# Readers and writers: how the analysis of reading behaviour provides a sound basis for designing L2-hypertexts

**Isabelle De Ridder – Gert Rijlaarsdam – Luuk Van Waes**

**University of Antwerp – University of Amsterdam**

## Abstract

This paper reports on experimental research in which 56 students were asked to read two different texts on a computer screen in an educational setting. In the experiment we try to gain an insight into the characteristics of this kind of reading in order to provide a sound basis for designing suitable hypertexts for L2-learning. More specifically, we investigated whether electronic glosses are indispensable to the optimisation of foreign language reading and should thus be provided by the hypertext. If so, choosing a suitable (i.e., non disruptive but engaging) signalling-device becomes essential. The results indicate that electronic glosses are a surplus value to language learning. As for signalling, no significant results could be established.

## Introduction

With the advent of new technologies in language learning, every possible language-learning tool has become available at the click of a mouse. All kinds of help functions, exercises and examples are ready-made and waiting to be activated in order to improve acquisition. As for the surplus value of these additional features for L2-language learning, results are sometimes inconclusive (e.g., Brett, 1997; Knight, 1994). When it comes to deciding which tools to include, writers-designers are mostly left to their own devices and often act on their intuition. Reading research might provide writers-designers of this specific kind of hypertext with essential information concerning the use readers make of the software and the effectiveness or redundancy of some particular textual features.

In this paper, we present empirical research conducted in an attempt to evaluate the effect of textual glosses on the readers' L2-text comprehension and incidental vocabulary learning in order to formulate some guidelines for the writers-designers involved.[1] In the literature, attention is often drawn towards the impact of full electronic dictionaries or multimediatic glosses (Chun & Plass, 1996, 1997; Knight, 1994). In general, both are seen as an asset to language learning software (e.g., Knight, 1994; Chun & Plass, 1996, 1997; Hulstijn, Hollander, & Greidanus, 1996; Lomicka, 1998). The type of glosses that were used in this experiment differ from the ones mentioned above in a sense that they were presented in pop-up windows containing a dictionary definition and a translation. Both contextual and general meaning was presented, so that text comprehension problems could be solved *ad hoc*.

[1] Incidental vocabulary learning refers to the kind of vocabulary that is "picked up" during normal L1 or L2 reading or listening activities (e.g., Hulstijn et al., 1996; Nagy, Herman, & Anderson, 1985; Nation, 1990; Sternberg, 1987; Ellis, 1994). It is called "incidental" because it is considered to be a "by-product" of the main cognitive activity (reading).

If it is the case that this kind of textual glosses improves incidental vocabulary learning and text comprehension drastically, then a suitable signalling-device (i.e., the way in which the software indicates that glosses are provided) should be available to the writer-designer. Indeed, since studies have indicated that readers often seem reluctant to interrupt their reading and therefore tend to ignore unfamiliar words (Hulstijn, 1993), this device should encourage readers to consult the glosses provided when reading becomes problematic. From this point of view, highlighting a difficult word with a conspicuous colour, thus indicating a link towards the gloss, might be a plausible solution.

However, this conclusion might be premature. Indeed, an L1- study by Black, Wright, Black, & Norman (1992) suggested that readers use electronic definitions only for short-term purposes. A possible explanation for this could be that the easy accessibility of electronic glosses conflicts with "deep processing" (De Ridder, 2002). When relying on the theory of "cognitive depth" and high mental effort, one assumes that the more actively we work out a solution to the problem, the more likely we are to store this information permanently (Craik & Lockhart, 1972). Hulstijn (1992), for instance, indicates that inferred meanings are still remembered better than given meanings. Keeping this in mind, the software should prevent the reader from superficial, excessive clicking (called "click happy behaviour" by Roby, 1999, p. 98). It is in trying to establish this balance between "well-determined"[2] and "excessive" clicking, that the signalling-device of electronic glosses plays an important part. Indeed, a screen design with highly visible links could incite users to click excessively, whereas a condition with for instance invisible links could encourage students to make a more careful use of the context to derive the meaning of unknown words.

In this study, we therefore evaluated two kinds of signalling-devices: visible versus invisible links. In the condition with visible links, the glossed words were in blue and were underlined, which at the moment is the most common way of marking hyperlinks in on-line environments. In the condition with invisible links, the glossed words were not underlined and remained in the same colour as the rest of the text. In both conditions, the links led to small pop-up windows with dictionary definitions and translations. The pop-ups appeared in such a way that they never hided the sentence with the activated hyperlink, providing a suitable context for interpreting the gloss. For an overview of both conditions, see Appendix.

Two research questions are thus the focus of our investigation:

1. What is the effect of textual electronic glosses on the readers' text comprehension and incidental vocabulary learning?
2. If it is the case that electronic glosses are indispensable to high quality language learning software, what kind of signalling-device should writers-designers opt for?

An experiment was set up in which one group of students (group I) read two glossed texts: one with highlighted hyperlinks (marked) and one with invisible links (unmarked).[3] The links led to dictionary definitions and translations. Another group of similar students (group II) read the same texts in the same conditions (marked and unmarked) but with blind glosses (i.e., the highlights of the hyperlinks remained but the glosses were not accessible). In this way, a possible effect of marking stood apart from a possible effect of glossing. It became feasible to examine an effect of glossing, marking and a combination of both. After reading the texts, the students took an expected text comprehension test and an unexpected vocabulary test, which was repeated one week after reading in order to measure any long-term word retention (effect of time and possible interaction effects with marking and glossing).

[2] Hulstijn et al. (1996) suggest that dictionary use should remain "well-determined".
[3] This experiment is the third in a series of three. For an overview of the first two, see (De Ridder, 1999; 2000; De Ridder & Van Waes, 2000; De Ridder, 2002).

## Hypotheses

***Expected effect on clicking behaviour***

*Marking increases clicking* – This hypothesis only applies to group I, since the other group did not have access to the dictionary definitions. Black et al. (1992) investigated glosses for L1 in the context of languages for specific purposes. In their study, students consulted significantly more definitions when these were indicated by a black spot behind the word. Our own previous research confirmed these findings, since in the condition with highly visible links, students clicked significantly more than in the condition with invisible links (De Ridder & Van Waes, 2000; De Ridder, 2000; 2001).

***Expected effects on reading time***

*Marking increases reading time* – Since students were expected to spend significantly more time on consulting glosses in the marked than in the unmarked condition, we expected that reading time would increase as well.

*Glossing increases reading time* – It is difficult to predict whether it takes more time to consult glosses in pop-up windows than to infer the meaning of unknown words from the context. Nevertheless, we expected that the text version without glosses would shorten the reading time, because in this condition there is no explicit interactive focus on vocabulary. In the text version without glosses we expected that students would only try to infer the meaning of words that are crucial to a global comprehension of the text and ignore others.

***Expected effects on vocabulary learning***

*Electronic glosses increase incidental vocabulary learning*
The electronic glosses that were evaluated in previous studies are said to improve the pick-up rates of incidental vocabulary learning (e.g., Ellis, Tanaka, & Yamazaki, 1994; Hulstijn, 1992; Knight, 1994; Hulstijn, Hollander & Greidanus, 1996). We expected to confirm these findings with the type of glosses that we chose to include.

*Marking increases incidental vocabulary learning*
*Group I* – Since marking is expected to increase the students' willingness to consult the provided glosses and since glosses are said to have a positive effect on incidental vocabulary learning (e.g., Ellis, Tanaka, & Yamazaki, 1994; Hulstijn, 1992; Knight, 1994; Hulstijn, Hollander & Greidanus, 1996), we expected students to score significantly better on the vocabulary test in the condition using highlighted hyperlinks.
*Group II* – In reading research, some studies also indicate that when words are made salient, they are remembered better (Chun & Plass, 1996, 1997; Liu & Reed, 1995). In the literature on marking in texts on paper, it is said that texts passages that are made graphically salient are remembered better (e.g., van Hout Wolters, 1986). Because of all these indications, we expected marking (graphic salience) to have a positive effect on vocabulary learning.

*Vocabulary knowledge decreases in time*
Since vocabulary knowledge is known to decrease in time, we did not expect this to be any different in this experiment.

***Expected effects on text comprehension***

*Glossing leads to a better text comprehension* – Previous studies have indicated that text comprehension improves with electronic (multimediatic) glosses (e.g., Chun & Plass, 1996, 1997; Lomicka, 1998). We expected to confirm these findings.

*Marking leads to poorer text comprehension* – Schriver (1996, p. 284) states that typography could give readers clues about two key features of documents: information about the text structure (the hierarchy, the organisation, and divisions among levels of the content) and information about the roles of text elements (the purposes and relations among the text elements, and their function to one another). However, in the marked condition in our experiment the words were only typographically highlighted to indicate that they probably did not belong to the readers' active vocabulary knowledge. The marking was not used to make some content items more prominent and did not provide information about the textual hierarchy. This way of presenting a text on a screen could lead to a disturbed Gestalt of the text, especially because the density of the hyperlinks was quite high. We expected that the abundant hyperlink markings would distract the students from the argument of the text and would demand more cognitive effort from them to 'reconstruct' the (hierarchical) content of the text.

***Expected interaction effects***

***Interaction effect between glossing and time***

*Glossing leads to a more superficial learning of vocabulary* – The results of the study by Black et al. (1992) suggest that readers tend to use electronic glosses for short-term word retention, as a kind of immediate first aid to clarifying the meaning of a sentence. We expected to confirm these findings in this experiment.

***Interaction effect between time and marking***

*Marking leads to a more superficial learning of vocabulary* – Vocabulary knowledge is said to decrease in time; marking is said to increase the willingness to consult glosses and glossing is expected to lead to a more superficial knowledge of vocabulary (restricted in time), therefore we expect marking to reinforce the effect of vocabulary loss in time.

***Interaction effect between marking and glossing***

*Marking reinforces the effect of glosses in vocabulary learning and text comprehension* – Because marking and glossing are both expected to positively influence vocabulary learning, we expected to find a strong interaction effect between marking and glossing (reinforcing each other's positive effect).

## Method

### Subjects

Fifty-six second-year students of economics (university level; 35 females, 21 males) voluntarily participated in this experiment. These students were between 19 and 21 years old and none of them had participated in any previous experiments. All of them were computer literate and were native speakers of Dutch. None of them were bilingual but all had a fairly advanced level of French (9 years of French as a foreign language). Moreover, they were just completing a course of business French incorporated in their curriculum (30 hours/academic year).

## Design

The subjects were randomly assigned to two groups of twenty-eight students each. One group had access to electronic glosses, the other did not. Both groups were subjected to a Latin-square design, as clarified in Table 1. This specific kind of design controls for text and condition order.

| | group | First text read | | Second text read | |
|---|---|---|---|---|---|
| Latin-square design 1: glossed condition *N* = 28 | 1 (*N*=7) | Text 1 | Marked | Text 2 | Unmarked |
| | 2 (*N*=7) | Text 2 | Marked | Text 1 | Unmarked |
| | 3 (*N*=7) | Text 1 | Unmarked | Text 2 | Marked |
| | 4 (*N*=7) | Text 2 | Unmarked | Text 1 | Marked |
| Latin-square design 2: not glossed condition *N* = 28 | 5 (*N*=7) | Text 1 | Marked | Text 2 | Unmarked |
| | 6 (*N*=7) | Text 2 | Marked | Text 1 | Unmarked |
| | 7 (*N*=7) | Text 1 | Unmarked | Text 2 | Marked |
| | 8 (*N*=7) | Text 2 | Unmarked | Text 1 | Marked |

Table 1 - Design of the Experiment

## Material

***Reading Materials*** – All students read two French economic texts, comparable in length (about 2,000 words each), grammatical difficulty, and vocabulary load. Both texts were also used in the first and the second experiment. The texts had been selected after a pilot study involving four texts, 28 students and two faculty members of French. Within this pilot study, the texts were evaluated for interest and level of difficulty: the two texts with the most similar score were included in the study. The first text dealt with the diamond industry in Antwerp; the second one was about human resources in business.

Twenty-eight students read the texts with the help of electronic glosses. A word was glossed whenever one student of the pilot study did not know the Dutch translation of this word in a post-reading vocabulary test. This procedure led to the creation of 109 different glosses in the first text and 116 in the second one. The students could easily access the glosses by clicking on the defined word. A pop-up window with a Dutch translation and a French definition (separated by a horizontal line) would then appear. We chose to offer the translation and the foreign language definition, since preferences for one or the other are said to be highly individual (Jacobs, Dufon, & Fong, 1994). The pop-up window did not cover up the portion of the text in which the glossed word was found (see Roby, 1999; Stark, 1990). Context-bound explanations were given first; more general information was given at the end between brackets. This difference was explained to the students before they started to read (see Widdowson, 1978). An impression of the chosen screen design can be found in the Appendix. Two versions of each text were created: a marked one and an unmarked one. In the marked condition the glossed words were in blue and underlined. In the unmarked condition the hyperlinks were invisible (i.e., the glossed words were typed in the same colour as the rest of the text and were not underlined). The glosses remained identical in both conditions.

The other half of the students (twenty-eight) read the texts with so-called blind glosses. In the marked condition, the difficult words remained highlighted but the students could not click on the marked words. In the unmarked condition, the students read a text without highlights and without glosses.

The texts appeared in a black Times New Roman font, 12 points, on a white screen. This screen design was based on the "Recommendations for Basic Typography and Spatial Factors" of Scott Grabinger & Osman-Jouchoux (1996, p. 194-196). The authors give recommendations for screen design, based upon both print and computer screen research. In

the study carried out, the following aspects were divergent from these guidelines: full justification (instead of left justification) and line length (83 characters). Scott Grabinger and Osman-Jouchoux propose a line length of 60 characters, while other authors (e.g., Dyson & Kipping, 1998) find that the reading rate increases with a greater number of characters per line.

***Instrumentation and Observation***

*Clicking Behaviour and Total Reading Time* – An *Internet Explorer* specific java-script made it possible to register how much time the students spent clicking. With these log files it could be determined which glosses the students consulted and how long they consulted them for. The students were shown how the glosses worked before starting to read and they were explicitly told to close each pop-up window after consultation.

*Vocabulary Tests* – One week before the experiment, the students took an unexpected vocabulary test of 100 items during their regular courses of French. We explicitly chose to detach the pre-test from the actual reading exercise by delaying the experiment and letting the students take the first test during regular courses. The experiment itself was on a voluntary basis, outside of the courses. We decided to separate the pre-test and the actual experiment in order not to compromise the incidental nature of the vocabulary learning (since the pre-test might be an indication to the students that the focus of the experiment might be on vocabulary). The 100 items of the test were the items of the two vocabulary tests used in the experiment in addition to other items of the same or +1 - difficulty level. A reliability test of the pre-test yielded a Cronbach's alpha value of .79 (N of items = 100; N of cases = 56).

After having read each text, the students took an unexpected vocabulary test of 38 items. All of these items were chosen on the basis of the original pilot study and were thus glossed words. The test was a multiple-choice test where the students were asked to find the correct translation of the French words given. For every item, three Dutch words were proposed, an "I don't know"-option and a "no answer is correct"-option.[4] The Cronbach's alpha values indicated that these tests were highly reliable (text 1, $\alpha = .76$; text 2, $\alpha = .75$; N of items = 38, N of cases = 56). The results on this vocabulary test were expressed as a "learning coefficient", i.e., a coefficient that indicates what students have learned by reading the texts. This coefficient (LC) was defined as follows:

$$LC = \frac{\Sigma \text{ words learned from reading the text}^{5}}{\Sigma \text{ words correctly answered in the test}}$$

This coefficient always has a value between 0 and 1. A coefficient that approximates to one indicates that the student in question has learned many words, whereas a coefficient close to zero suggests that the reader has not learned many new words. In this way, by determining how much vocabulary students have actually learned, their pre-knowledge is taken into consideration.

One week after having read the text, the students took a delayed vocabulary test (the same items but in a scrambled order) in order to measure any long-term retention of the vocabulary possibly acquired (Cronbach's alpha values: text 1, $\alpha = .73$; text 2, $\alpha = .77$; N of items = 38, N of cases = 56).

*Text comprehension* – After reading the texts, the students took a text comprehension test. This test contained 14 open questions per text and proved to be fairly reliable (Cronbach's

[4] In three out of 38 times, this last option was the correct answer. This option was included in order to reduce guessing.

[5] These are the words that are answered wrongly in the pre-test but correctly in the post-test.

alpha values: text 1, $\alpha$ = .82; N of items = 27; N of cases = 56; text 2, $\alpha$ = .77; N of items = 23, N of cases = 56)[6].

## Procedure

The students read both texts in one session of two hours. The students came to the computer-lab to read the two texts consecutively (with a short break in between, if they so chose). The reading task was a fairly general task in which the students were asked to read the text for meaning. After reading the texts, they first took the tests (vocabulary and text comprehension) that related to the text they had read first and then the tests referring to the second text. One week later, they took the delayed vocabulary test during a regular class period of business communication.

## Results

### Effect of marking on clicking behaviour

These results only apply to the group of students that read the texts with the help of electronic glosses (group I). Table 2 shows the percentage of words clicked on in the glossed condition, analysed with a paired samples t-test (two tailed). The results show that marking has a significant effect on students' clicking behaviour: t=2.65, df=27, $p < .05$.[7] Thus, students click significantly more in the marked than in the unmarked condition, as was hypothesised.

| Condition | Mean | *SD* | Min | Max | *t* | df | p (2-tailed) | *N* |
|---|---|---|---|---|---|---|---|---|
| Marked | 44.66 | 26.82 | 10.07 | 110.79 | | | | 28 |
| Unmarked | 32.24 | 18.47 | 5.63 | 70.42 | | | | 28 |
| t-test | | | | | 2.65 | 27 | .013 | 28 |

Table 2 - Percentage of Words Clicked in the Glossed Condition (Group I)

### Effect of marking and glossing on reading time

Table 3 presents the total time the students spent on reading. These results were analysed with a General Linear Model, Repeated Measures with marking as a within-subjects variable and glossing as a between-subjects variable.[8] These results show that whereas glossing has a positive effect on reading time ($F$ (1,54)= 15.37, $p$ <.05), marking has no effect. Thus in the glossed condition, students spent significantly more time reading. Whether students read a marked or an unmarked text, this did not increase their reading time (in the glossed nor in the unglossed condition), contrary to what was hypothesised. In addition, no interaction effects were found.

| Condition | Marked | | | | Unmarked | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | *SD* | Min | Max | Mean | *SD* | Min | Max | *N* |
| Group I (glossed) | 1425.14 | 311.35 | 926 | 2283 | 1406.21 | 318.48 | 618 | 2278 | 28 |
| Group II (not glossed) | 1195.39 | 276.37 | 831 | 1988 | 1205.75 | 232.27 | 860 | 1740 | 28 |
| Total | 1310.27 | 313.88 | 831 | 2283 | 1305.98 | 294.12 | 618 | 2278 | 56 |

[6] The model answers to the open questions were divided in different content units. The number of items reported here refers to the total sum of content units in the comprehension test and resembles the maximum score.
[7] All statistical tests were performed at the .05 level, unless otherwise indicated.
[8] A General Linear Model, Repeated Measures is an Analysis of Variance (an ANOVA for Repeated Measures). Using the GLM-procedure, one can test null hypotheses about the effects of both the within-subjects and the between-subjects factors. A within-subjects design means that each participant provides more than one response. With a between-subjects variable, every set of responses comes from a different group of subjects.

| GLM, Repeated Measures | *SS* | *MS* | *df* | *F* | *p* |
|---|---|---|---|---|---|
| Effect of marking (within-variable) | 514.29 | 514.29 | 1 | .006 | 936 |
| Effect of glossing (between-variable) | 1295590.321 | 1295590.321 | 1 | 15.37 | .000* |
| Interaction effect marking*glossing | 6003.57 | 6003.57 | 1 | .075 | .785 |

Table 3 – Total reading time in seconds

### Effect of marking, glossing and time on vocabulary learning

Table 4 presents the results of the vocabulary tests, taken immediately after reading and one week later (expressed as a learning coefficient, cf. *supra*). These results were analysed with a General Linear Model, Repeated Measures with marking and time (test immediately after reading and delayed test) as within-subjects variables and glossing as a between-subjects variable. The results show that on the one hand, marking has no significant effect on vocabulary learning. On the other hand, time and glossing do have a significant effect on vocabulary learning: time, $F$ (1,54)=254.315, $p$ <.05; glossing, $F$ (1,54)=27.841, $p$ <.05. There are no interaction effects. Both groups thus score significantly lower in the delayed vocabulary test. Group I, which read the text with the help of glosses scores significantly better on the vocabulary test than group II. Thus, glossing increases incidental vocabulary learning (as was hypothesised), whereas marking does not have a specific influence (contrary to what was hypothesised). Moreover, marking does not lead to more superficial, short-term retention of vocabulary learning, since there is no significant interaction effect between time and marking (contrary to what was hypothesised).

| Condition | Marked | | | | Unmarked | | | | |
|---|---|---|---|---|---|---|---|---|---|
| IMMEDIATELY AFTER READING | Mean | *SD* | Min | Max | Mean | *SD* | Min | Max | *N* |
| Group I (glossed) | .44 | .13 | .27 | .83 | .43 | .11 | .25 | .72 | 28 |
| Group II (not glossed) | .30 | .15 | .00 | .60 | .32 | .14 | .00 | .62 | 28 |
| Total | .37 | .16 | .00 | .83 | .38 | .14 | .00 | .72 | 56 |

| Condition | Marked | | | | Unmarked | | | | |
|---|---|---|---|---|---|---|---|---|---|
| ONE WEEK LATER | Mean | *SD* | Min | Max | Mean | *SD* | Min | Max | *N* |
| Group I (glossed) | .31 | .12 | .07 | .68 | .30 | .12 | .12 | .61 | 28 |
| Group II (not glossed) | .19 | .13 | .00 | .47 | .15 | .12 | .00 | .50 | 28 |
| Total | .25 | .14 | .00 | .68 | .23 | .14 | .00 | .61 | 56 |

| GLM, Repeated Measures | *SS* | *MS* | *df* | *F* | *p* |
|---|---|---|---|---|---|
| Effect of marking (within-variable) | 2.188 | 2.188 | 1 | .094 | .760 |
| Effect of time (within-variable) | 1.042 | 1.042 | 1 | .254.315 | .000* |
| Effect of glossing (between-variable) | .986 | .986 | 1 | 27.841 | .000* |
| Interaction effect marking*glossing | 1.143 | 1.143 | 1 | .005 | .944 |
| Interaction effect marking*time | 1.290 | 1.290 | 1 | 2.815 | .099 |
| Interaction effect time*glossing | 1.116 | 1.116 | 1 | .272 | .604 |
| Interaction effect marking*glossing*time | 9.257 | 9.257 | 1 | 2.020 | .161 |

Table 4 – Results on the Vocabulary Test (expressed as a "learning coefficient") in All Conditions (Groups I and II) and Analysed Effects

### Effect of marking and glossing on text comprehension

Table 5 presents the results of the text comprehension test in percentage. The results were also analysed with a General Linear Model, Repeated Measures with marking as a within

variable and glossing as a between variable. These results show that glossing has a positive effect on text comprehension ($F(1,54) = 13.06$, $p < .05$) which was in line with our hypothesis. Marking has no influence on text comprehension and no interaction effects could be established, contrary to what was hypothesised.

| Condition | Marked | | | | Unmarked | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Mean | *SD* | Min | Max | Mean | *SD* | Min | Max | *N* |
| Group I (glossed) | 43.88 | 17.55 | 7.41 | 77.78 | 43.79 | 19.10 | 13.04 | 78.26 | 28 |
| Group II (not glossed) | 31.73 | 16.18 | 3.70 | 66.87 | 33.13 | 17.20 | 3.70 | 78.26 | 28 |
| Total | 37.80 | 17.81 | 3.70 | 77.78 | 38.46 | 18.80 | 3.70 | 78.26 | 56 |

| GLM, Repeated Measures | *SS* | *MS* | *df* | *F* | *p* |
|---|---|---|---|---|---|
| Effect of marking (within-variable) | 12.22 | 12.23 | 1 | .036 | .850 |
| Effect of glossing (between-variable) | 3645.72 | 3645.72 | 1 | 13.06 | .001* |
| Interaction effect marking*glossing | 15.41 | 15.41 | 1 | .046 | .831 |

Table 5 – Results on the text comprehension test in percentage

## Discussion and conclusions

The results of the experiment indicate that students spent significantly more time reading in a glossed condition. Glossing also has a strong positive effect on incidental vocabulary learning and text comprehension, which confirms the results of research literature (cf. *supra*). The textual glosses that were used in this experiment (dictionary definitions in the target language and translations) clearly increased the vocabulary learned incidentally and the quality of text comprehension.

Highlighting hyperlinks that lead to dictionary definitions increases readers' willingness to consult glosses. However, this particular clicking behaviour did not increase the time students spent reading. Neither did this intensified clicking affect vocabulary learning or text comprehension, since no significant differences were established between the marked and the unmarked condition. Students seem to pick up the same amount of new words in the marked and unmarked conditions and seem to understand the text equally well, contrary to what we hypothesised. In the second group of students –the one that was confronted with the blind glosses– no effect of marking was established either. This means that making difficult words graphically salient does not increase the chance of those words being incidentally learned by the reader. Nor does marking have an influence on text comprehension. Moreover, graphical salience (marking) did not interact with glossing, which means that the effect of glossing was not reinforced by the highlights.

Vocabulary learning decreases in time, since the results on the delayed vocabulary test were significantly lower in all conditions. However, the amount of words learned in the marked condition did not decrease more strongly than the words learned in the unmarked condition. Similarly, words that were learned in the glossed condition were not forgotten more easily than words that were derived from the context (group II), since no interaction effects could be established. These results indicate that glossing or marking does not necessarily lead to a more superficial, time-restricted knowledge of vocabulary, as was suggested by Black et al. (1992).

On the one hand, our results indicate that glossing difficult words in language learning software (with textual glosses as described above) seems justified from a pedagogical point of view (cf. the positive results as far as incidental vocabulary learning and text comprehension are concerned). On the other hand, marking does not seem to have influenced learning.

The fact that quite a lot of the hypotheses that were put forward in this paper, could not be confirmed in the study, demonstrates that one should be careful to act on intuitions and that extrapolating findings from one reading condition to another is often inadvisable (e.g., reading from paper vs. reading from a screen).

We also think that focussing on very specific elements in document design in controlled situations is a worthwhile method of conducting research. It allows the researchers to focus on well-defined characteristics of a text and leads to more precise answers that could guide writers-designers in their production process. Of course, the narrow focus of a study like the one presented here has also its drawbacks. For instance, one could wonder whether highlighting hyperlinks that have other functions (e.g., hyperlinks that lead to –multimediatic-background information, additional examples or graphs) or direct the reader to different (non-linear) hierarchical levels in a text will have the same effect as hyperlinks that lead to pop-ups with vocabulary glosses. Also the specificity of the environment (i.e., foreign language learning) and the reading task should be taken into account.

Finally, it goes without saying that a clearer understanding of the reading process will lead to a better insight in certain aspects of the writing process. The central role Hayes (1996) attributes to reading in his revised model of the writing process confirms the importance of the complementarity of these activities. In the discussion of this model for instance he stresses the fact that some of the reading activities that take place when writing are focused on building representations of the text as a spatial display. Because at the moment most of the writing occurs in digital environments, it is of course important to distinguish between reading from paper and from a computer screen (cf. Haas and Hayes, 1986). In order to gain a better understanding of this specific reading activity, we think that it is of the utmost importance to consider the effects of those design features that are specific for on-line texts (e.g., highlighted hyperlinks). By analysing the effects of this particular feature, we believe to have delivered a small contribution to the analysis of this complex reading activity.

## Further research

While we were able to establish an effect of glossing on vocabulary learning and text comprehension, effects of marking failed to occur. Nevertheless, it is possible that marking is related to individual preferences and individual learner profiles (cf. the rather high standard deviations (SD) in the percentage of words clicked, see Table 2). The results of the current research indicate that the learners involved were rather flexible and adapted easily to the software material they were confronted with. However, the question can still be raised whether users should adapt to the design of the screen and not the other way around. Visible (highlighted) hyperlinks might be helpful for some learners while disturbing for others, as far as vocabulary learning, text comprehension and the reading process are concerned. Students whose learning style is one of external regulation and step-by-step processing might benefit more from a highlighted condition than students whose learning style is one of self-regulation and deep-processing (Vermunt & Van Rijswijk, 1987). Furthermore, glossing might also be related to intra-personal differences. Knight (1994), for instance indicates that students with a different verbal ability use electronic glosses differently. In this sense, marking and glossing could be helpful to students with a low verbal ability. Since Knight (1994) worked with much shorter texts, it will be interesting to see whether the results from her study and the one presented here can be compared in any way. The kind of research that involves cognitive flexibility and cognitive profiles would undoubtedly be useful, especially in the light of today's technological developments.

In another line of research, we would like to link reading and writing more explicitly. The aim of this project would be to identify the cognitive processes that characterise the transfer from receptive to productive vocabulary knowledge. Furthermore, it addresses the question whether these processes are similar when they involve words learned incidentally through reading or when they have been picked up during a normal listening activity. Again, learning profiles are included in order to explain possible differences (post-hoc).

## References


Black, A., Wright, P., Black, D., & Norman, K. (1992). Consulting on-line dictionary information while reading. *Hypermedia 4*(3), 145-169.

Brett, P. (1997). A comparative study of the effects of the use of multimedia on listening comprehension. *System 25*(1), 39-53.

Chun, D. M., & Plass, J .L. (1997). Research on text comprehension in multimedia environments. *Language Learning & Technology 1*(1), 60-81.

Chun, D. M., & Plass, J. L. (1996). Effects of multimedia annotations on vocabulary acquisition. *Modern Language Journal 80*(2), 183-198.

Craik, F., & Lockhart, R. S. (1972). Levels of processing: A framework for memory record. *Journal of verbal learning and verbal behaviour 11,* 67-84.

De Ridder, I. (1999). Are we still reading or just following links? In K. Cameron (Ed.), *CALL and the learning community* (pp. 195-116). Exeter, UK: ELM Bank Publications.

De Ridder, I. (2000). Are we conditioned to follow links? Highlights in CALL-materials and their impact on the reading process. *Computer Assisted Language Learning 13*(2), 183-195.

De Ridder, I. (2002). Visible or invisible links: does the highlighting of hyperlinks affect incidental vocabulary learning, text comprehension and the reading process? *Language Learning & Technology 6(1)*: 123-146.

De Ridder, I., & Van Waes, L. (2000). Gemarkeerde hyperlinks en hun invloed op het leesproces, de woordenschatverwerving en het tekstbegrip [Marked hyperlinks and their influence on reading process, vocabulary acquisition and text comprehension]. In N. Ummelen, R. Neutelings, & F. Maes (Eds.), *Over de grenzen van de taalbeheersing. Lezingen gehouden op het VIOT congres in Delft* [Over the borders of language competence. Proceedings of the VIOT conference in Delft, The Netherlands] (pp. 381-391). Den Haag, The Netherlands: SDU.

Dyson, M. C., & Kipping, G. J. (1998). The effects of line length and method of movement on patterns of reading from screen. *Visible Language 32*(2), 151-181.

Ellis, R. (1994). Factors in the incidental acquisition of second language vocabulary from oral input: a review essay. *Applied Language Learning* 5: 1-32.

Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of word meanings. *Language Learning 44*(3), 449-491.

Haas, C. & Hayes, J.R. (1986). What did I just say? Reading problems in writing with the machine. *Research in the Teaching of English*, 20(1), 22-35.

Hayes, J.R. (1996). A new framework for understanding cognition and affect in writing. In: C.M. Levy & S. Ransdell (eds.), *The Science of Writing: Theories, Methods, Individuals Difference, and Applications* (pp. 1-28). Mahwah, NJ: Lawrence Erlbaum Associates.

Hulstijn, J. (1993). When do foreign language readers look up the meaning of unfamiliar words? The influence of task and learner variables. *Modern Language Journal 77,* 139-147.

Hulstijn, J. H. (1992). Retention of inferred and given meanings: Experiments in incidental vocabulary learning. In P. J. L. Arnaud & H. Béjoint (Eds.), *Vocabulary and applied linguistics* (pp. 113-125). London: Macmillan.

Hulstijn, J. H., Hollander, M., & Greidanus, T. (1996). Incidental vocabulary learning by advanced foreign language students: The influence of marginal glosses, dictionary use and reoccurence of unknown words. *The Modern Language Journal 80*(3), 327-339.

Jacobs, G. M., Dufon, P., & Fong, C. H. (1994). L1 and L2 vocabulary glosses in L2 reading passages: Their effectiveness for increasing comprehension and vocabulary knowledge. *Journal of Research in Reading 17,* 19-28.

Knight, S. (1994). Dictionary: The tool of last resort in foreign language reading? A new perspective. *The Modern Language Journal 78,* 285-299.

Liu, M., & Reed, W. M. (1995). The effect of hypermedia-assisted instruction on second-language learning. *Journal of Educational Computing Research 12*(2), 159-175.

Roby, W. B. (1999). What's in a gloss? *Language Learning and Technology 2*(2), 94-101. Retrieved August 13, 2001, from http://llt.msu.edu/vol2num2/roby/

Schriver, K.A. (1996). *Dynamics in document design: Creating text for readers.* New York: John Wiley & Sons.

Scott Grabinger, R., & Osman-Jouchoux, R. (1996). Designing screens for learning. In H. van Oostendorp & Sj. de Mul (Eds.), *Cognitive aspects of electronic text processing* (pp. 181-212). Norwood, NJ: Ablex Publishing Corporation.

Stark, H. A. (1990). What do readers do to pop-ups, and pop-ups do to readers? In R. McAleese & C. Green (Eds.), *Hypertext: State of the art* (pp. 2-9). Oxford, UK: Intellect Limited.

Sternberg, R. J. (1987). Most vocabulary is learned from context. In M. G. McKeown & M. E. Curtis (Eds.), *The nature of vocabulary acquisition* (pp. 89-105). Hillsdale, NJ: Erlbaum.

van Hout Wolters, B. (1986). *Markeren van kerngedeelten in studie-teksten [Marking core elements in study texts]*. Lisse: Swets & Zeitlinger.

Vermunt, J. D., & Van Rijswijck, F. A. (1987). *Inventaris leerstijlen voor het Hoger Onderwijs* [Inventory of learning styles for higher education]. Tilburg, The Netherlands0: Katholieke Universiteit Brabant

Widdowson, H. G. (1978). *Teaching language as communication.* Oxford, UK: Oxford University Press.

## Appendix

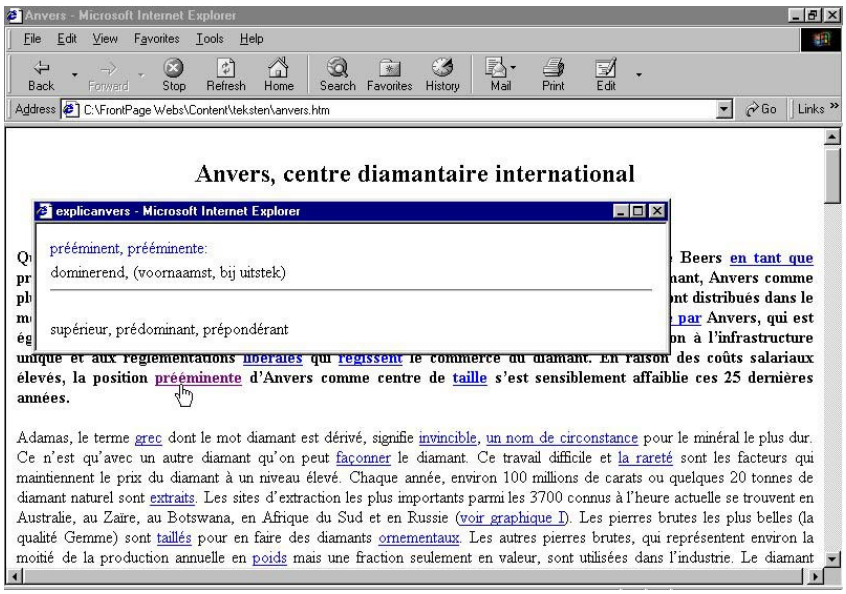


Figure 1 – Screen shot of the condition using visible links

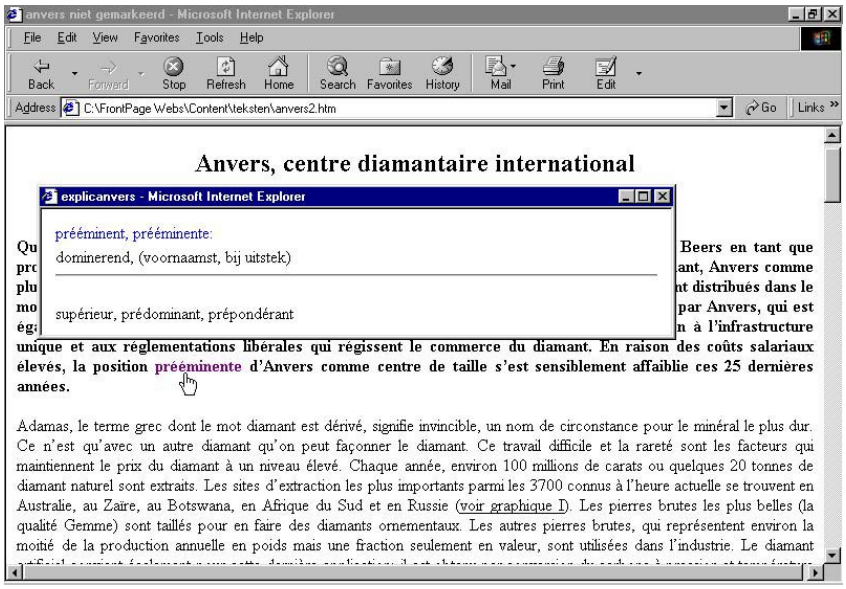


Figure 2 – Screen shot of the condition using invisible links