Does adding pictures to glosses enhance vocabulary uptake from reading?

CrossMark

Frank Boers [a], [*], Paul Warren [a], Lin He [b], Julie Deconinck [c]

[a] *Victoria University of Wellington, New Zealand*
[b] *Xi'An International Studies University, China*
[c] *Vrije Universiteit Brussel, Belgium*

ARTICLE INFO

ABSTRACT

This article reports three trials of a pen-and-paper experiment where adult L2 learners' recollection of glossed words was tested after they had read a text with or without pictures included in the glosses. Unlike previous studies in which a superiority of multimodal glosses over text-only glosses was claimed, the experiment furnished no evidence that the addition of pictures helped the learners to retain the glossed words any better than providing glosses containing only verbal explanations. When learners were prompted to recall of the written form of the words, the gloss condition with pictures in fact led to the poorest performance. The results suggest that the provision of pictures alongside textual information to elucidate the meaning of novel words may reduce the amount of attention that L2 readers give to the words proper.

© 2017 Elsevier Ltd. All rights reserved.

1. Introduction

Glosses or annotations that accompany text to clarify the meaning of unfamiliar words not only facilitate text comprehension but can also promote learners' acquisition of the glossed words (Hulstijn, Hollander, & Greidanus, 1996). Glosses draw attention to words that might otherwise be overlooked by the learner, they reiterate the word of interest within the gloss, and they ensure adequate interpretation — provided the gloss is sufficiently informative and the information is presented in a way that is accessible to the learner. Many studies in this area have compared the benefits of different kinds of glosses. For example, one line of research compares the benefits of annotations in L1 (translations) and in L2 (definitions) (Jacobs, Dufon, & Hong, 1994; Ko, 2012; Yoshii, 2006). Another line evaluates ways of stimulating cognitive engagement with the glosses, for example by incorporating an interpretation challenge in the gloss itself (Boers, 2000; Nagata, 1999; Watanabe, 1997). A third line of research concerns the potential benefits of multimodal glossing, defined here as the combination of textual clarification and pictorial elucidation of word meaning. That is the line of research to which the present article aims to contribute.

Several studies (see below) have reported findings that appear to support the thesis that multimodal information helps L2 readers retain the meaning of glossed words better than textual clarifications alone. The present article first evaluates those findings and the conclusions drawn from them, and points to a need for approximate replication studies to re-assess the benefits of multimodal glosses. This is followed by a report of such a study, the results of which cast doubt on the proclaimed superiority of multimodal glosses over text-only glosses for L2 vocabulary uptake.

* Corresponding author. School of Linguistics and Applied Language Studies, Victoria University of Wellington, PO Box 600, Wellington, New Zealand.
E-mail addresses: frank.boers@vuw.ac.nz (F. Boers), paul.warren@vuw.ac.nz (P. Warren), helin@xisu.edu.cn (L. He), julie.deconinck@vub.ac.be (J. Deconinck).

2. Literature review

2.1. Studies on the benefits of providing more than one gloss for a word

When pictures are used in printed materials for second language learning such as text books, they are presented to learners together with the textual input, typically on the same page, in a single gloss. It is that co-presentation of words and pictures in printed materials that is the object of the experiment we report further below. However, a fair amount of the evidence that has been interpreted in favour of adding pictures to verbal clarifications of word meaning comes from research conducted in the context of computer-assisted reading where participants consulted separate textual and pictorial annotations consecutively, by mouse-clicking on highlighted words. Chun and Plass (1996) is an early study of this kind. In their experiment, half of the targeted words only had a textual annotation while the other half had both a pictorial and a textual annotation. In post-tests where the students were asked to match the L2 words with their meaning, the words for which both pictorial and textual annotations were available generated the best scores. While this attests to the advantage of providing a textual annotation in addition to a textual one, the authors did not explore whether this might be due to multiple look-ups (in the case of two available annotations) rather than the multimodality *per se* of the information that was made available for half of the words. In a similar study, Plass, Chun, Mayer, and Leutner (1998) showed that participants were indeed more likely to retain the meaning of target words if they consulted two annotations about a word than if they consulted only one.

Jones and Plass (2002) investigated the benefits of word annotations in the context of computer-assisted L2 listening practice. Learners listened (in a self-paced manner) to a text with its transcript appearing on the computer screen. Again, some of the words in the transcript were highlighted as having annotations, accessible by mouse-clicking on them. For some of the learners only one annotation (either text or picture) was available for each of these words, while for other learners both a textual and a pictorial annotation was made available. The students who inspected two annotations outperformed those who were provided with only one in a post-test about the meaning of the L2 words.

Akbulut (2007) is another study where L2 readers could mouse-click on highlighted words to access annotations. Again, the participants' retention of the meaning of the annotated words was found to be better after reading a text in a condition where pictorial annotations were available in addition to textual ones than in a condition where only a textual annotation was available per target word. Although the author explains that the computer software recorded how often participants accessed given annotations, this data is unfortunately not included in the article, and so it is again impossible to tell whether the better post-test performance under the multimodal condition is to be attributed to the multimodality *per se* or rather to the greater number of look-ups prompted by the availability of more than one annotation.

In sum, what this handful of studies suggests is that making more than one annotation available stimulates look-ups. A word whose meaning is looked up twice also receives attention twice, and the amount of attention given to a word is known to be one of the predictors of word learning (e.g., Godfroid, Boers, & Housen, 2013; Godfroid & Schmidtke, 2013; see Schmitt, 2008, for a review that emphasizes the role of attention or "engagement" in vocabulary learning). What the above studies do *not* show, strictly speaking, is that it matters whether the available annotations which stimulate multiple look-ups and thus multiple episodes of engagement with the same word also include pictorial annotations.

2.2. Studies comparing the benefits of multimodal and text-only glosses for L2 vocabulary uptake

Let us now turn to studies where pictures and textual information were co-presented together in a single gloss, which is the reading condition that the present article aims to re-evaluate. In Kost, Fost, and Lenzini (1999), L2 learners read a short narrative text with marginal glosses for 14 unfamiliar words. Three groups of learners encountered a different version of the glosses: a version with L1 translations, a version with pictures, and a version with both an L1 translation and a picture in each gloss. In post-tests, the students were presented with the L2 words and asked (a) to provide the L1 translation, (b) to match the word with its corresponding picture, and (c) to match the word with its corresponding L1 translation. No significant differences between the three groups' scores were found in the first test (i.e., where the students were required to supply the meaning of the L2 words). On the picture recognition test the students who had received glosses containing a picture outperformed the students who had received only translations. This is not very surprising as the latter students had not seen those pictures before.[1] On the third test, where students were asked to match the L2 words with their L1 translation, it was the group that had received picture-only glosses that performed most poorly. Again, this is not so surprising, as these students had not seen the translations before. Interestingly, however, the combination of translation and picture in glosses yielded the best scores on this third test, which suggests that multimodal clarifications of word meaning were helpful for the learners in Koss et al.'s experiment.

A few other studies where visuals and verbal clarifications were combined in single glosses were conducted in computer-aided reading contexts again. In Al-Seghayer (2001) seven words in a text came with a definition only, for another seven words this definition was accompanied by a picture, and for yet another seven words the definition was accompanied by a video clip. In the post-test, which gauged students' recognition and recall of the meaning of the target words, the set of words that had been annotated only by means of a definition received the lowest mean score. It needs to be conceded, though, that it is hard to create sets of target words that are perfectly matched with regard to factors that influence their learnability (Ellis & Beaton, 1993,

---

[1] See Jones (2004) for evidence of input-test congruency effects in the context of word learning from single-mode and multimodal annotations.