that young students combine both word and semantic cues to select hyperlinks, and that the use of semantic overlap cues critically depends upon students' comprehension skills. Specifically, in this section we address the results related to two research questions: the first related to the role of relevance cues in students' hyperlink selection and the second with the purpose of testing the role of comprehension skills in students' decisions to select particular links in a document.

Relevance cues and hypertext navigation

Our results from initial link selections support the assumption that students combine both word matching and semantic overlap to assess hyperlinks' relevance and to navigate. When both cues point to a particular link, this is selected more often than competing links that are only activated via one cue. This is particularly evident in literal-link questions in which students face two competing links, one relevant link that is cued via word matching and semantic overlap, and an irrelevant link that is only cued via word matching. The fact that students select less often an irrelevant link that is cued by a word matching can not be explained by a word matching view. However, the fact that students select more often the relevant link in this scenario could be interpreted as evidence that students prioritize semantic cues over word matching, as proposed by a semantic view of link relevance.

Additional evidence that can only be explained by a semantic and word matching view comes from the results in the paraphrase-link questions. In this case, the relevant link is only cued via semantic overlap (i.e. students need to paraphrase the link to match it with the question goal), and an irrelevant link is only cued via word matching. Our results show that in this scenario young students erroneously select initially the irrelevant link cued via word matching quite often, i.e. to a similar degree than the relevant link. It is important to note that in this scenario students largely ignore other available uncued links, which suggests that the selection of links was not random. This pattern of results indicate that both word matching and semantic cues operate in combination and students, in principle, don't prioritize one over the other in scenarios in which the available hyperlinks are activated only via a single cue. The semantic view can not account for the fact that in these situations students preferred in approximately half of the cases to select a link that was only cued via word matching rather than a link only primed via semantic overlap. Similarly, the word matching view can not explain the fact that in many occasions students selected links cued via semantic overlap when a competing link cued via word matching was also available.

Comprehension skills, relevance cues and hypertext navigation

Results replicate prior studies indicating that selecting a relevant link facilitates to answer correctly the retrieve and integrate questions (e.g. Naumann et al., 2007, 2008). Thus, in complex

scenarios where several relevance cues are presented, young students must critically assess the available cues to determine the optimal navigation path. In such cases just relying on word matching might lead students to an erroneous path. Emphasizing semantic overlap, by contrast, is a flexible way to relate the task demands to the available link resources. Why then students at all use a less efficient cue such as word matching over semantic overlap to assess link relevance to any substantial degree?

Our results reveal that only students that possess very high comprehension skills may to some extent prioritize semantic overlap over word matching to identify relevant links. In contrast, poor comprehenders prioritize word matching cues over semantic overlap in situations where the semantically overlapping link is not as well literally matched. This result is in line with the general finding that students who are good comprehenders in the print medium are better at selecting relevant information (Naumann et al., 2007, 2008), discarding irrelevant information (Cerdán et al., 2011), and crossing the selected information into a semantically cohesive order (Salmerón & García, 2011).

In addition, our results extend Rouet et al. (2011) findings in showing that good comprehenders prioritize more informative link content cues over other distracting information (link typography). When the task contains two competing relevance cues (word match and semantic overlap), students who are good comprehenders are to some extent able to analyze the wording of a paraphrased relevant hyperlink so that it becomes apparent whether or not it will provide information useful for the task at hand. Note that, in any case, across tasks they still have trouble inhibiting irrelevant word cues in many cases. Bad comprehenders, on the contrary, could be overwhelmed by the word decoding demands of the task. For that reason, they might rely on less demanding relevance cues, such as word matching or link typography.

Future research however will have to clarify what are the specific mechanisms by which comprehension skills enter web navigation. At least two different models are obvious: An inference generalization model and a working memory load model. The inference generalization model is the one we assumed in our analysis in this article. This model presupposes that basically the same inferencing and prediction mechanisms that are required in traditional text processing enable proficient readers to make good predictions concerning the usefulness of a hyperlink (cf. Coiro &

Dobler, 2007). It may, however, be that this is not the case, as would be assumed by a working memory load model. Such a model would assume that good readers' working memory is not as much burdened with text processing as poor readers' working memory (Just & Carpenter, 1992). As a consequence, good readers have available cognitive resources to judge the usefulness of a hyperlink. The kind of inferences used in these judgments, however might be of a different nature than those employed in reading traditional text. Although there is some evidence that basic comprehension skills contribute to the selection of relevant links independently from working memory capacity (Naumann et al., 2008), future research will have to sort out in which way comprehension skills on the text level and individual working memory contribute to navigation, and through navigation to comprehension of web-based texts.

Cognitive models of hypertext navigation

According to SNIF-ACT (Fu & Pirolli, 2007), users compare their goal with the semantic information conveyed in a link, by means of a semantically based spreading activation mechanism. This mechanism could easily incorporate word matching as an additional relevance cue, for example, by assuming automatic activation of not only word meanings but also word surface. The model proposes that the perceived relevance is compared to the scent of the already assessed links. This might explain why an irrelevant link cued via a word matching is selected quite often in literal-link questions: students could have identified it as a link with enough information scent as compared to the other previously read links in that scenario. Future research might analyze students' eye-movements data in relation to links' relevance cues to support this view.

Instructional applications

Our study stresses the importance of selecting relevant links to improve students' comprehension. Although we do not focus on the consequences of selecting irrelevant links, we can expect that navigating through irrelevant paths by using word matching might also have detrimental consequences in the midterm. Different individual and task characteristics may prevent students from relying on word matching cues without any active control of its effectiveness. For example, prior background knowledge on the text topic might be necessary to suppress the activation of irrelevant

information in a particular learning context. In this line, Rouet, et al.(2011) have found that providing young students with a summary of the relevant background topic information before searching the Web prevents them from relying on superficial cues, such as typography, when assessing link relevance. Future research may explore how prior background knowledge may participate on the rejection of irrelevant links primed by word matching.

Limitations and future research

Our study comes with certain limitations that should be addressed in future research. First, the questions used mention the relevant link in the question in the first place (because it is either a subject or a direct object), whereas the irrelevant link is always mentioned at the end (because it is a dispensable complement). This issue may raise concerns about the potential primacy and recency effects of link presentation order on students' use of hyperlinks. Future research should investigate this issue, while controlling for the potential confounding of the syntactical function played by the relevant and irrelevant links in the questions. Second, our study doesn't focus on complex hyperlink selection patterns, such as that of a student that for a particular question first selects the relevant link, but subsequently selects the irrelevant mentioned and the irrelevant non-mentioned link. Understanding such complex interactions demands for additional methodologies, such as think aloud protocols. Finally, we have only collected visit times from the log files, and these only represent the time students spend with a particular screen open. A fine grained analysis of reading times analysis, such as those obtained by means of eye-tracking, would allow to qualify the processes underlying students' hyperlink selection. For example, one student may have clicked on a relevant link after a close reading of the initial paragraph, whereas other may have done that after skimming a few words. Our current efforts go in this direction.

References

- Cerdán, R., Gilabert, R., & Vidal-Abarca, E. (2011). Selecting information to answer questions. Strategic individual differences when searching texts. *Learning and Individual Differences*, 21, 201-205. doi:10.1016/j.lindif.2010.11.007

- Coiro, J., & Dobler, E. (2007). Exploring the online comprehension strategies used by sixth-grade skilled readers to search for and locate information on the Internet. *Reading Research Quarterly, 42*, 214-257. doi:10.1598/RRQ.42.2.2
- Fu, W-T., & Pirolli, P. (2007). SNIF-ACT: A cognitive model of user navigation on the World Wide Web. *Human-Computer Interaction, 22*, 355-412. doi:10.1080/07370020701638806
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*, 122-149. doi:10.1037/0033-295X.99.1.122
- Juvina, I., & van Oostendorp, H. (2008). Modeling semantic and structural knowledge in web navigation. *Discourse Processes, 45*, 346-364. doi: 10.1080/01638530802145205
- Kammerer, Y., & Gerjets, P. (2012). Effects of search interface and internet-specific epistemic beliefs on source evaluations during web search for medical information: An eye-tracking study. *Behaviour & Information Technology, 31*, 83-97. doi:10.1080/0144929X.2011.599040
- Keil, F.C., & Kominsky, J.F. (2013). Missing links in middle school: developing use of disciplinary relatedness in evaluating Internet search results. *PLoS ONE, 8*: e67777. doi:10.1371/journal.pone.0067777
- Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. New York: Cambridge University Press.
- Lewis, M. R., & Mensink, M. C. (2012). Prereading questions and online text comprehension. *Discourse Processes, 49*, 367-390. doi:10.1080/0163853X.2012.662801
- Naumann, J., Richter, T., Christmann, U., & Groeben, N. (2008). Working memory capacity and reading skill moderate the effectiveness of strategy trainings in learning from hypertext. *Learning and Individual Differences, 18*, 197-213. doi:10.1016/j.lindif.2007.08.007
- Naumann, J., Richter, T., Flender, J., Christmann, U., & Groeben, N. (2007). Signaling in expository hypertexts compensates for deficits in reading skill. *Journal of Educational Psychology, 99*, 791-213. doi:10.1037/0022-0663.99.4.791

- Organisation for Economic Co-operation and Development (OECD). (2009). PISA 2009 assessment framework - Key competencies in reading, mathematics and science. Retrieved from <http://www.oecd.org/dataoecd/11/40/44455820.pdf>
- Pan, B., Hembrooke, H., Joachims, T., Lorigo, L., Gay, G., & Granka, L. (2007). In Google we trust: Users' decisions on rank, position, and relevance. *Journal of Computer-Mediated Communication*, *12*, 801–823. doi: 10.1111/j.1083-6101.2007.00351.x
- Perfetti, C. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, *11*, 357-383. doi:10.1080/10888438.2011.564244
- Rouet, J.-F., Ros, C., Goumi, A., Macedo-Rouet, M. & Dinet, J. (2011). The influence of surface and deep cues on primary and secondary school students' assessment of relevance in Web menus *Learning and Instruction*, *21*, 205-219. doi:10.1016/j.learninstruc.2010.02.007
- Salmerón, L., Cañas, J.J., Kintsch, W., & Fajardo, I. (2005). Reading strategies and hypertext comprehension. *Discourse Processes*, *40*, 171 – 191. doi:10.1207/s15326950dp4003_1
- Salmerón, L., & García, V. (2011). Comprehension skills and children's navigation strategies in hypertext. *Computers in Human Behavior*, *27*, 1143-1151. doi:10.1016/j.chb.2010.12.008
- Salmerón, L., Kammerer, Y., & García-Carrión, P. (2013). Searching the Web for conflicting topics: page and user factors. *Computers in Human Behavior*, *29*, 2161–2171. doi:10.1016/j.chb.2013.04.034
- Tapia, J. (2004). *La Revolución Francesa [The French Revolution]*. Retrieved from <http://ntic.educacion.es/w3/eos/MaterialesEducativos/bachillerato/historia/revfran/index.htm>
- Vidal-Abarca, E., Gilabert, R. Martínez, T., Sellés, P., Abad, N., & Ferrer, C. (2007). *TPC. Test de Procesos de Comprensión* (Test of Comprehension Strategies). Madrid, ICCE.
- Wiley, J., Goldman, S., Graesser, A., Sanchez, C., Ash, I. & Hemmerich, J. (2009). Source evaluation, comprehension, and learning in internet science inquiry tasks. *American Educational Research Journal*, *46*, 1060-1106. doi: 10.3102/0002831209333183

Footnotes

¹ One case was excluded from the analyses including comprehension skills as a covariate, which proved to be an outlier. Checks of residuals had revealed non-normal residuals distributions. Removal of that case normalized the distribution.

² To rule out a potential effect of the position of a link on navigation we analyzed students' initial link selection as a function of link position (first, second or third). Results showed no differences based on link position, $F < 1$.

³ On average, students clicked on hyperlinks in 60% of the questions ($SD = 22.71\%$). Note that students were not forced to click hyperlinks. Thus, in some cases they could have answered after reading the main Wikipedia page, or just from memory. 5 participants only read the main Wikipedia page without clicking any hyperlink.

Table 1

Predictions of the three models tested in regard to initial clicks to relevant and irrelevant hyperlinks as a function of question wording

Question wording	Semantic view	Word matching view	Semantic and word matching view
Question paraphrases relevant link and literally mentions irrelevant link	Irrelevant < Relevant	Irrelevant > Relevant	Irrelevant = Relevant
Question literally mentions both relevant and irrelevant link	Irrelevant < Relevant	Irrelevant = Relevant	Irrelevant < Relevant

Table 2

Predictions of the three models tested in regard to the relation between comprehension skills and initial click of a relevant link as a function of question wording

Question wording	Semantic view	Word matching view	Semantic and word matching view
Question paraphrases relevant link and literally mentions irrelevant link	Facilitates	No effect	Facilitates
Question literally mentions both relevant and irrelevant link	Facilitates	No effect	No effect

Figure captions.

Figure 1: Screenshot of the main Wikipedia page used in the study.

Figure 2: Three-way interaction of question type, hyperlink type and comprehension skills on initial hyperlink selection.



Revolución Francesa

La **Revolución Francesa** fue el cambio político más importante que se produjo en Europa, a finales del siglo XVIII. No fue sólo importante para Francia, sino que sirvió de ejemplo para otros países. Esta revolución significó el triunfo de un pueblo pobre, oprimido y cansado de las injusticias, sobre los privilegios de la nobleza feudal y del estado absolutista que habían existido durante la época del Antiguo Régimen.

Contenido [ocultar]

- 1 Causas de la Revolución
 - 1.1 Nuevas ideas sociales
 - 1.2 Transformaciones sociales
 - 1.3 Crisis económica
- 2 Desarrollo de la Revolución
 - 2.1 Los Estados Generales
 - 2.2 La Asamblea Nacional
 - 2.3 Las Revueltas Populares en las ciudades
 - 2.4 Las Revueltas Populares en el campo
- 3 Consecuencias de la Revolución
 - 3.1 Nuevo sistema político
 - 3.2 Pérdida de poder de la Iglesia
 - 3.3 Pérdida de poder del Rey
- 4 Enlaces externos



Toma de la Bastilla, 14 de julio de 1789.

Causas de la Revolución

Nuevas ideas sociales

El siglo XVIII conoce grandes transformaciones intelectuales. Si en el siglo XVII se había iniciado la llamada *revolución científica*, durante el siglo XVIII surge en Francia un nutrido grupo de pensadores que creen que la razón humana resolverá muchos de los problemas que complican la vida de los seres humanos. Los líderes intelectuales de este movimiento, denominado *La Ilustración*, se consideraban a sí mismos como la élite de la sociedad, cuyo principal propósito era liderar a los países de la *Europa occidental* y del resto del mundo hacia el progreso, sacándolo de un largo periodo de tradiciones, superstición, irracionalidad y tiranía. Esta corriente abogaba por la razón como la forma de establecer un sistema autoritario ético, en oposición a la *sociedad pre-revolucionaria*.

Transformaciones sociales

El aumento de la población durante el siglo XVIII hizo que en ocasiones no hubiera alimentos para todos. A estas etapas se las llamaba crisis de subsistencia y eran cada vez más frecuentes. Como consecuencia de estas crisis el número de pobres, con frecuencia campesinos, crecía sin parar.

Además, la *nobleza pre-revolucionaria*, pese a su *poder político* y sus grandes posesiones, se encontraba en crisis ante el empuje de la burguesía, grupo social cuyo papel económico no paraba de crecer y que, sin embargo, veía cómo su fuerza política era muy reducida ante los deseos de la nobleza de ejercer un verdadero monopolio de los cargos del estado.

Por tanto, el *descontento social* de la burguesía y del campesinado (por diferentes razones) se convertirá en fuerza impulsora de la revolución.

Crisis económica

En vísperas del estallido de la Revolución, la economía francesa no pasaba por buenos momentos. La circunstancia de fondo era la mala situación de la *economía rural*, sector responsable entonces de dos tercios de la riqueza de Francia. Los años 1787 y 1788 se caracterizaron por desfavorables condiciones meteorológicas que pusieron en marcha una durísima crisis de alimentos, algo común en el antiguo sistema feudal. La escasez de alimentos condujo a una escalada de precios y con esta llegó una gran agitación política y el descontento social de los grupos más desfavorecidos.

Con todo, muchos historiadores han dado más peso como causa económica de la revolución no a la crisis agraria, sino a la crisis fiscal por la que pasaba el estado francés. Así es sabido que desde hacía décadas la *hacienda francesa* era incapaz de recaudar con sus impuestos las cantidades que se gastaban.



La libertad guiando al pueblo. Cuadro de Delacroix.

Desarrollo de la Revolución

Los Estados Generales

Los Estados Generales eran organismos políticos que representaban por separado a los tres estamentos del Reino: la nobleza, la iglesia y el pueblo o Tercer Estado. El rey no convocaba a los Estados Generales desde hacía más de un siglo. Así que ante las presiones de los estamentos del país, Luis XVI, Rey de Francia, se vio obligado a convocarlos. La reunión de los Estados Generales se realizaría a comienzos de mayo de 1789.

Francia se vio inundada por numerosos panfletos de carácter político. En ellos se concluía que el Tercer Estado, la inmensa mayoría de la nación francesa, formado por quienes trabajan y sostienen con sus impuestos el estado, debía tener en sus manos el control político de Francia.

En junio de 1789 reunidos ya los Estados Generales por separado, desde el Tercer Estado se cursó una invitación a los miembros de la Iglesia y la Nobleza para unirse bajo un solo techo. A mediados de ese mes los representantes del Tercer Estado decidieron constituirse en Asamblea Nacional.



Insignia revolucionaria de la Revolución Francesa.

La Asamblea Nacional

La Asamblea Nacional fue una institución enmarcada en la Revolución francesa que existió desde el 17 de junio de 1789 y sirvió de transición entre las instituciones de antes de la revolución y las nuevas *instituciones revolucionarias*.

En junio de 1789, reunidos ya los Estados Generales por separado, desde el Tercer Estado se cursa una invitación a los miembros de la Iglesia y la Nobleza para unirse bajo un solo techo. A principios de julio de 1789 la mayor parte de los nobles y los eclesiásticos ya se han unido a la llamada Asamblea Nacional y trabajan en la redacción de una Constitución y de una declaración de derechos. Tras la famosa toma de la Bastilla la revolución está plenamente en marcha, pues parece quedar claro que la autoridad del Rey quedaría limitada por las nuevas *leyes revolucionarias*. La monarquía absoluta se tambalea.

Las Revueltas Populares en la ciudad

Pero desde esos primeros días de julio se advierte que parte de la Corte no está dispuesta a aceptar la nueva situación, habiendo un despliegue de tropas en Versalles y París. Esta acción será respondida con movimientos populares de miles de personas buscando armas y construyendo barricadas ante el rumor de que la monarquía piensa terminar con las acciones revolucionarias de los Estados Generales.

En este contexto se producirá un acontecimiento que simbolizará para siempre el comienzo de las acciones revolucionarias: el 14 de julio de 1789 una enorme masa de parisinos se lanza a la toma de La Bastilla, fortaleza que servía de depósito de pólvora y prisión estatal. Los revolucionarios tomaron la fortaleza sin demasiada resistencia y, lo que es más importante, sin que intervinieran las tropas reales, que permanecieron impasibles ante el temor de que los soldados se unieran a los revolucionarios.

Las Revueltas Populares en el campo

En los días siguientes el estallido se extenderá por toda Francia numerosas revueltas campesinas que protestarán por sus *duras condiciones de vida*. En el campo se producirá una enorme agitación, con masas de campesinos tomando castillos de la nobleza al asalto con la intención de quemar los archivos donde se guardaban los documentos de propiedad señorial. Como respuesta a estas revueltas campesinas, los diputados de la Asamblea Nacional redactarán un decreto de abolición de los derechos feudales que, sin responder a todas las demandas, sirvió para pacificar a los *campesinos*.

Consecuencias de la Revolución

Nuevo sistema político

En el mes de agosto de 1789 la Asamblea Nacional avanza la redacción de la Constitución que contiene la reforma del Estado francés, quedando Francia configurada como una monarquía constitucional. La Constitución de 1789 organizaba el Estado francés mediante la separación de los poderes legislativo, ejecutivo y judicial. Esta distinción se basa en las teorías de los *grandes intelectuales franceses*.

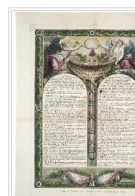
Dos años después del inicio de la revolución se celebraron elecciones en Francia y de ellas salió una nueva Asamblea Constituyente que recibiría el nombre de *Convención Nacional*, cuya primera tarea sería la redacción de una nueva Constitución. Los responsables políticos de la Convención decidieron suprimir la Monarquía y establecer un *sistema Republicano*.

Pérdida de poder de la Iglesia

El enfrentamiento entre los revolucionarios y la Iglesia es progresivo y se debe fundamentalmente a medidas tomadas por los poderes revolucionarios, como la nacionalización y venta de sus bienes, la supresión de órdenes religiosos y la obligación de todos los miembros de la Iglesia de realizar un juramento de fidelidad al nuevo estado. Casi la mitad de la Iglesia se resistirá a este juramento y parte de estas *revueltas contrarrevolucionarias* serán lideradas por los eclesiásticos en algunas *regiones francesas*, auspiciados por el Papa de Roma.

Pérdida de poder del Rey

Con la monarquía los problemas surgirán por la resistencia del Rey Luis XVI a aceptar la legislación revolucionaria de la Asamblea. Pronto las revueltas patrocinadas por la Iglesia y los *nobles emigrados* recibirán la ayuda de los *monarcas absolutos* de otros países europeos, y es en ese contexto cuando se produce la fuga de la familia real francesa. El rey huye en secreto, dispuesto a unirse a los rebeldes contrarrevolucionarios y a luchar contra las instituciones revolucionarias de su propio país. Después de varios años de gobiernos revolucionarios las instituciones francesas decidirán suprimir la Monarquía e instaurar la República francesa. Cuando finalmente se descubran pruebas de que Luis XVI está colaborando con los *ejércitos extranjeros* que luchan contra la Francia revolucionaria, los revolucionarios pedirán para él la pena de muerte. La ejecución pública del Rey se producirá en enero de 1793.



La Declaración de los Derechos del Hombre.

Enlaces externos

- [Análisis histórico sobre la Revolución Francesa](#)
- [Vídeos sobre la Revolución Francesa \(en inglés\)](#)
- [La Revolución Francesa \(historico digital\)](#)

