Pictures in Online Dictionaries: Shall We See Them?

Anna Dziemianko
danna@amu.edu.pl
Adam Mickiewicz University in Poznań, Poland

ABSTRACT

The aim of the paper is to examine the influence of different access paths to pictures in online English learners' dictionaries on meaning reception and retention. The article also explores the influence of pictures on the time of meaning comprehension. The following aspects of access to pictures in online learners' dictionaries are investigated: making pictures instantly visible in entries, hyperlinking pictures, providing no pictures. In an online experiment, upper-intermediate foreign learners of English took part. They explained infrequent English concrete nouns based on reference to purpose-built monolingual dictionary entries. Three experimental conditions were created, depending on access to pictures in the entries: definitions with pictures visible by default (instantly visible pictures), definitions with pictures available upon clicking a hyperlink (hyperlinked pictures), definitions only (no pictures). Meaning comprehension and retention were evaluated based on L1 equivalents of the target nouns provided by the participants. The results reveal that instantly visible pictures and hyperlinked pictures improve meaning comprehension to a similar extent in comparison with the no-picture condition. However, meaning retention checked immediately after exposure is the most successful when pictures are visible by default in entries. Hyperlinked pictures prove to be no more useful for learning meaning than definitions without any pictorial support. They also extend comprehension time the most. Instantly visible pictures, in turn, neither speed up the comprehension of meaning nor slow it down compared to entries with no pictures. Considering their significant contribution to meaning comprehension and retention, instantly visible pictures seem to be the most recommendable in online dictionaries for learners of English.

Keywords: pictures; hyperlinks; online dictionaries; comprehension; retention

INTRODUCTION

Today, online dictionaries are multimodal. Multimodality refers to a combination of several sensory and communicative modes (e.g., printed words, static images, moving images or sounds) to produce meaning (Dressman, 2019). In other words, "multiple modes are orchestrated to express meaning" (Lee & Hazita, 2021: 104). The electronic medium has largely extended the repertoire of modes employed to represent meaning in online dictionaries. Apart from the verbal mode, the following ones can be involved (cf. Lew 2010):

eISSN: 2550-2131
ISSN: 1675-8021

- audio recordings of verbal elements (notably headwords, definitions and examples),
- audio recordings of non-linguistic sounds (e.g., onomatopoeias as well as recordings of sounds produced by musical instruments (e.g., the violin or the sitar), animals (e.g., roaring or hooting) or humans (e.g., clapping or hiccup),
- pictorial illustrations (e.g., (line) drawings, photographs or graphs),
- animations,
- video clips.

Traditional verbal explanations in dictionaries can thus be complemented in various ways. The multiplicity of channels on the screen may be assumed to help learners pick up meaning in a foreign language (Lew, 2010; Dressman, 2019). Multimodality in dictionaries can be expected to engage users, enhance language comprehension and acquisition. Yet, research practice does not always confirm such predictions. For example, Lew and Doroszewska (2009) found that animations in dictionaries have a significant *negative* effect on retention. Possibly, they distract dictionary users from form-meaning relations, which are fundamental to retention. Another reason might be insufficient multimodal literacy among dictionary users, defined as the ability to successfully engage with materials that involve different semiotic (or meaning-making) resources (e.g., written-linguistic, visual or audio; Camiciottoli & Campoy-Cubillo, 2018, p. 1). Lew (2010) rightly notes that much more research is needed to identify optimal combinations of meaning representation modes in dictionaries. This remark holds also for including static pictures in dictionary entries.

Pictures tend to be remembered better than words, a phenomenon known as the picture superiority effect (Ensor et al., 2019). The mnemonic advantage of pictures is most often attributed to their characteristics and the cognitive processing they springboard. According to the dual-coding theory (Paivio, 2013) there are two independent, though interacting systems in human memory: verbal and visual. Pictures are typically more likely than words to be encoded in both of them, which leads to better retention. Besides, pictures have more unique visual features than text, which makes them both physically and conceptually distinct. Physical distinctiveness means that there is more visual variety from picture to picture than from word to word. Conceptual distinctiveness implies that pictures require more semantic elaboration than words. Such characteristics stimulate elaborative processing and benefit retention (Ensor et al., 2019; Van der Broek et al., 2021).

However, the existing literature makes it by no means clear that pictures should be present in dictionaries, and, if so, how they should be incorporated. Below follows a literature review related to the most important aspects of picture inclusion in online dictionary entries and learning materials.

LITERATURE REVIEW

Pictures are generally considered a welcome feature of online dictionaries, where their basic function is to support definitions and assist comprehension (Klosa, 2015; Lew et al., 2018). Unfortunately, empirical research into the actual usefulness of graphic illustrations in online dictionaries is scarce. Dziemianko (2022) showed that illustrations in online learners' dictionaries are generally valuable for reception and retention (both immediate and delayed), though their actual usefulness depends on the specific format of graphic support (pictures in color, greyscale or line drawings). Apart from that, little is known about the effect of graphic illustrations on understanding verbal explanations in online monolingual entries and remembering meaning.

eISSN: 2550-2131
ISSN: 1675-8021

The beneficial role of adding pictures to text is more extensively evidenced in non-lexicographic studies into decoding and vocabulary learning (recently reviewed by Ramezanali et al., 2021). Their results support the multimedia principle that people learn better from words and pictures than words alone (Mayer, 2021, p. 57).[1] Such findings also accord with the dual-coding theory, which states that the presentation of both verbal and visual materials encourages learners to create referential connections between the two forms of mental representation, which enhances learning (e.g., Paivio, 2013).

Nonetheless, there are investigations which reveal no significant effect of pictures on learning. For example, Çakmak and Erçetin, (2018: 41) conclude that dual modes (verbal and visual ones combined) are no more effective than single modes (verbal or visual) in vocabulary learning, and pictures alone may not be as effective as text alone. Consequently, they advise caution when incorporating both textual-plus-visual and visual-only modes into multimedia materials. Similarly, Boers et al. (2017) established that text combined with pictures does not support learning and retaining nouns any better than verbal explanations alone. It has also been shown that the simultaneous presentation of pictures and words can result in downright *worse* comprehension and recall than either verbal or pictorial information presented separately. Plass et al. (2003, p. 240) found that multiple representations of information hinder vocabulary learning by low-ability students, who experience high cognitive load imposed by the need to process visual information. A negative effect of pictures (both alone and combined with text) on learning has been attested, too. Acha (2009) observed that pictures alone or paired with verbal annotations in a multimedia program lead to much worse immediate and delayed recall than words alone. Consequently, she suggested caution when designing vocabulary learning multimedia programs to avoid "unneeded duplication of essential information" (Acha, 2009, p. 24). Such findings cast doubt on the abundant and often uncritical use of images in textbooks and CALL packages. On top of that, students often overestimate the contribution of pictures to learning; they tend to think they learn more with their help than they actually do (Carpenter & Geller, 2020; Van den Broek et al., 2021). In practice, pictures can generate considerable distraction and rivet attention when unfamiliar or difficult words are learnt, especially when learners tend to process vocabulary through imagery (Boers et al., 2017).

Overall, the role of pictures in learning emerges as anything but incontestable. The existing research provides divergent results and gives only partial justification for adding pictures to text in learning materials. Different research aims and methodologies employed in the studies prevent any broad and valid generalizations. The question as to whether to include pictures in online dictionaries for foreign learners is then still a valid one. It also needs to be remembered that The Internet abounds in readily available images, and learners feel inclined to resort to various picture-based, non-lexicographic online resources, which may further call into question the rationale behind adding pictures to dictionary entries.

No less relevant is the question of how to display pictures in online dictionaries (if at all). Optimizing access to pictures is important, since presentation space is severely constrained (if only by screen size) both in hand-held portables and regular computers, on which online dictionaries are often accessed. There appear to be two basic approaches to including graphic illustrations in online dictionaries: making pictures instantly visible in entries or hyperlinking them. On the one hand, a picture can be integrated with the entry and embedded inside dictionary text, and thus visible by default together with the definition. On the other hand, it can be separated from the entry

[1] The principle applies provided that graphics are relevant to words. Seductive illustrations (interesting but not relevant) or decorative ones (neutral but not relevant) do not assist the learning process (Sung & Mayer, 2012).

eISSN: 2550-2131
ISSN: 1675-8021

by means of a hyperlink to be clicked on demand. Then, only the anchor text of the hyperlink (e.g., picture of thimble) is visible on the screen together with the definition. Upon clicking, the link takes the user to the relevant image. Regrettably, little is known about the usefulness of these access paths to pictures in learners' dictionaries.

The latter entry design (whereby pictures are separated from dictionary text) no doubt saves valuable presentation space. It is also supported by the information-delivery theory of multimedia learning (Mayer, 2009, p. 141). The theory assumes that learning is a cumulative process and consists in adding information to memory. Learning occurs when information is received by the student. The more times information is delivered, the more successful learning becomes (Mayer, 2009, p. 103, 123). If the verbal and visual information is delivered separately (at different places and / or times), learners stand a better chance of remembering it, as they are exposed to the information twice. Two separate exposures to verbal and visual information are better than one, when learners have only one chance to store it in memory. The apparent weakness of this theory is that media channels are nothing but information delivery routes used to pass information to students, and learning boils down to the transmission of information to the learner.

The former approach to designing entries (whereby pictures are integrated with dictionary text), in turn, conforms with the cognitive theory of multimedia learning. The theory assumes learning to be much more than adding separate pieces of information to one's memory; it is an active and dynamic process of making sense of the presented pieces of information and building connections between them. In this process, mental representations are actively created within learners' information-processing systems. When words and pictures are shown together to learners, the chances are that they will be held in working memory at the same time, which encourages building mental connections between them (Mayer 2021). Put differently, "simultaneous presentations prime the learner to build connections between corresponding visual and verbal material, whereas successive presentations make this active cognitive processing much more difficult by creating extraneous processing" (Mayer (2009, p. 168). The extraneous effort in successive presentations results from searching the screen for visual illustrations or / and trying to access the separate graphics that correspond to the printed text. This effort appears to be needless and should rather be invested in active learning. In the integrated approach, no such effort is necessary and more cognitive resources can be devoted to active learning. When graphics and words are processed at the same time, connections between them can be built more successfully, which facilitates making sense of them and committing them to learners' working memory (Mayer, 2021). The obvious advantage of this perspective is that it fosters generative processing by making it easier for learners to build connections between words and pictures. This, in turn, may result in deeper understanding and learning.

It is important to point out that the integrated approach to designing instructional materials respects two fundamental principles for reducing extraneous processing in multimedia learning: the principles of temporal and spatial contiguity. They hold that learning is more effective when words and pictures are presented contiguously in time and space. Put differently, it is helpful to present words and pictures simultaneously (rather than successively) and near (rather than far from) each other (Fiorella & Mayer, 2021). It may be surmised that also in dictionaries, the proximity of illustrations and definitions should encourage interaction between the verbal and visual modes of defining. Therefore, it may be expected that pictures integrated in dictionary entries could promote learning.

There is a substantial body of research which provides strong and consistent support for the cognitive theory of multimedia learning as well as the temporal and spatial contiguity

principles (reviewed by Mayer, 2021). Yet, the separated interface is not always bound to impede retention. When the sequentially presented visual and verbal information segments are adequately short, learners may be able to hold both in their working memory without exceeding its capacity and be still involved in meaningful learning. In dictionaries, concise definitions and *single illustrations*, or one-off pictures showing single objects (Svensén, 2009, p. 303), could well be classified as short verbal and visual segments, respectively. It is not impossible, then, that the separated interface could be justifiable in dictionaries. Unfortunately, there are no empirical studies to date which would determine whether the integrated approach to presenting pictures in online dictionaries (whereby pictures are shown together with entry text) or the separated one (whereby pictures are isolated from dictionary text by hyperlinks) is more recommendable for meaning reception and learning.

Two important issues are related to separating pictures from definitions and making them available on demand. They concern consultation behavior and consultation time. So far it is known that pictures integrated with dictionary entries complement verbal explanations; they do not monopolize users' attention or marginalize definitions. Lew et al. (2018) found that in monolingual entries with single pictures, users normally consult both definitions and pictures, with the latter consuming about 40% of look-up time. It is a matter of personal preference whether pictures or definitions are examined first, and users switch between them a small number of times. The existing research does not reveal, however, whether pictures separated from entry text by hyperlinks would also attract users' attention. There is a risk that the hyperlinks might be ignored and make hardly any difference to users. Entries with hyperlinked pictures would then be looked up virtually like those without pictures. On the other hand, hyperlinks no doubt allow the learner to decide whether they want to see pictures or not. This seems to be an asset considering the fact that many online instructional materials make learners process *all* visual and verbal information available on the screen, and do not allow the learner to request only the *relevant* information. As for the influence of pictures on consultation time, it has been found so far that entries without pictures take significantly longer to consult than those with single pictures integrated in the microstructure (Dziemianko, 2022). Yet, it might be conjectured that opening hyperlinks to pictures, which takes some time, might extend look-up.

Not only does clicking hyperlinks take time; it is also more effortful than seeing pictures in the entry right away. The Involvement Load Hypothesis (Hulstijn & Laufer, 2001) assumes that greater effort put in accessing information results in better learning. It follows that accessing hyperlinked pictures in online dictionaries might improve the retention of meaning. Quite intriguingly, Lew et al. (2018) found that the effect of consultation time (taken to reflect cognitive effort) on meaning retention depends on the type of illustration. They noted that longer time of processing an entry, thus greater mental exertion, entails better retention for objects illustrated out-of-context (in isolation). Yet, no such effect was observed when pictures show objects in their typical context.[2] The study did not concern, however, illustrations separated from the entry, but only embedded ones.

Overall, while the existing research reveals a few important and interesting facts concerning users' behavior and consultation time when pictures and entry text are integrated, little is known about the possible effect of separating pictures from definitions.

[2] Svensén (2009, p. 305) calls such pictures *structural illustrations* (e.g., a picture of the banister in the context of a staircase).

eISSN: 2550-2131
ISSN: 1675-8021

AIM

The aim of the paper is to determine if access to pictures in online dictionaries (instant/default visibility, hyperlinking, no pictorial support) affects meaning comprehension and retention. An attempt is also made to establish whether the time of reception is influenced by pictures in online dictionary entries. Three research questions are posed:

1. Is meaning comprehension affected by access to pictures in entries (immediately visible pictures, hyperlinked pictures, no pictures)?
2. Is the retention of meaning conditioned by the presence of pictures in entries and their access path (immediately visible / hyperlinked pictures)?
3. Is the time of meaning comprehension dependent on pictures in dictionary entries?

METHODS

INSTRUMENTS

To achieve the aims of the paper, an online experiment was conducted. Three data collection instruments were employed: a pre-test, a main test and a post-test. Each of them involved 15 English concrete nouns (*awning, barrel, bib, copse, gavel, hubcap, mortar, nightstick, pickax, playpen, plunger, rolling pin, thimble, trough, turnstile*). The following criteria were used for selecting the target items: they had to be concrete, infrequent English nouns which would lend themselves to illustration, unlikely to be known to upper-intermediate EFL learners.

Participants' familiarity with the test items was verified with the help of the first data collection instrument, i.e., the pre-test. In the pre-test, the target nouns were simply listed, one per web page. Their meaning had to be explained in the answer window below a test word. The participants were requested to supply equivalents in their native language (*anonymized*) relying only on their knowledge of English. No access to any resources was allowed.

In the main test, which constituted the second data collection instrument, the same task had to be done following the consultation of purpose-built, monolingual online entries. Each noun was shown on a separate page together with its entry compiled on the basis of the most prestigious online English learners' dictionaries: the *Oxford Advanced Learner's Dictionary* and the *Longman Dictionary of Contemporary English*. Each entry included a definition of the headword noun, its phonetic transcription and part of speech. All the definitions were analytical and framed within the scope of the defining vocabularies of the abovementioned dictionaries. Great attention was paid to the wording of the definitions and its possible effect on the results. Any genus which could be part of the L1 equivalent was replaced. For example, the OALD definition of *gavel*: *a small hammer used by a person in charge of a meeting or an auction, or by a judge in court, in order to get people's attention* contains the genus *hammer*, whose (*L1 anonymized*) equivalent (*L1 anonymized*) is part of the equivalent of *gavel* (*L1 anonymized*). Thus, the genus *hammer* was replaced by *tool* to avoid any undue influence of the wording of the definition on participants' answers.[3] Three versions of the main test were created, which differed only in access to pictures in the supplied entries (see Appendices A-C):

[3] The definitions used in the study are collated in Appendix D.

eISSN: 2550-2131
ISSN: 1675-8021

A) entries with pictures visible by default (instantly visible pictures),
B) entries with pictures available by clicking hyperlinks (hyperlinked pictures),
C) entries with no pictures.

The three test versions made it possible to determine if meaning reception was affected by access to pictures in the supplied dictionary entries. In the two versions with pictorial support (points A and B above), pictures and hyperlinks were given below the definition. The pictures in both test versions were the same; they were in color and offered only the images of the headwords with no background, and thus represented single illustrations (see the Introduction). In the test version with hyperlinks to pictures, the links opened in a new browser. In this way pictures were separated from entry text. By contrast, in the test version with pictures visible by default, the pictures were integrated into the entry structure and visible together with verbal explanations.

The post-test, the third data collection instrument, was not announced in advance. Like in the pre-test, the participants saw only the test items without any dictionary support and were not allowed to consult any sources on their own. They were expected to provide L1 equivalents of the target nouns from memory. The answers obtained in the post-test made it possible to measure how much meaning was retained after dictionary consultation in the main test.

To reduce the learning effect, the sequence of the target nouns was randomized in each data collection instrument (the pre-test, the main test and the post-test). The time of performing the task in the main test was for each participant automatically logged by the system. Logs were also used to check whether hyperlinks to pictures were accessed.

SUBJECTS AND PROCEDURES

A total of 238 upper-intermediate learners of English as a foreign language (B2 in CEFR) took part in the experiment. All of them were students at (*anonymized*) university. Their proficiency level was established on the basis of the final exam in English taken at the end of the academic year and the level of teaching materials used in class.

The experiment took place during regular university classes. The students were seated in a computer lab at uniform desktop PCs. First, they were given the pre-test and supplied L1 equivalents of the target nouns relying only on their knowledge of English. Immediately afterwards, they took the main test, in which L1 equivalents of the same words had to be given after reference to purpose-built monolingual dictionary entries. At this stage, the participants were randomly assigned to the experimental conditions, which corresponded to the three versions of the main test (88 students took the test with instantly visible pictures, 74 did the one with hyperlinked pictures, and 76 were given the test with no pictures). To divide the subjects into dictionary groups, students in each class were requested to count off from the left by threes and remember their number. Then, each number was assigned one test version. The assignment of numbers (1, 2 and 3) to the test versions varied across student groups; in one group a number (e.g., 1) was assigned the version with instantly visible pictures, in another group – the one with hyperlinked pictures, and in still different – the one with no pictures. The main test was followed by the immediate retention test. Throughout the whole experimental session, the students were closely monitored and their Internet activity was controlled.

The experiment was learner-paced in which the participants themselves managed the time of task performance; they decided how much time to devote to a test item and, in the hyperlink condition, when (if at all) to open the link to the relevant picture. Yet, they had to move

sequentially through the questions and could not return to a previous one or modify any already submitted answer.

SCORING

A participant's answer was considered correct and scored one point if it corresponded to any equivalent of the target nouns in the leading bilingual English-(*target language anonymized*) dictionaries, available in paper form or online. Spelling mistakes were not taken into account as long as they did not affect the comprehension of the supplied equivalents. Rarely (in less than 1% of all cases) did the participants supply descriptive explanations of the target nouns in their native language. Such responses did not score any point because they were always translations of the English definitions provided in the entries. It was thus impossible to assume that they testified to the comprehension of target noun meanings.

Participants' answers were evaluated independently by two raters, both native speakers of (*L1 anonymized*) proficient in English, who taught English at the academic level. The raters graded the responses independently. Perfect inter-rater reliability was obtained (r=1.00), with no cases of divergent evaluation.

EXPERIMENTAL DESIGN AND DATA ANALYSIS

Access to pictures was a between-groups independent variable (categorical predictor), as each participant was assigned to only one test version. The factor had three levels (instantly visible pictures, hyperlinked pictures, no pictures).

First, the effect of access to pictures on two dependent variables (*meaning reception* and *retention*) was investigated. To analyze the data, a one-way MANOVA was conducted. To further explore differences between the three experimental conditions (instantly visible pictures, hyperlinked pictures, no pictures) for each dependent variable (*meaning reception* and *retention*), univariate ANOVAs were conducted. The Bonferroni adjustment was introduced to control for the family-wise error when running multiple comparisons; each ANOVA was tested at a .025 (.05/2) alpha level. Significant ANOVA results were then analyzed with the help of the Bonferroni post-hoc test.

Next, to analyze the effect of pictures on the time of meaning comprehension, a one-way ANOVA was run with *access to pictures* as the three-level categorical predictor and *time* as the dependent variable. Significant results were investigated by means of the Bonferroni post-hoc test.

The analyses were run with the help of the *Statistica* package. The MANOVA and all ANOVAs were computed using the default sigma-restricted parameterization as a method of coding the levels of the categorical predictor variable, and the default Effective hypothesis Type VI sum of squares decomposition. In all the analyses, exactly the same results were obtained for other types of sums of squares decomposition (Types I, II and III). In other words, there were no differences in the results between the types of sums of squares.

eISSN: 2550-2131
ISSN: 1675-8021

RESULTS

RECEPTION AND RETENTION

First of all, it is necessary to note that the careful choice of test items was very felicitous; only 0.42% of meaning explanations provided in the pre-test were appropriate. Also, in the main test, the vast majority (97%) of the participants assigned to the test version with hyperlinked pictures consistently opened the links. Only two participants neglected them in most entries.

Results of the MANOVA indicate that there was a statistically significant difference between the three groups (participants consulting dictionaries with immediately visible pictures, hyperlinked pictures and without any pictures) on the combined dependent variables (Wilk's lambda = 0.418, F = 11.205, $p < .001$, partial $\eta^2 = 0.582$).

The results of the follow-up univariate ANOVAs show that meaning reception was dependent on access to pictures (F = 21.228, $p < .001$, partial $\eta^2 = 0.503$). It was the most successful when entries offered pictures either visible by default (82.15%) or hyperlinked (81.98%), with no difference between these two conditions ($p = 1.000$). In each of them, meaning reception was about one third better than when no pictures were available (61.06%; 82.15*100/61.06 = 134.33; 81.98*100/61.06 = 134.26), and these differences were statistically significant ($p < .001$, see Figure 1).

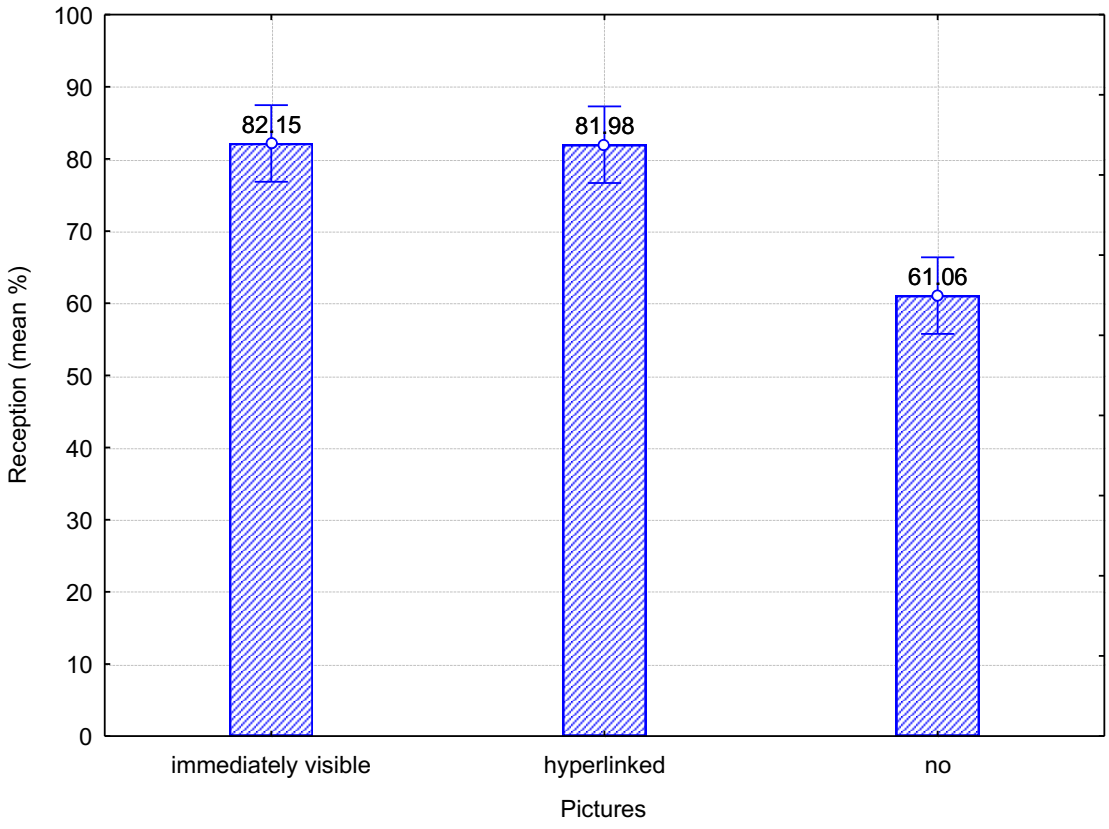


FIGURE 1. Reception by experimental group

Learning meaning was also affected by access to pictures (F = 7.991, p = .001, partial η^2 = 0.276). It was largely facilitated by pictures visible by default in the entry (62.07%). Entries with hyperlinked pictures (48.21%) were no more useful than those without any pictorial support (41.04%, p = .562). In these two conditions, meaning retention was, respectively, about one fourth (48.21*100/62.07 = 77.67, p = .039) and one third (41.04*100/62.07 = 66.13, p = .001) worse than when pictures were instantly visible, and these differences were statistically significant (see Figure 2).

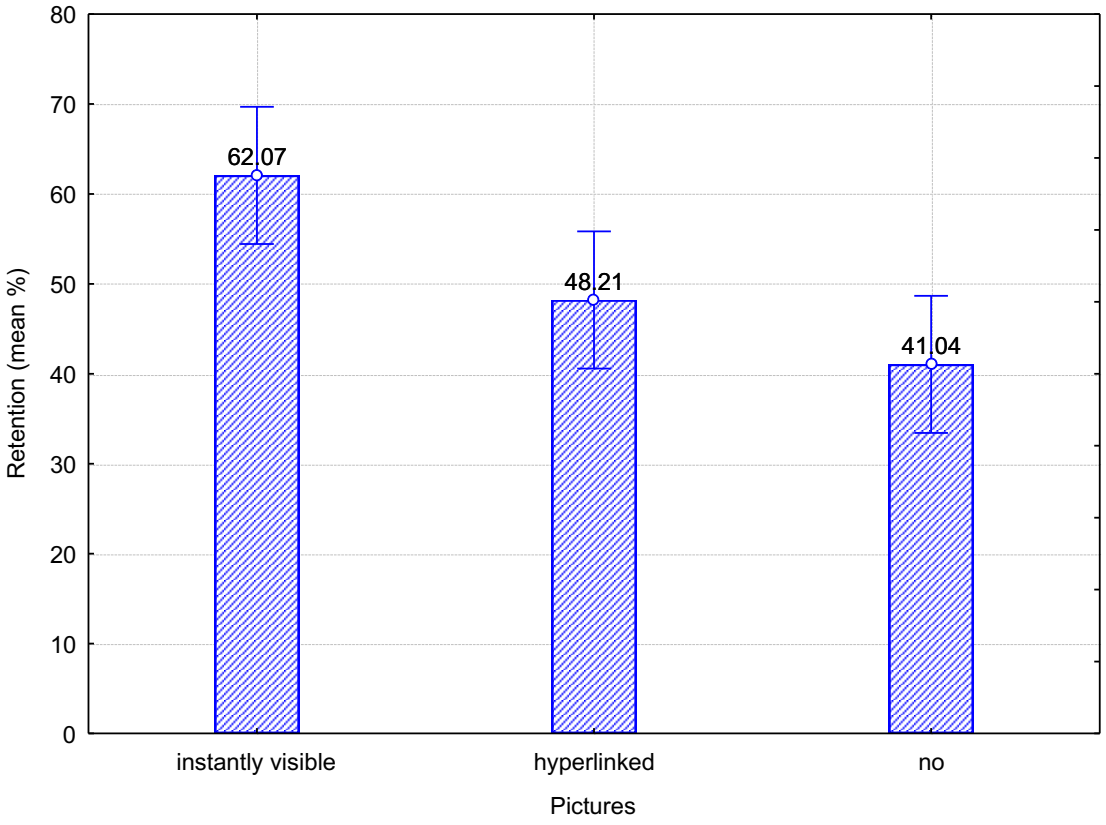


FIGURE 2. Retention by experimental group

TIME

The ANOVA results show that the time of task performance in the main test was dependent on access to pictures (F = 18.738, p < .001, partial η^2 = 0.472). The effect was quite large; over 47% of the variability in the time devoted to explaining meaning can be attributed to pictures. The activity proved to be the most time-consuming in the group dealing with hyperlinked pictures, who needed on average about 1673 seconds to supply L1 equivalents. That time was significantly longer than when pictures were displayed by default (1226 seconds, p <.001) or when there were no pictures in entries (1370 seconds, p = .001, see Figure 3). In these two conditions, task performance proved to be comparably time-consuming (p = .181).

eISSN: 2550-2131
ISSN: 1675-8021

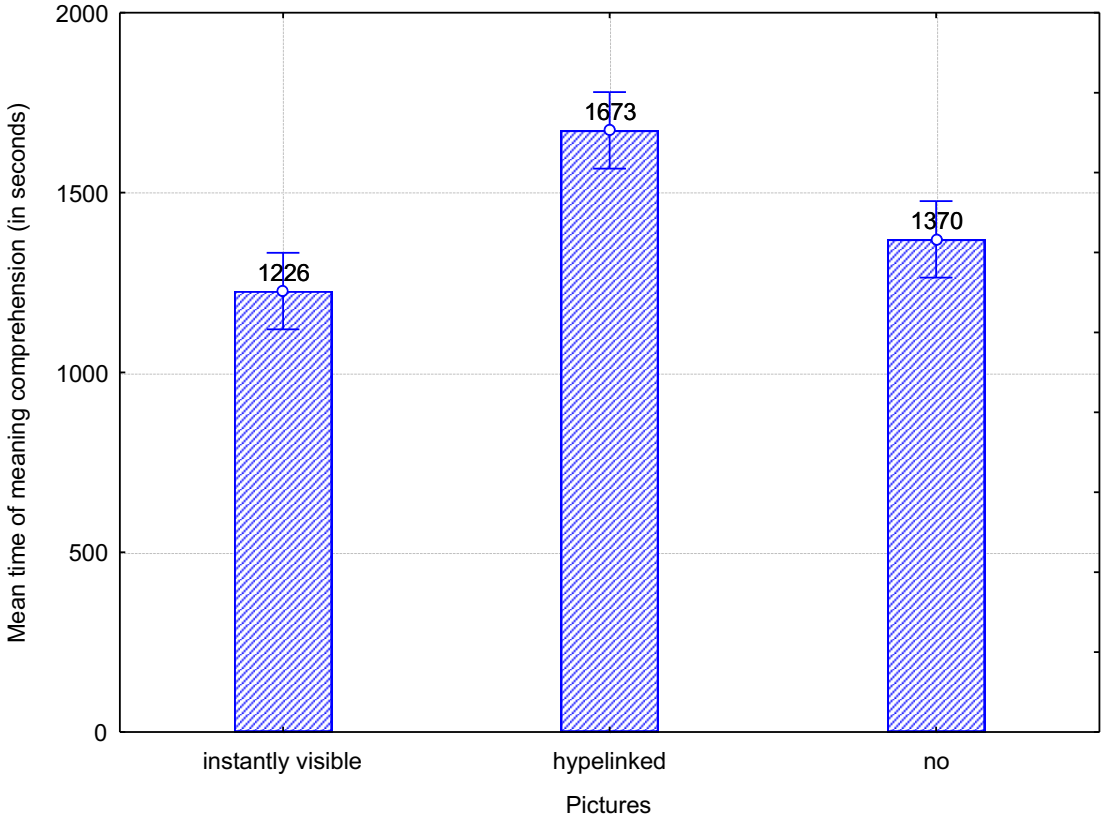


FIGURE 3. Mean time of meaning comprehension by experimental group

SUMMARY OF THE RESULTS

Table 1 symbolically summarizes the obtained results for reception, retention and decoding time in the three experimental conditions. Any results which are significantly different from each other are assigned different numbers (1 and 2). By the same token, those which share a number (either 1 or 2) are comparable. Significantly better results are marked with 1, worse ones with 2.

TABLE 1. A symbolic summary of the results

Dependent variable	Instantly visible pictures	Hyperlinked pictured	No pictures
Reception	1	1	2
Retention	1	2	2
Time	1	2	1

First, the obtained results show that understanding meaning is affected by access to pictures in entries. Comprehension is the least successful when no pictures are provided, and boosted to a comparable extent (by about one third) when pictures are available by default and hyperlinked (research question one). Second, remembering meaning is dependent on whether pictures are instantly visible or hyperlinked. The former significantly enhance retention, while the latter do not; they prove to be only as good as entries without pictures (research question two). Thus, pictures immediately visible in online entries emerge as the most recommendable for learning meaning.

Third, pictures affect the time of meaning comprehension. Hyperlinked ones are the least recommendable in this respect as they significantly extend the process of decoding meaning. Pictures visible by default neither speed it up nor slow it down in comparison with the no picture condition (research question three).

DISCUSSION

It seems surprising that virtually all participants consulting entries with hyperlinks accessed pictures. Possibly, they felt insecure trying to understand the supplied dictionary definitions and decided to see pictures to (better) understand meaning. Alternatively, they may have understood the definitions, but still had problems with visualizing the objects defined, and just wanted to verify their predictions. The participants may also have had problems with finding L1 equivalents on the basis of the definitions. Pictures might have helped them to name the objects in their native language. On the other hand, it is also possible that hyperlinks in dictionaries play an organizational role. They may impose a rigid hierarchy that provides a directional guide preventing dictionary users from disoriented performance. The resulting hierarchical entry structure may induce a specific mechanical response (or reference behavior) in dictionary users: reading definitions first and clicking hyperlinks next. If so, this non-linear entry structure may override the freedom or personal preference when deciding where to look first in entries (cf. Lew et al., 2018). Finally, it is conceivable that the participants simply felt they were expected to open the hyperlinks to see pictures, and they did so to perform as required and not to disappoint the researcher. The Hawthorne effect might thus have come into play; the awareness of being involved in a study with hyperlinks might have made the participants adapt their behavior accordingly (cf. Tight, 2022).

The research shows that meaning comprehension, operationalized as providing L1 equivalents, was the least successful in the no-picture condition. This result might be explained by the so called selective interference effect. The effect occurs when people are simultaneously involved in at least two tasks which require some manipulation of representations from the same code or system (e.g., either verbal or imagery/visuo-spatial, like doing two verbal tasks, or two imagery/visuo-spatial tasks). Then, experimental results tend to be poorer than when one task (e.g., verbal) is done together with another which calls upon simultaneous reference to the other code (e.g., imagery/visuo-spatial; cf. De Beni & Moè, 2003). The rationale for this effect can be found within the framework of the dual-coding theory; tasks which involve the same code create interference because the same representational and processing resources are called upon. Thus, when L1 equivalents had to be provided solely on the basis of verbal definitions, only one (verbal) code was involved, which led to the selective interference effect and, consequently – significantly worse performance than in either condition with pictures.

Better retention results obtained by the group who consulted entries with instantly visible pictures can be explained when the spatial and temporal proximity of the verbal and visual information is considered. Instantly visible pictures were integrated in the entries and visible together with definitions, which, in keeping with the cognitive theory of multimedia learning, supports learning. In the hyperlink condition, separating images from dictionary text through hyperlinks might have produced the split attention effect. As Ayres and Sweller (2021) explain, split attention occurs when learners need to divide their attention between a few sources of physically or temporally separate information, each of which is crucial to grasping the material, and then mentally integrate them. Such an attention split violates two principles of the cognitive theory of multimedia learning: the principles of spatial and temporal contiguity, described in the

Literature Review. When these principles are compromised and attention is split, extraneous cognitive load (associated with confusing presentation) is increased due to the need to mentally integrate the disparate sources of information, and the chances of learning are reduced. Possibly, the lack of both spatial and temporal contiguity of images and definitions in the hyperlink condition made the subjects split their attention, which negatively affected learning meaning. Overall, the retention results seem to support the cognitive theory of multimedia learning, which justifies integrating pictures with definitions in dictionary entries, rather than the information-delivery theory, according to which separated interfaces for text and images would be more recommendable (cf. Mayer, 2009, 2021).

The current study reveals that learners working in the no-picture condition needed virtually as much time to decode meaning as those who saw pictures together with definitions; the difference of 144 seconds in favor of the latter condition was not statistically significant. This observation diverges from previous findings that meaning reception takes considerably longer in the absence of pictures from entries than in their presence (Dziemianko, 2022). The current study also challenges the conclusion that the best meaning retention is related to the longest consultation time (provided that pictures show objects without typical context; cf. Lew et al., 2018).[4] The best retention was now observed for entries with instantly visible pictures (62.07%), which were actually consulted for the shortest time (1226 seconds). When consultation was the longest (1673 seconds in the hyperlink condition), the retention rate (48.21%) proved to be no significantly better than the worst result (41.04%), obtained in the no-picture condition. Possibly, clicking hyperlinks not only extended look-up, but also introduced some distraction from the task and extraneous load connected with the less straightforward presentation. Maybe it made dictionary users momentarily lose focus and negatively affected retention. Consequently, even though more effortful than integrated pictures, hyperlinks did not yield better retention, in contrast to what could be expected in the light of the Involvement Load Hypothesis (Hulstijn & Laufer, 2001). However, caution is necessary when trying to make any associations between consultation time and retention in the current study. The measurement of time in the present investigation is a less accurate approximation of actual consultation (let alone attentional focus) than in the case of Lew et al. (2018), where, thanks to eye tracking, gaze fixations more reliably reflected cognitive effort invested in processing lexicographic information. In the current research, logs only made it possible to measure the combined time of both reading an entry and submitting an answer. It cannot be guaranteed, either, that the whole of it was spent on the task.

LIMITATIONS

There is a chance that some subjects who grasped definitions (and even saw pictures) were still unsure what to call the defined objects in their mother tongue. Such problems might be put down to the selection of the target words. The nouns denote objects not often used by the participants and typically absent from their closest environment. This seems to be a natural, though unfortunate consequence of choosing infrequent concrete English nouns, likely to be unfamiliar to upper-intermediate learners.

As noted above, the measurement of time captured not only reading definitions and (where applicable) consulting pictures, but also coming up with equivalents. Regrettably, in view of software limitations, it was technically impossible to separate the time of processing dictionary

[4] Only such pictures were used in the present investigation.

eISSN: 2550-2131
ISSN: 1675-8021

information from the time of supplying equivalents. Without recourse to eye-tracking, it was unfeasible to identify (and exclude from time counts) momentary concentration lapses, either.

Unfortunately, no learner characteristics were taken into consideration apart from the proficiency level. In particular, it was not determined whether the participants were visualizers or verbalizers. Visualizers and verbalizers tend to perform differently depending on whether learning materials contain images or not. It has been shown that learning outcomes improve when learning materials match individual cognitive styles (e.g., Thomas & McKay, 2010). It is impossible to tell whether the assignment of the participants to the experimental conditions in the current study corresponded with their learning preferences or cognitive styles. Also, pictures are known to induce positive mood and feelings of satisfaction as well as reduce the perceived difficulty of the material to be covered (Sung & Mayer, 2012), a set of affective-motivational outcomes that the present investigation did not address, either.

FURTHER RESEARCH

In further research into illustrations in online dictionaries it might be recommendable to focus on more common words whose L1 equivalents would be easier to come up with. This might mean that subjects who are less proficient in English should be involved.

It might also be useful to see if access to pictures matters when dictionaries are consulted on mobile devices. Then, the size of the screen might make lexicographers particularly willing to hide pictures under a brief hyperlink anchor text, or provide miniature pictures expandable upon tapping (see Figure 4). The latter, also known as *scenic illustrations* (Svensén, 2009, p. 310; Klosa, 2015, p. 519), present a set of thematically related objects in a given context, and the assembled components are identified by labels. Obviously, users have to put in some effort to find a picture of the looked up word in a set of related pictures. Empirical research is needed to see whether hyperlinks to single pictures or expandable scenic pictures are more recommendable in online dictionaries consulted on small screens.

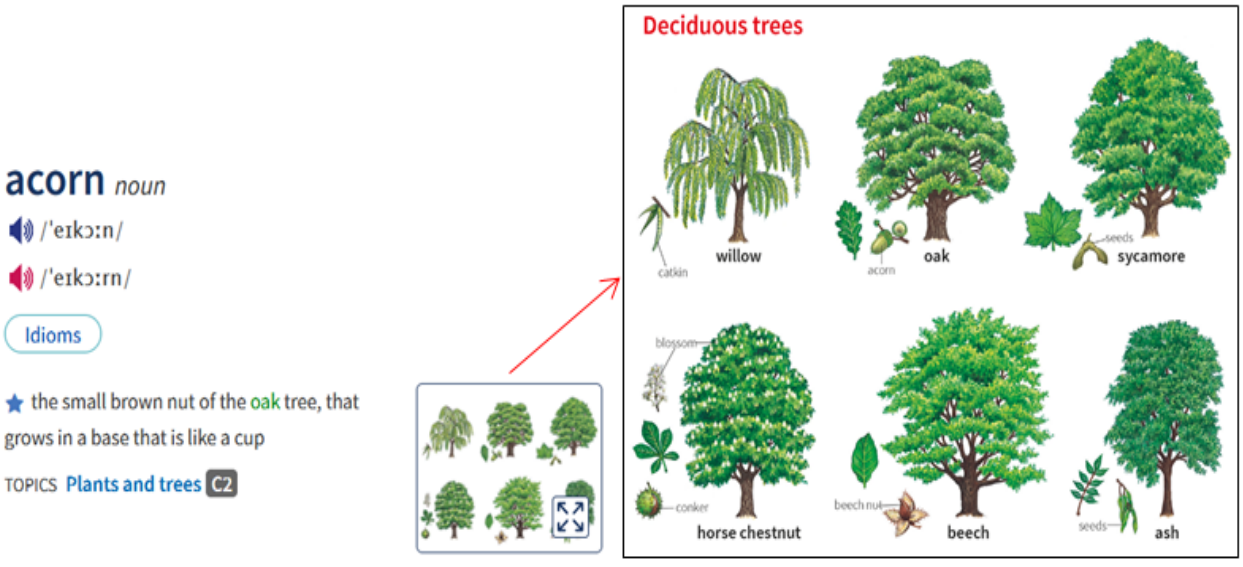


FIGURE 4. A miniature scenic picture in the entry for *acorn* in the *Oxford Advanced Learner's Dictionary* and its expanded view

eISSN: 2550-2131
ISSN: 1675-8021

It is not certain, either, that in a different task and / or for a different set of test items, hyperlinked pictures would be accessed by users. If the task did not require the provision of L1 equivalents, or if more commonly used nouns were selected for test items, users might not be so eager to open hyperlinks to see pictures of the words defined. Besides, a more naturalistic study is necessary to see if links are (willingly) accessed when dictionary users do not feel they are expected to do so.

It also seems worthwhile to investigate the usefulness of entries with pictures only. It may be argued that pictures, which define words by ostension, or showing the objects which the words designate, give basically the same information as verbal definitions. This might generate redundancy, which can occur when the same information is presented simultaneously in different forms. Redundancy imposes heavy demands on working memory; it makes learners unnecessarily coordinate and relate multiple forms of the same information, wastes limited cognitive resources and inhibits learning (Kalyuga & Sweller, 2021). Yet, what is redundant is relative and depends on who is taught, what information is presented and how. Definitions may prove too difficult (thus useless and redundant) for less advanced learners, but important for the proficient, who comprehend them easily. Besides, to make definitions potentially redundant, pictures in dictionaries need to be self-contained, unambiguous and fully intelligible without additional verbal support. Not all words lend themselves well to be illustrated in this way; abstract nouns, for example, are typically non-picturable. There are also different illustration formats available, like color and greyscale pictures or line drawings, which differ in usefulness (Dziemianko, 2022). Lexicographers also have at their disposal more illustration types than the aforementioned single, structural and scenic pictures (cf. Svensén, 2009). Consequently, depending on the illustrability of words and lexicographers' design decisions, the verbal and pictorial entry content may only partly overlap in lexicographic practice, which could generate *partial* rather than full redundancy (Kalyuga & Sweller, 2021).[5] Partial redundancy can be beneficial for learners who require additional support (e.g., less skilled readers). However, as expertise increases, material and design principles which are essential for less expert learners may become redundant and disadvantageous for more knowledgeable ones, and may even hinder learning, a phenomenon known as the expertise reversal principle (Kalyuga & Sweller, 2021). Thus, to investigate picture-definition redundancy in monolingual dictionary entries, different proficiency groups and degrees of partial redundancy between the visual and verbal content should be considered.

Mayer (2021, p. 156) has recently called for investigations into the kinds of pictures that should be added in instructional materials, how they should be added and for whom they should be added. It appears that all aspects of the call are immediately relevant also for further research into the usefulness of pictures in online English learners' dictionaries.

[5] Interestingly, Mayer (2021) holds that words and pictures are qualitatively different and, by their natures, cannot be informationally equivalent. It follows that they can only be at best complementary, but never substitutable.

eISSN: 2550-2131
ISSN: 1675-8021

CONCLUSION

In the contemporary visual culture, pictures are omnipresent. They have also found their way to online dictionaries, which have become truly multimodal. Images are conventionally considered to be superior to words inasmuch are they tend to better stick in memory. Yet, as the literature review shows, empirical research does not always confirm their beneficial role in language learning; adding pictures to text may, but does not have to, bring educational benefits. Supplementing dictionary definitions with pictures was thus taken up as a research topic to see whether the visual mode affects meaning reception and retention as well as the time of meaning construal. Considering the limitation of presentation space in the various electronic devices on which online dictionaries are accessed, an optimal way of incorporating pictures in online dictionary entries was sought. The effectiveness of two access paths to pictures was investigated: hyperlinking and default visibility in entries. The condition with no access to pictures in entries was also included in the design.

The results obtained from an online experiment justify giving an affirmative answer to the question posed in the title of the paper: yes, we shall see pictures in online dictionaries, and they should be instantly visible. Admittedly, providing pictures either as immediately visible images or hyperlinks improves meaning comprehension to a similar extent (by about one third) in comparison with the absence of any visual support. Yet, if meaning is to be remembered, pictures should be displayed by default in entries. Only then do they support learning. Hyperlinked pictures, by contrast, do not bring any substantial retention gains. They are also the least recommendable considering the time of decoding, which they significantly extend.

Broadly speaking, the findings confirm the cognitive theory of multimedia learning, but not the information-delivery theory (cf. Mayer, 2021). They justify integrating the visual (pictorial) and verbal (textual) modes in the online dictionary interface rather than separating them. Apart from the clear lexicographic implications, the findings may hopefully have broader pedagogical applications and may be relevant for instructional materials design.

Although rigorously designed and conducted, the study is not free of limitations. An important concern refers to the selection of test items and the operationalization of meaning comprehension, which may have made the participants lost for words in their native language (i.e., L1 equivalents). Another weakness relates to the measurement of meaning construal time, which could be precise enough only with resort to eye-tracking. Hopefully, the limitations of the study can soon morph into research inspirations and opportunities to further explore the role of pictures in online dictionaries.

REFERENCES

Acha, J. (2009). The effectiveness of multimedia programmes in children's vocabulary learning. *British Journal of Educational Technology*. *40*(1), 23-31. https://doi.org/10.1111/j.1467-8535.2007.00800.x

Ayres, P. & Sweller, J. (2021). The split-attention principle in multimedia learning. In R. Mayer & L. Fiorella (Eds.), *The Cambridge Handbook of Multimedia Learning* (pp. 199-211). Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139547369.011

Boers, F., Warren, P., He, L. & Deconinck, J. (2017). Does adding pictures to glosses enhance vocabulary uptake from reading? *System*. *66*, 113-129.

Çakmak, F. & Erçetin, G. (2018). Effects of gloss type on text recall and incidental vocabulary learning in mobile-assisted L2 listening. *ReCALL*. *30*(1), 24-47. https://doi.org/10.1017/S0958344017000155

Camiciottoli, B. C. & Campoy-Cubillo, M. C. (2018). Introduction: The nexus of multimodality, multimodal literacy, and English language teaching in research and practice in higher education settings. *System*. *77*, 1-9.

Carpenter, S. K. & Geller, J. (2020). Is a picture really worth a thousand words? Evaluating contributions of fluency and analytic processing in metacognitive judgements for pictures in foreign language vocabulary learning. *Quarterly Journal of Experimental Psychology*. *73*(2), 211-224. https://doi.org/10.1177/1747021819879416

De Beni, R. & Moè, A. (2003). Imagery and rehearsal as study strategies for written or orally presented passages. *Psychonomic Bulletin & Review*. *10*, 975-980. https://doi.org/10.3758/BF03196561

Dressman, M. (2019). Multimodality and language learning. In M. Dressman & R. W. Sadler (Eds.), *The Handbook of Informal Language Learning* (pp. 39-55). John Wiley & Sons Ltd. https://doi.org/10.1002/9781119472384.ch3

Dziemianko, A. (2022). The usefulness of graphic illustrations in online dictionaries. *ReCALL. 34*(2), 218-234. doi:10.1017/S0958344021000264

Ensor, T. M., Bancroft, T. D. & Hockley, W. E. (2019). Listening to the picture-superiority effect: Evidence for the conceptual-distinctiveness account of picture superiority in recognition. *Experimental Psychology*. *66*(2), 134-153. https://doi.org/10.1027/1618-3169/a000437

Ensor, T. M., Surprenant, A. M. & Neath, I. (2019). Increasing word distinctiveness eliminates the picture superiority effect in recognition: Evidence for the physical-distinctiveness account. *Memory and Cognition. 47*(1), 182-193. https://doi.org/10.3758/s13421-018-0858-9

Fiorella, L. & Mayer, R. (2021). Principles for reducing extraneous processing in multimedia learning: Coherence, signaling, redundancy, spatial contiguity, and temporal contiguity principles. In R. Mayer & L. Fiorella (Eds.), *The Cambridge Handbook of Multimedia Learning* (pp. 185-198). Cambridge: Cambridge University Press. doi:10.1017/9781108894333.019

Hulstijn, J. H. & Laufer, B. (2001). Some empirical evidence for the involvement load hypothesis in vocabulary acquisition. *Language Learning*. *51*, 539-558.

Kalyuga, S. & Sweller, J. (2021). The redundancy principle in multimedia learning. In R. Mayer & L. Fiorella (Eds.), *The Cambridge Handbook of Multimedia Learning* (pp. 212-220). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108894333.021

Klosa, A. (2015). Illustrations in dictionaries: Encyclopaedic and cultural information in dictionaries. In P. Durkin, (Ed.), *The Oxford Handbook of Lexicography* (pp. 515-531). Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199691630.013.37

Lee, S. S. & Hazita, A. (2021). Integrating Facebook as a Web 2.0 tool in a responsive pedagogy for multimodal oral presentation skills. *GEMA Online® Journal of Language Studies*. *21*(3), 103-124. http://doi.org/10.17576/gema-2021-2103-06

Lew, R. (2010). Multimodal lexicography: The representation of meaning in electronic dictionaries. *Lexikos*. *20*, 290-306. https://doi.org/10.4314/lex.v20i1.62717

Lew, R. & Doroszewska, J. (2009). Electronic dictionary entries with animated pictures: Lookup preferences and word retention. *International Journal of Lexicography*. *22*(3), 239-257, https://doi.org/10.1093/ijl/ecp022

Lew, R., Kaźmierczak, R., Tomczak, E. & Leszkowicz, M. (2018). Competition of definition and pictorial illustration for dictionary users' attention: An eye-tracking study. *International Journal of Lexicography*. *31*(1), 53-77. https://doi.org/10.1093/ijl/ecx002

Longman Dictionary of Contemporary English. https://www.ldoceonline.com/

Mayer, R. E. (2009). *Multimedia Learning* (2nd edition). New York: Cambridge University Press.

Mayer, R. E. (2021). The multimedia principle. In R. Mayer & L. Fiorella (Eds.), *The Cambridge Handbook of Multimedia Learning* (pp. 145-157). Cambridge: Cambridge University Press. https://doi.org/10.1017/9781108894333.015

Oxford Advanced Learner's Dictionary. https://www.oxfordlearnersdictionaries.com/

Paivio, A. (2013). *Imagery and Verbal Processes*. New York: Psychology Press.

Plass, J. L., Chun, D. M., Mayer, R. E. & Leutner, D. (2003). Cognitive load in reading a foreign language text with multimedia aids and the influence of verbal and spatial abilities. *Computers in Human Behavior*. *19*(2), 221-243. https://doi.org/10.1016/S0747-5632(02)00015-8

Ramezanali, N., Uchihara, T. & Faez, F. (2021). Efficacy of Multimodal glossing on second language vocabulary learning: A meta-analysis. *TESOL J. 55*, 105-133. https://doi.org/10.1002/tesq.579

Sung, E. & Mayer, R. E. (2012). When graphics improve liking but not learning from online lessons. *Computers in Human Behavior*. *28*(5), 1618-1625. https://doi.org/10.1016/j.chb.2012.03.026

Svensén, B. (2009). *A Handbook of Lexicography: The Theory and Practice of Dictionary-making*. Cambridge: Cambridge University Press.

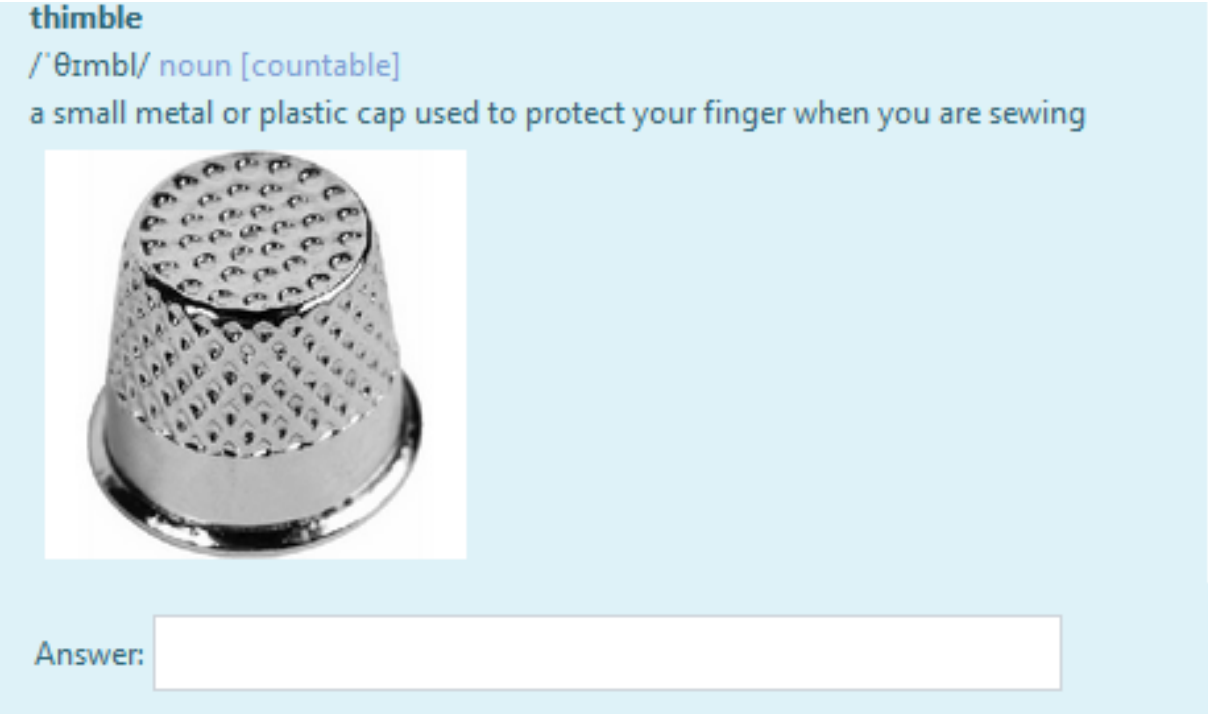
Thomas, P. R. & McKay, J. B. (2010). Cognitive styles and instructional design in university learning. *Learning and Individual Differences.* *20*(3), 197-202. https://doi.org/10.1016/j.lindif.2010.01.002

Tight, M. (2023). Positivity bias in higher education research. *Higher Education Quarterly*. *77*(2), 201-214.

Van den Broek, G. S. E., van Gog, T., Jansen, E., Pleijsant, M. & Kester, L. (2021). Multimedia effects during retrieval practice: Images that reveal the answer reduce vocabulary learning. *Journal of Educational Psychology*. *113*(8), 1587-1608. https://doi.org/10.1037/edu0000499

APPENDIX A

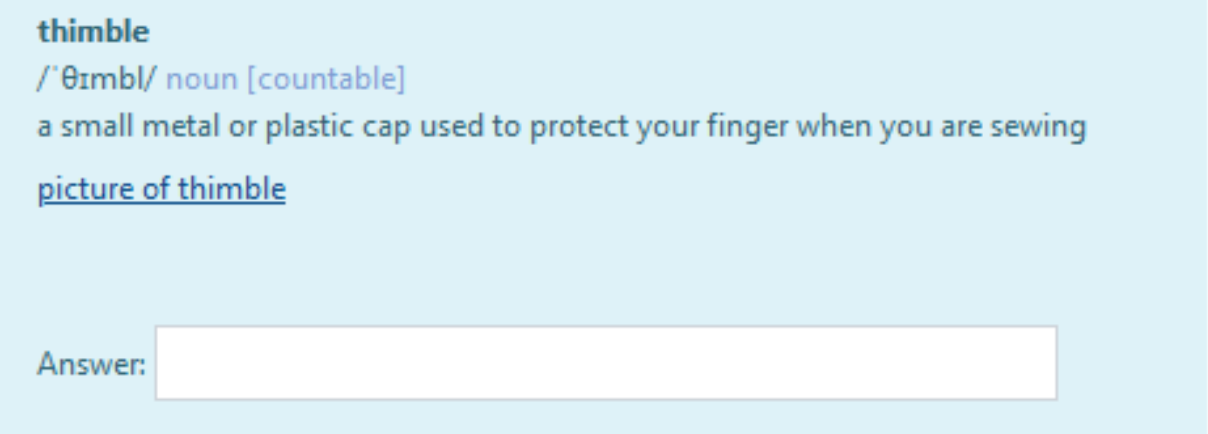
The entry for *thimble* with an instantly visible picture



eISSN: 2550-2131
ISSN: 1675-8021

APPENDIX B

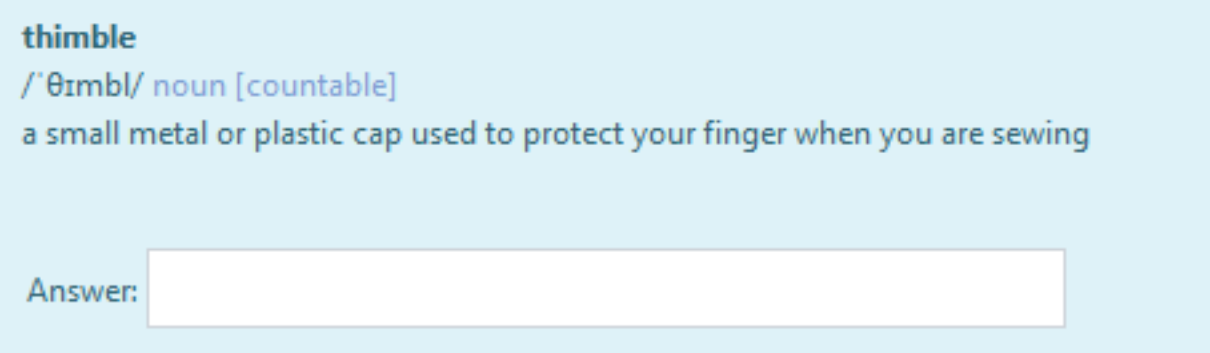
The entry for *thimble* with a hyperlinked picture



eISSN: 2550-2131
ISSN: 1675-8021

APPENDIX C

The entry for *thimble* without any picture



eISSN: 2550-2131
ISSN: 1675-8021

APPENDIX D

Dictionary definitions used in the experiment

Headword	**Definition**[6]
awning	a sheet of strong cloth that stretches out from above a door or window to keep off the sun or rain
barrel	the part of a gun like a tube through which the bullets are fired
bib	a piece of cloth or plastic that you fasten around a baby's neck to protect its clothes while it is eating
copse	a small area of trees growing together
gavel	a small hammer [tool] used by a person in charge of a meeting or an auction, or by a judge in court, in order to get people's attention
hubcap	a round metal cover for the centre of a wheel on a vehicle
mortar	a stone bowl in which substances are crushed with a pestle (=tool with a heavy round end)
nightstick	a short thick stick [tool] that police officers carry as a weapon
pickax	a tool for breaking hard surfaces, with a long wooden handle and a curved metal bar with a sharp point
playpen	a frame with wooden bars or netting that surrounds a small area in which a baby or small child can play safely
plunger	a piece of equipment used for clearing kitchen and bathroom pipes, that consists of a rubber cup fixed to a handle
rolling pin	a long tube-shaped piece of wood used for making pastry flat and thin before you cook it
thimble	a small metal or plastic cap used to protect your finger when you are sewing
trough	a long, narrow open container for animals to eat or drink from
turnstile	a gate [barrier] at the entrance to a public building, stadium, etc. that turns in a circle when pushed, allowing one person to go through at a time

[6] Substitutes of genus terms too close to (*L1 anonymized*) equivalents are marked in square brackets.

eISSN: 2550-2131
ISSN: 1675-8021

ABOUT THE AUTHOR

Anna Dziemianko is a university professor at Adam Mickiewicz University in Poznań, Poland. Her research interests include English lexicography, user studies, syntax in dictionaries, defining strategies and the presentation of lexicographic data on dictionary websites. Recently she has been conducting comparative user research into AI, web search and dictionaries.

eISSN: 2550-2131
ISSN: 1675-8021