Why One and Two
Do Not Make Three:
Dictionary Form Revisited

Anna Dziemianko, *Department of Lexicology and Lexicography,*
Faculty of English, Adam Mickiewicz University,
Poznań, Poland (danna@wa.amu.edu.pl)

Abstract: The primary aim of the article is to compare the usefulness of paper and electronic versions of OALDCE7 (Wehmeier 2005) for language encoding, decoding and learning. It is explained why, in contrast to Dziemianko's (2010) findings concerning COBUILD6 (Sinclair 2008), but in keeping with her observations (Dziemianko 2011) with regard to LDOCE5 (Mayor 2009), the e-version of OALDCE7 proved to be no better for language reception, production and learning than the dictionary in book form.[1] An attempt is made to pinpoint the micro- and macrostructural design features which make e-COBUILD6 a better learning tool than e-OALDCE7 and e-LDOCE5. Recommendations concerning further research into the significance of the medium (paper vs. electronic) in the process of dictionary use conclude the study. The secondary aim which the paper attempts to achieve is to present the status of replication as a scientific research method and justify its use in lexicography.

Keywords: PAPER DICTIONARIES, ELECTRONIC DICTIONARIES, DICTIONARY USE, ENCODING, DECODING, RETENTION, RESEARCH METHODS, REPLICATION, MENUS, HIGHLIGHTING, NOISE, ACCESS, ENTRY LENGTH

Opsomming: Waarom een en twee nie gelyk is aan drie — woordeboekvorm herbeskou. Die primêre doel van die artikel is om die bruikbaarheid van papier- en elektroniese weergawes van OALDCE7 (Wehmeier 2005) te vergelyk vir taalenkodering en -dekodering en die aanleer van taal. Daar word verduidelik waarom, in teenstelling met Dziemianko (2010) se bevindinge betreffende COBUILD6 (Sinclair 2008), maar in ooreenstemming met haar waarnemings (Dziemianko 2011) met betrekking tot LDOCE5 (Mayor 2009), die e-weergawe van OALDCE7 geblyk het nie beter te wees vir taalresepsie en -produksie en die aanleer van taal as die woordeboek in boekvorm nie.[1] 'n Poging word aangewend om die mikro- en makrostrukturele ontwerpkenmerke aan te stip wat e-COBUILD6 'n beter onderrighulpmiddel maak as e-OALDCE7 en e-LDOCE5. Aanbevelings betreffende verdere navorsing oor die belangrikheid van die medium (papier vs. elektronies) in die proses van woordeboekgebruik sluit die studie af. Die sekondêre doel wat die artikel probeer bereik, is om die status van replisering as 'n wetenskaplike navorsingsmetode aan te bied en die gebruik daarvan in die leksikografie te regverdig.

Sleutelwoorde: PAPIERWOORDEBOEKE, ELEKTRONIESE WOORDEBOEKE, WOORDEBOEKGEBRUIK, ENKODERING, DEKODERING, RETENSIE, NAVORSINGSMETODES, REPLISERING, KIESLYSTE, BELIGTING, GERAAS, TOEGANG, LENGTE VAN INSKRYWING

1.   Introduction

1.1   The usefulness of paper and electronic dictionaries

Electronic counterparts of printed monolingual English learners' dictionaries, available on CD-ROMs, online, or — increasingly often — on portable electronic devices, are taken for granted these days. Some of them appear to be quite close to their predecessors in book form (Rogers 1996, Nesi 1999). However, it is strongly stressed that e-dictionaries should not be just electronic remakes of existing printed dictionaries, but should rather be compiled from scratch as genuine electronic tools and take advantage of the wide array of technological possibilities (Nielsen and Mourier 2005: 110). Although contemporary electronic dictionaries, also those based on paper ones, do employ various functionalities offered by the electronic medium and/or the Web technology, further improvements are suggested (Müller-Spitzer et al. 2011, Prinsloo et al. 2011, Lew: In press, Kwary 2012). While the compilation of electronic dictionaries for foreign learners of English independently of (or in place of) paper dictionaries might be just a matter of time, the coexistence of the two media at present raises an obvious question of their relative usefulness in different linguistic tasks.

There is a vast body of studies where the effectiveness of paper and electronic dictionary use is compared.[2] Unfortunately, the results do not permit easy generalisation due to the wide range of user- and task-variables as well as different functionalities and lexicographic data available in the diverse electronic dictionaries used in research. Worse yet, even when the design, dictionary and user differences are neglected, hardly any general picture emerges, either.

First, as regards decoding, no effect of paper and electronic dictionary conditions was found by Nesi (2000), Kobayashi (2007), Koyama and Takeuchi (2007) and Chen (2010, 2012). Electronic dictionaries were however observed to significantly facilitate language reception by Osaki et al. (2003), Osaki and Nakayama (2004) or Dziemianko (2010). In the first two of the abovementioned studies, they also proved to significantly help in identifying contextually appropriate meanings.

Second, different conclusions follow also from the few studies where the influence of paper and electronic dictionaries on language production was tested. In the study by Chen (2010), the subjects were requested to formulate sentences with low-frequency words on the basis of the information found in dictionaries available on hand-held electronic devices and on paper. The results obtained in the encoding task did not depend on the dictionary used. In the study by Dziemianko (2010), in turn, the results from the production task, which consisted in supplying prepositions missing from sentences, were significantly better in the group working with the online version of COBUILD6 than in the one consulting COBUILD6 on paper.

Third, conclusions from studies concerned with the role of paper and electronic dictionaries in vocabulary retention are no less confusing. On the one hand, there are investigations which point to no significant effect of the medium on retention (Koyama and Takeuchi 2003, Osaki et al. 2003, Osaki and Nakayama 2004, Kobayashi 2007, Xu 2010, Chen 2010, 2012). There are also those where the medium proved consequential in this respect. The research conducted by Koyama and Takeuchi (2004) revealed that paper dictionary use resulted in better retention than reference to a portable electronic dictionary. Dziemianko (2010), by contrast, concluded that the consultation of COBUILD6 online resulted in better retention of meaning and collocations than the use of the dictionary in book form. Interestingly enough, the authors of both studies refer to the Involvement Load Hypothesis to account for their findings. Koyama and Takeuchi (2004) suppose that the more demanding process of paper dictionary search is beneficial to retention, in line with the assumption that greater effort means deeper processing, which stimulates retention. Dziemianko (2010), in turn, presumes that the saliency of a dictionary entry on the computer screen as well as the lack of distractions in the form of entries irrelevant to the task at hand, which are bound to be seen on the page of a paper dictionary, induce the cognitive involvement which enhances retention.

Finally, even the replication of a study on the usefulness of paper and electronic dictionaries yields results divergent from those obtained in the original investigation. Dziemianko (2011) adopted the same conditions as those in her pervious study (Dziemianko 2010), except for the dictionary. Instead of COBUILD6, the paper and free online versions of LDOCE5 were offered for consultation. Importantly, the subjects who comprised the other sample were as proficient in English and familiar with paper and electronic dictionaries as those who used COBUILD6 (B2-C1 in CEFR). Despite the same tasks in both experiments, the results from the replication do not confirm previous conclusions. Whereas in the 2010 study it was found that the electronic medium enhanced reception, production and the retention of meaning and collocations, in the more recent investigation dictionary format proved to be inconsequential to the scores on the very same language tasks. In other words, success rates in encoding, decoding and retention were comparable across the two dictionary conditions, i.e., LDOCE5 on paper and online.

To account for the results, Dziemianko (2011) points out that in the free online version of LDOCE5 excessive noise in the form of colourful widgets or animated tower advertisements dwarfs lexicographic data. Such unsolicited (promotional) information in loud colours and different shapes must have diverted the subjects' attention away from dictionary information, which became less prominent and quite inconspicuous. Possibly, then, discerning lexicographic information and extracting it from the glutted website became no less difficult than locating it in a paper dictionary. Unfortunately, neither p-LDOCE5 search nor e-LDOCE5 noise contributed to strengthening the memory trace in a way which could positively influence retention. E-COBUILD6, by

contrast, is much clearer and more neatly organised. In particular, there are no advertisements on its website, and dictionary information looks salient on the screen. Possibly that is why it was more useful than COBUILD6 in book form.

The above brief overview of selected recent studies on paper and electronic dictionary use reveals no obvious conclusions concerning the relative usefulness of these media for language reception, production and retention. As already pointed out above, the investigations differ in tasks, subjects, sampling methods, monitoring dictionary use or quantification, which naturally raises serious comparability issues. Unfortunately, the role of dictionary form in other respects, not discussed in the present paper, such as the speed of dictionary consultation, entry navigation, access paths or even dictionary appreciation is no clearer, either (Dziemianko: In press).

1.2   The role of replication

The wide variety (and inconclusiveness) of research into the relative usefulness of paper and electronic dictionaries highlights the need for systematic replication. Commonly seen as merely repeating a study to see if the same results can be obtained (Lindsay and Ehrenberg 1993: 217, Abbuhl 2012: 296), replication constitutes a crucial scientific method. If carefully designed and conducted, it leads to results that can be generalised, rather than just isolated findings (Lindsay and Ehrenberg 1993: 216). It also increases confidence in the results and helps to establish the reliability of research (Seidlhofer 2003: 215, Gass et al. 2011: 210-211). It is even claimed that "the soundest empirical test of the reliability of data is provided by replicating" (Sidman 1960: 70) and "an isolated study remains virtually meaningless and useless in itself" (Lindsay and Ehrenberg 1993: 218).

Gast (2009: 112) gives three reasons why it is worthwhile to replicate previous studies: to assess the reliability of findings (i.e., internal validity), to assess the generality of findings (i.e., external validity) and to look for exceptions (i.e., conditions under which the original findings do not apply). It is thanks to replication that the margin of error is reduced and confidence that findings are not accidental is strengthened. Systematic replication (whereby a researcher carries out a planned series of studies with systematic changes from one study to another and identifies them as a series) is particularly valuable as it makes it possible to establish the generality of findings, or see how broadly the results can generalise beyond the original experiment (Gast 2009: 111-112, 116, 121). Currently, statistical significance is taken for the ultimate objective of a study, rather than just the first step. A statistically significant result means that it is unlikely to be a product of the sampling error and that it is probably real inasmuch as it is likely to be achieved if the whole population is tested. Yet, "[s]ignificance cannot and does not tell us whether the same result would hold again in a different population or under different conditions. To establish that would require much explicit replication" (Lindsay and Ehrenberg 1993:

218). Put differently, "one statistically significant finding cannot be accepted as 'the truth'; only when results are repeated in other studies can we have greater confidence that our decision to accept or reject a hypothesis is correct" (Abbuhl 2012: 306).

Apart from justifying the need for replications, it is necessary to reflect on how research can be replicated. Replications can be plotted along a continuum which extends from exact, through approximate, to conceptual replications, depending on how closely they resemble the original study (Abbuhl 2012: 297-300). Exact replications (also known as literal, strict or virtual), which consist in repeating the original study exactly or as exactly as possible, are mostly unreal, since no groups of subjects with all their idiosyncratic characteristics and experiences can be duplicated (Lindsay and Ehrenberg 1993: 200, Macaulay 2003: 78). In the case of approximate replications, also known as replications with changes (Abbuhl 2012: 298), the original study is repeated, but some (typically non-major) variables are modified, e.g., population, setting or task, yet comparability is not lost. The aim of such replications is to verify the generalisability of the results from the original study to a new population, setting or modality. In fact, the differences in the conditions of the consecutive studies are of the essence; it is they that make it possible to see whether results hold nevertheless (Lindsay and Ehrenberg 1993: 217).[3] Finally, conceptual or constructive replications diverge from the original study to the largest extent; the same research question is investigated, but a different design is followed. In other words, the findings from an existing study supply the starting point, but researchers develop their own methodology. Such replications make it possible to distinguish between method-specific results and those which can be generalised, but the more variables are changed, the less comparable the original study and its conceptual replication become (Abbuhl 2012: 304).

Unfortunately, replication is held in relatively low esteem; it is considered to be inferior to original research (Umapathy 1987: 170) and lacking in prestige (Campbell 1986: 122). The "pressure to be original" (Park 2004: 194) and the mistaken view that any replication boils down to merely repeating an existing study exactly (Lindsay and Ehrenberg 1993: 220) contribute to the low regard for replication as a scientific method. Although its role in theory development cannot be overestimated, irrespective of whether it supports the tested theory or, perhaps even more importantly — not, replication is seldom undertaken.

As regards research into dictionary use, the value of replication seems to be recognised; the method is claimed to be helpful for improving dictionaries and their usability for language learners (McCreary 2002: 182). However, there are relatively few studies openly acknowledged to be replications of some previous investigations, conducted with different degrees of modification (e.g., Greenbaum et al. 1984, Nesi and Meara 1991, Horst 1995, McCreary and Dolezal 1999, McCreary 2002, McCreary and Amacker 2006, Lew and Doroszewska 2009, Lew and Dziemianko 2006, Lew 2010b, Dziemianko 2011, Chen 2012).[4] Admittedly, the study by Greenbaum et al. (1984), which replicates the

survey by Quirk (1974), shows that the method has been employed in user-centred research for at least three decades. Yet, the small number of replications cannot be unmotivated. It might result from the fact that many studies on dictionary use are simply non-replicable (Hartmann 1987: 27). The low esteem which replication has is probably another factor which discourages researchers. Besides, it is by no means easy to ensure that the original study and its replication are closely comparable. Although replications are considered advisable when the researcher's aim is to make a new study parallel to an existing one (Lew 2002), direct comparisons can still be quite difficult to perform. For one thing, as pointed out above, exact replications are virtually nonexistent. For another, approximate replications, where the conditions whose influence is of particular interest are purposely varied, obviously give a chance for systematic comparison, provided that the other conditions remain unchanged. Yet, it takes time and effort to control the latter, which makes approximate replications difficult to accomplish successfully. Finally, the fact that not many researchers openly wish their investigations could be replicated in the future (McCreary and Dolezal 1999, Al-Ajmi 2002, Dziemianko 2006, Lew and Dziemianko 2006, Koyama and Takeushi 2007, Tono 2011) suggests that, in fact, the awareness of the benefits which can be derived from replication might need to be raised. It is tacitly assumed that replication "carries more risk than potential reward for both the replicator and the originator of the research" (Park 2004: 194). After all, failure to obtain the same result might be seen as a proof that the latter was wrong, or that the former is incompetent (Lindsay and Ehrenberg 1993: 218).

Indeed, although replications are said to be crucial "to distinguish the spurious from the real" (Abbuhl 2012: 306), there is a strong bias against negative findings. The file-drawer syndrome prevents the publication of many replications which do not support previous findings (Lindsay 1990, Park 2004: 194). Admittedly, confirming replications (whose results agree with those from the original investigation) are valuable inasmuch as they make the corroborated findings more credible. Yet, disconfirming replications are by no means worthless. Assuming that research is conversation, they prove that there is still a need to discuss the issue which turns out to be more complex than it seemed (Lindsay and Ehrenberg 1993: 218, Abbuhl 2012: 306). Besides, accounting for the divergent results provides ample scope for originality.

In an attempt to meet the need for systematic replication in research into dictionary use, the next part of the paper describes the second approximate replication of the study by Dziemianko (2010) and the obtained results.

2.   The replication

2.1   Aim

As mentioned above, Dziemianko (2010) found that e-COBUILD6 was more useful in L2 reception, production and learning (retention of meaning and col-

locations) than COBUILD6 on paper. The results were not confirmed by the first approximate replication carried out by the author herself, where the paper and free online versions of LDOCE5 were employed. No statistically significant differences between the results obtained in the paper and electronic dictionary conditions were noted then in any task (Dziemianko 2011).

The aim of the present study is twofold. First, an attempt is made to investigate the usefulness of OALDCE7 in paper and electronic form for language reception, production and learning. Second, Dziemianko's (2010) findings concerning COBUILD6 are compared with those obtained from both replications.

The following research questions are answered:

1.   Which version, paper or electronic, of OALDCE7 is more useful for L2 reception, production and learning (retention of meaning and collocations)?

2.   Which dictionary (OALDCE7, LDOCE5 or COBUILD6) and in which form is most helpful in dealing with receptive and productive tasks, and which is the best learning tool?

The CD-ROM and regular printed versions of OALDCE7 were used. The choice of the seventh edition of the dictionary, rather than the latest one, was motivated by the number of copies of the dictionary in book form available in the experimental setting as well as by the functionalities of the electronic version. For one thing, there were enough paper copies of OALDCE7 to go around in the groups in which the study was conducted. For another, the CD-ROM version of OALDCE7 made it possible to see whether some search facilities which it offers (such as automatic scrolling or highlighting the entry for the looked up word, not available in the online versions of LDOCE5 and COBUILD6) matter to dictionary users.

2.2   Materials and subjects

The materials used by Dziemianko (2010), i.e., the pretest, questionnaire, test and unexpected delayed post-test, were employed. The subjects did the same receptive and productive tasks as in the original study. In the receptive task, they explained the meaning of nine nouns and phrases (*backgammon, booby prize, clampdown, collateral damage, down under, dream ticket, flapjack, onus, outcrop*). The productive part consisted in completing sentences with prepositions removed from nine collocations (*on the blink, in cahoots with, up the creek, at gunpoint, wreak havoc on, in the offing, in the pipeline, under sedation, on the trot*). Both tasks featured in the pretest, test proper and retention test. The pretest served to sift out the cases where the subjects knew correct answers. It was accompanied by a questionnaire to gain an insight into the subjects' familiarity with dictionary formats. Once the pretest and the questionnaire had been completed, the test was administered. In the test, the subjects did the same tasks as in the pretest, but with access to either paper or electronic OALDCE7. In the

delayed retention test conducted two weeks later, the sequence of the target structures was reshuffled and no access to dictionaries was allowed. The study was carried out in regular class time (45 minutes).

Great care was taken to ensure that the subjects were as proficient as those in the original research. Overall, 86 students of English (B2-C1 in CEFR) at Poznań University took part in the study; 42 of them consulted p-OALDCE7 and the other 44 the e-OALDCE7. The subjects' proficiency was determined on the basis of the grammar test in the practical English exam taken on a yearly basis. Importantly, the information obtained from the questionnaire indicates that in both experimental conditions the proportions of subjects consulting paper and electronic dictionaries as a matter of routine were comparable (the p-OALDCE7 group: students using paper dictionaries 66.7%, students using electronic dictionaries 69.0%, p=0.83; the e-OALDCE7 group: students using paper dictionaries 63.6%, students using electronic dictionaries 68.2%, p=0.68; Z test for dependent samples, non-significant, alpha-level=0,05).

2.3   Results

2.3.1   Research question one (the usefulness of OALDCE7)

The mean proportions of correct answers in the main and retention tests are illustrated in Figure 1. The results of the repeated-measures ANOVAs for both tests are given in Table 1.[5]
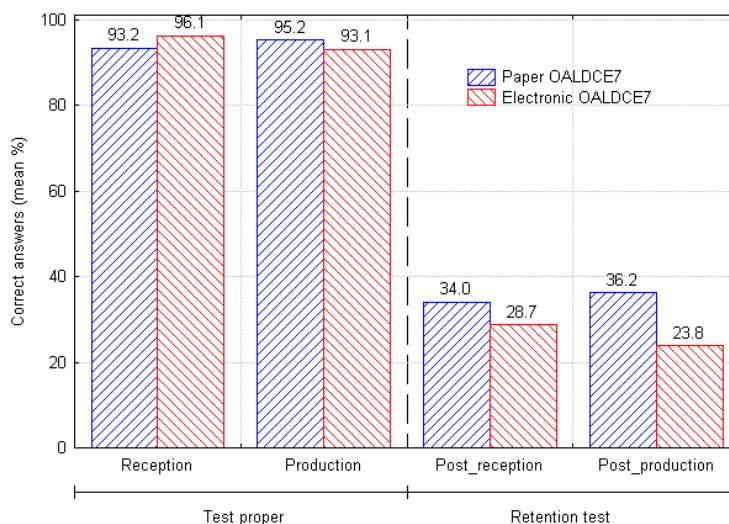

Figure 1: Results obtained in the main and retention tests (OALDCE7)

Factors	Test proper			Retention		
	F	**p**	**partial η²**	**F**	**p**	**partial η²**
FORM	0.0	0.880	0.001	1.3	0.272	0.075
TASK	0.0	0.880	0.001	0.1	0.790	0.005
TASK*FORM	0.7	0.422	0.041	0.5	0.499	0.029

Table 1: ANOVA summary results (main and retention tests): OALDCE7

In each test, the scores on each task were comparable among the users of paper and electronic versions of OALDCE7 at the accepted level of significance (alpha=0.05). In the main test, the subjects provided over 90 percent of correct answers in each task. The differences in the main test scores between the paper and electronic conditions approximated 3 percent for reception (paper dictionary (PD): 93.2%, electronic dictionary (ED): 96.1%) and production (PD: 95.2%, ED: 93.1%). In the retention test, in turn, active recall in the paper dictionary group (PD: 36.2%) was about half as good again as in the electronic dictionary group (ED: 23.8%). For passive recall, the difference, still in favour of the paper dictionary, amounted to 18 percent (PD: 34% vs. ED: 28.7%). While the differences were statistically insignificant in the light of the ANOVA, their scale seems to suggest that if the sample had been bigger, they might have gained significance. Yet, the low values of the estimate of effect size (partial η²) computed for the retention test show that the size of each investigated main and interaction effect was very small, which means that only a modest proportion of the respective variance can be accounted for by a given (main or interaction) effect. In particular, only 7.5% of the between subjects variance in retention scores can be attributed to dictionary form (FORM, partial η²=0.075).

2.3.2 Research question two (comparative usefulness of OALDCE7, LDOCE5 and COBUILD6)

2.3.2.1 Test proper

ANOVA results for the main test scores achieved by OALDCE7, LDOCE5 and COBUILD6 users are given in Table 2.

Factor	F	p	partial η²
DICTIONARY	0.3	0.728	0.013
FORM	2.8	0.099	0.056
DICTIONARY*FORM	3.5	0.039*	0.127
TASK	0.0	0.999	0.000
TASK*DICTIONARY	0.0	0.965	0.001
TASK*FORM	0.4	0.553	0.007
TASK*DICTIONARY*FORM	0.4	0.690	0.015

Table 2: ANOVA (main test): OALDCE7, LDOCE5 and COBUILD6

The data show that only one interaction (DICTIONARY x FORM) was statistically significant (p=0.039, alpha=0.05; partial η^2=0.127). To explore it in more depth, Table 3 shows the results of the Tukey Honest Significant Difference test.[6] Figure 2 illustrates the interaction graphically.

DICTIONARY	FORM	Mean %		
COBUILD6	paper	92.1	****	
OALDCE7	paper	94.2	****	****
OALDCE7	electronic	94.6	****	****
LDOCE5	electronic	95.1	****	****
LDOCE5	paper	95.9	****	****
COBUILD6	electronic	98.6		****

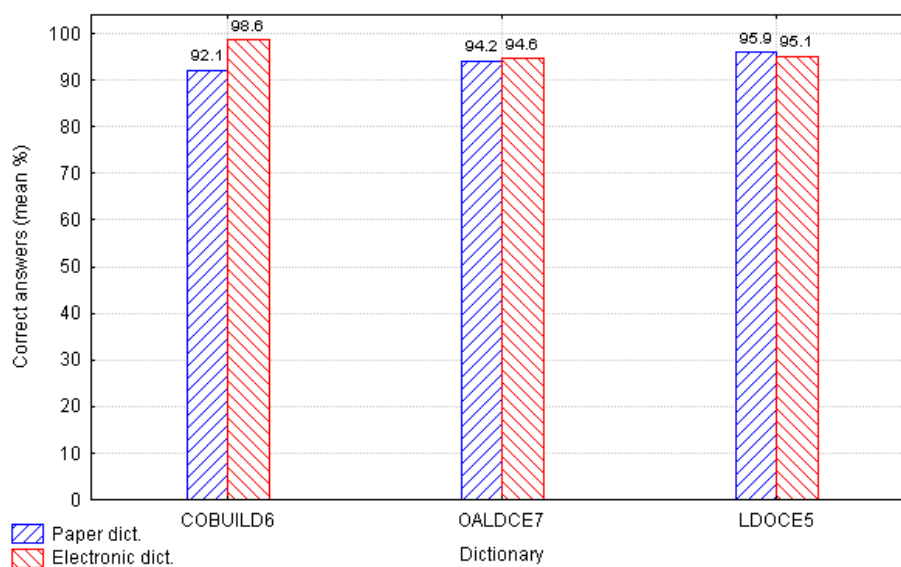**Table 3:** Tukey HSD test: DICTIONARY x FORM (main test)

Figure 2: DICTIONARY x FORM: Correct answers (mean %) in the main test

The results of the Tukey HSD test reveal that in the main test, e-COBUILD6 (98.6%) was more useful than COBUILD6 on paper (92.1%, cf. Dziemianko 2010). However, both versions of LDOCE5 and OALDCE7 were comparably helpful.

2.3.2.2   Retention test

Summary ANOVA results for the retention test are collated in Table 4.

Factor	F	p	partial η²
DICTIONARY	20.9	0.000**	0.465
FORM	0.3	0.564	0.007
DICTIONARY*FORM	3.1	0.054	0.115
TASK	8.8	0.005**	0.155
TASK*DICTIONARY	1.8	0.184	0.068
TASK*FORM	0.1	0.717	0.003
TASK*DICTIONARY*FORM	0.3	0.745	0.012

Table 4: ANOVA (retention test): OALDCE7, LDOCE5 and COBUILD6

The data indicate that the main effects produced by DICTIONARY (p=0.000) and TASK (p=0.005) were statistically highly significant at alpha=0.05. Also, the effect sizes associated with these factors were large and medium, respectively (DICTIONARY: partial η^2=0.465, TASK: partial η^2=0.155). Table 5 gives the results of the Tukey HSD test for the two significant effects, illustrated graphically in Figure 3.

DICTIONARY	Mean %			TASK	Mean %	
OALDCE7	30.7	****		Post_production	39.0	****
LDOCE5	37.5	****		Post_reception	47.9	
COBUILD6	62.2		****			

Table 5: Tukey HSD test: DICTIONARY and TASK (retention test)
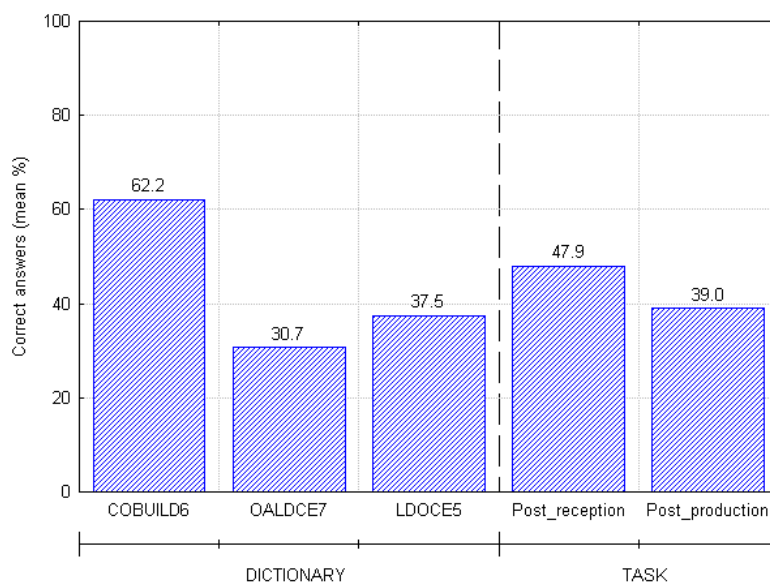

Figure 3: DICTIONARY and TASK: Correct answers (mean %) in the retention test

First, the best retention was observed in the COBUILD6 group, where it exceeded 62% and was significantly better than in the other dictionary conditions. The retention results obtained after reference to OALDCE7 (30.7%) and LDOCE5 (37.5%), only about half as good as among COBUILD6 users, were comparable. Second, meaning (47.9%) proved much easier to remember than collocations (39.0%); passive recall was over one fourth more successful than active recall, and the difference was statistically significant at alpha=0.05.

The interaction Dictionary x Form, which is not quite statistically significant but approaches significance (p=0.054, alpha=0.05; partial η^2=0.115, Table 4), also merits further investigation. Results of the Tukey HSD test for the interaction in question are collated in Table 6. The relevant mean proportions are illustrated in Figure 4.

DICTIONARY	FORM	Mean %			
OALDCE7	electronic	26.3	****		
OALDCE7	paper	35.1	****	****	
LDOCE5	electronic	37.4	****	****	
LDOCE5	paper	37.6	****	****	
COBUILD6	paper	54.0		****	****
COBUILD6	electronic	70.3			****

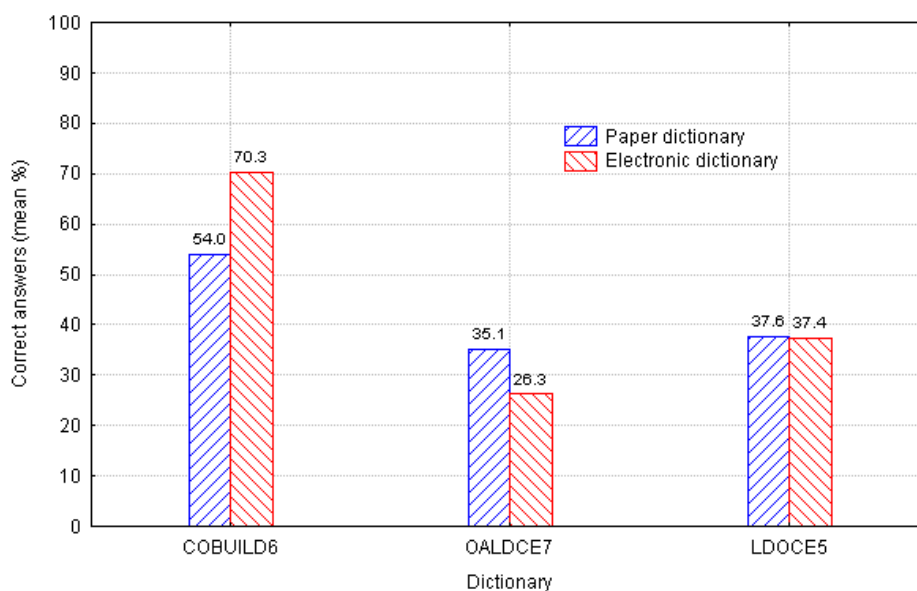**Table 6:** Tukey HSD test: DICTIONARY x FORM (retention test)

Figure 4: DICTIONARY x FORM: Correct answers (mean %) in the retention test

Three main conclusions follow from the data. First, it transpires that there were no significant differences in retention between the users of paper and electronic versions of LDOCE5 and OALDCE7. Second, reference to e-COBUILD6 yielded significantly better retention results than reliance on the other e-dictionaries; e-COBUILD users (70.3%) remembered about 90 and 170 percent more than the subjects who referred to e-LDOCE5 (37.4%) and e-OALDCE7 (26.3%), respectively. Third, retention among the users of LDOCE5, OALDCE7 and COBUILD6 on paper was comparable. Even though reference to p-COBUILD6 (54.0%) yielded retention results which were about half as good again as those obtained after the consultation of p-OALDCE7 (35.1%) and p-LDOCE5 (37.6%), on the Tukey HSD test, the difference was not statistically significant at alpha=0.05.

3.   Discussion

Obviously, the replications led to conclusions different from those obtained in the original study. First of all, in contrast to Dziemianko's (2010) findings concerning COBUILD6, the e-versions of OALDCE7 and LDOCE5 proved to be no better for language reception, production and learning than the dictionaries in book form. Second, e-COBUILD6 was found to be a better learning tool than e-OALDCE7 and e-LDOCE5. It is thus necessary to reflect on the micro- or macrostructural features and factors not intrinsic to any dictionary structures which contributed to the success achieved with the help of e-COBUILD6 and prevented e-OALDCE7 and e-LDOCE5 from being likewise useful.

First of all, it is worth noting that the e-COBUILD6 website is quite crude; it is made up of the search window followed by the entry for the looked up word and a few buttons on the right (to be clicked if users wish to expand their vocabulary, customise the dictionary or get help). In e-OALDCE7, in turn, the entry for the looked up word, if short enough, is displayed along with the entries which follow it. This form of presentation resembles the paper dictionary and diverges from the approach adopted by e-COBUILD6, where only the entry for the looked-up word can be seen on the screen. Undoubtedly, the view of entries in a sequence must have naturally dispersed the subjects' attention and disturbed concentration. Such interface dissimilarities might be a reason why the retention scores of e-COBUILD6 users were better than those of the e-OALDCE7 group. The same factor might also account for the lack of any statistically significant difference between the results obtained with the help of the electronic and paper versions of OALDCE7 in the main and retention tests. In e-LDOCE5, by contrast, the entries for the headwords which follow the looked up word are not displayed, but the website overflows with noise, thereby deflecting users from the dictionary itself and making lexicographic data much less salient and distinct (cf. Dziemianko 2011 and section 1.1). This could be a possible reason why e-LDOCE5 was no more helpful in any experimental task than p-LDOCE5.

Apart from the examination of interfaces, item analysis was conducted with a view to explaining the observed results. Looking at the data for individual target items, Dziemianko (2011) drew interesting conclusions about the role of clickable menus in e-LDOCE5, i.e., vertical menus which consist of several matches, each of which is hyperlinked to an entry or subentry. Figure 5 shows such a menu for *blink*.[7]
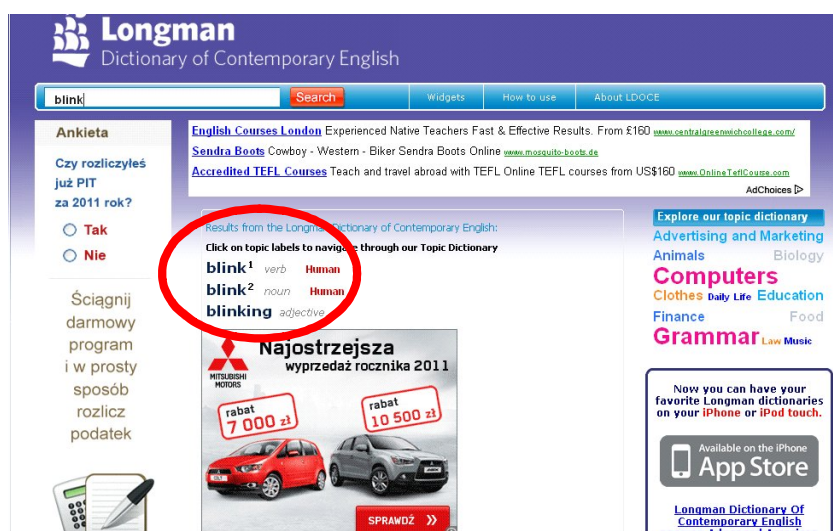

Figure 5: The menu for *blink* in e-LDOCE5 (circled)

Accessing noun phrases through clickable menus in e-LDOCE5 was found to severely impede reception in comparison with p-LDOCE5. No similar effect of clickable menus was identified on production. However, they proved seriously detrimental to passive and active recall (in comparison with the menu-less access paths in e-COBUILD6). Dziemianko (2011) hypothesised that the mechanical rather than cognitive effort invested into coping with the hierarchical, step-wise outer access structure in e-LDOCE5, at which stage *relevant* semantic information is not processed yet, did not strengthen the memory trace, but actually prevented successful reception and retention.

OALDCE7 does not feature clickable menus similar to those in e-LDOCE5, but it offers a different functionality — automatic scrolling whereby the looked up compound, phrase or idiom not given the headword status is immediately shown at the top of the screen. It is worth remembering that the results obtained by e-OALDCE7 users in the receptive task in the test proper were on average 3 percent better than in the group consulting p-OALDCE7 (cf. Figure 1). The largest difference in decoding scores between the experimental conditions was observed for *down under*, which in the paper version is given as the

sixth of the seven idioms explained at the end of the 12-sense entry for *down* (adv). In the electronic version, in turn, *down under* is immediately shown at the top of the computer screen, its identical placement in the entry for *down* notwithstanding. Automatic scrolling to the phrase resulted in 29 percent better score. Even though not quite statistically significant (p=0.080, Z test for independent samples, two-tailed, alpha=0.05), the difference was much beyond the aforementioned average (3 percent).

Interesting observations can be made about active recall, which was on average four times better among e-COBUILD6 users than among the subjects consulting e-OALDCE7. Item analysis reveals that this difference owes most to the collocation *up the creek*, retained over 13 times more often by the e-COBUILD6 group. This tremendous and statistically significant difference (p=0.000, Z test for independent samples, two-tailed, alpha=0.05) results most probably from the fact that the search for *creek* in e-OALDCE7 yields two matches. The first of them, a proper name irrelevant to the task at hand (*Creek — a member of a Native American people, many of whom now live in the US state of Oklahoma*), is highlighted, as shown in Figure 6. In e-COBUILD6, by contrast, *up the creek* constitutes the third subentry of *creek*, none of which is highlighted. The tentative conclusion which can be drawn from the data is that highlighting the entry for the searched word by default does not pay off when its homograph, treated in a separate entry (which is not highlighted), happens to be what dictionary users need. In such a case, default highlighting can result in immensely poorer retention.

Figure 6: The highlighted entry for *Creek* in e-OALDCE7

The second largest difference in active recall between the groups using e-COBUILD6 and e-OALDCE7 was observed for *on the trot*. The subjects who consulted the latter dictionary found the phrase in the section devoted to idioms, located at the end of the entry which consists of four verb senses, a sub-entry for the phrasal verb *trot (sth) out* and two noun senses. In e-COBUILD6, in turn, *on the trot* constitutes the third (final) subentry, but the two preceding verb subentries are quite short. Judging by the number of senses which separate the headword from the target phrase, the search path in e-OALDCE7 is three times longer than in e-COBUILD6. This might be a reason why e-COBUILD6 users were about 7 times more successful in active recall than the subjects who referred to e-OALDCE7. Apparently, then, the effort exerted to locate the phrase, as measured by entry length, is inversely related to active recall. In other words, the longer the entry is, the lower the chances of successful retention becomes. Yet, this hypothesis needs to be verified in further studies. It is worth noting that the results obtained in the main test for *on the trot* indicate that the phrase was extracted with comparable success from both dictionaries (97% in e-COBUILD6 and 97.7% e-OALDCE7, p=0.840 Z test for independent samples, alpha=0.05). Such an observation supports the surprising findings by Nesi and Tan (2011), who noted that the senses at the end of the entry are identified with the greatest speed and accuracy by dictionary users, followed by those which are given first. The regularity observed in the entry for *trot* not only confirms the saliency of the entry-final position, but also suggests that the effect persists regardless of entry length.[8] Nonetheless, it transpires that the saliency of entry-final positions has widely different consequences for entry navigation (i.e., finding the needed information) on the one hand, and retention on the other.

The foregoing discussion makes it possible to formulate a few suggestions for further research into e-dictionary use. First, it appears that the role of noise on dictionary websites is worth looking into. It goes without saying that advertisements make online dictionaries accessible to anyone free of charge. No wonder, then, that ad-supported online dictionaries are enjoying considerable popularity.[9] Nonetheless, it is open to question whether dictionary websites with and without advertisements are comparably useful. The tentative conclusion following from the present investigation is that unsolicited promotional material diverts users' attention from lexicographic data and actually deprives an online dictionary of much of its usefulness. Second, the effect of the hierarchical nature of data display in electronic dictionaries on retention is another promising area of research. The above assessment of the possible influence of clickable menus on retention, and active recall in particular, is quite pessimistic, but systematic manipulation of fabricated microstructures is necessary to get a deeper insight into the actual significance of clickable menus in electronic dictionaries. Admittedly, research into clickable menus as access facilitators was taken up by Lew and Tokarek (2010), who concluded that such tools help lower-level students navigate a dictionary entry and get to the right

sense, but are of no real benefit to advanced users. Apart from regular clickable menus, the authors looked into the usefulness of clickable menus where the target sense was automatically highlighted. Such menus proved comparably useful at both proficiency levels. However, no attempt has yet been made to investigate the effect of menus in paper or electronic dictionaries on retention (cf. Nesi and Tan 2011, Tono 2011, Lew 2010b). Third, it might be useful to explore the influence which highlighting entries in electronic dictionaries exerts on active and passive recall in the case of homographs treated in different entries, only one of which is highlighted. While highlighting entries by default seems attractive, it transpires that bringing out the entry which does not feature the information that a user wishes to find has a negative impact on retention. At this stage it is worth distinguishing between highlighting entries and highlighting specific senses. The latter was found a welcome navigation enhancement in polysemous microstructures, where it assists users in reaching the relevant sense more quickly and accurately (Lew and Tokarek 2010).

Unfortunately, the present study is not free from limitations. First, a number of subject variables were not controlled. Only the subjects' familiarity with dictionary formats and proficiency in English were taken into consideration, since they were considered most likely to immediately affect dictionary use and language skills. Besides, it needs to be remembered that real dictionaries rather than systematically manipulated microstructures were employed. Such an approach resulted in a naturalistic task, but it made it difficult to pin down specific factors responsible for the observed effects. To establish the role of selected factors, entries need to be fabricated and systematically manipulated, which no doubt creates more tightly controlled, albeit more artificial, conditions. The use of actual paper and electronic dictionaries also means that dictionary form alone may not be the key factor which determines the effectiveness of dictionary consultation. Specific solutions adopted and form-independent typographical structural indicators (Gouws 2003), such as font size and colour, line spacing or layout, which remained beyond control in the studies discussed above, can play an important role in dictionary use (cf. Lew 2010a: 294, Nesi: In press). To reduce their influence, printouts of the electronic dictionary screen display could be used instead of a real paper dictionary. Such task operationalisation could help to isolate the factor of dictionary form (on-screen vs. paper) and free it of the effect produced by typographical structural indicators (cf. Chen 2012). Nonetheless, in this way the paper dictionary user is also largely helped inasmuch as only mini-dictionaries covering the key items rather than complete paper dictionaries are typically produced from printouts, which seriously limits and simplifies outer access (Bergenholtz and Gouws 2007: 243).[10]

All in all, whereas the present study proved to be quite exploratory in nature at the stage of item analysis, it made it possible to develop a few testable hypotheses which merit further attention. In this way it hopefully confirmed that replication as a research method does not entail lack of originality. Impor-

tantly, it also showed that approximate replication helps to validate theories and substantiate generalisations. Ultimately, it is replications that contribute to making research a truly accretive process whereby knowledge is accumulated and consolidated over time, and, by the same token, prevent a discipline from being composed of scattered hypotheses and observations (cf. Santos 1989).

Endnotes

1. In the article, where differences between the dictionaries and their forms are of the utmost importance, the aforementioned, generally accepted acronyms are used for the sake of convenience. In the list of references, full bibliographic information is provided under the names of the respective dictionary editors, not repeated below: Mayor (2009) — LDOCE5, Sinclair (2008) — COBUILD6 and Wehmeier (2005) — OALDCE7.
2. For an overview, see Dziemianko (In press).
3. Naturally, the greater the differences are, the higher the risk that the effect will not be replicated. Yet, if it is confirmed, its generality increases (Gast 2009: 111). By the same token, "failure to replicate or follow up on studies with different populations and in different contexts may lead to de facto generalisation" (Duff 2006: 71).
4. Compare similar remark made by Chi (2009: 14), who also notes the paucity of replications in the field of dictionary use.
5. In any ANOVA discussed below, TASK was the repeated-measures factor.
6. All the means connected by (****) in one column are not different from each other at p=0.05.
7. The screenshot also gives an insight into the amount of noise on the e-LDOCE5 website.
8. Only five-sense entries were employed in the study by Nesi and Tan (2011).
9. See also Lew (2011).
10. Proponents of the Involvement Load Hypothesis would no doubt claim that simplified outer access can affect retention results, the assumption being that any effort invested in word search, including mechanical page turning and scanning running heads, can increase the chances of successful retention. On the other hand, it is suggested that not any involvement, but only semantic involvement affects vocabulary retention in the process of dictionary use. The aforementioned, largely automatic stages of paper dictionary look-up, might not yet evoke adequate semantic or cognitive involvement to influence vocabulary retention (cf. Craik and Lockhart 1972, Dziemianko: In press). Besides, printouts of an electronic dictionary prevent users from scanning entries close to the target ones, which might also affect retention (Chen 2012).

References

Abbuhl, R. 2012. Why, When, and How to Replicate Research. Mackey, A and S.M. Gass (Eds.). 2012. *Research Methods in Second Language Acquisition: A Practical Guide*: 296-312. Oxford: Blackwell Publishing.

Al-Ajmi, H. 2002. Which Microstructural Features of Bilingual Dictionaries Affect Users' Look-Up Performance? *International Journal of Lexicography* 15(2): 119-131.

Bergenholtz, H. and R.H. Gouws. 2007. The Access Process in Dictionaries for Fixed Expressions. *Lexicographica* 23: 237-260.

Campbell, D.T. 1986. Science's Social System of Validity-Enhancing Collective Belief Change and the Problems of the Social Sciences. Fiske, D.W. and R.A. Shweder (Eds.). 1986. *Metatheory in Social Science: Pluralisms and Subjectivities*: 108-135. Chicago: University of Chicago Press.

Chen, Y. 2010. Dictionary Use and EFL Learning. A Contrastive Study of Pocket Electronic Dictionaries and Paper Dictionaries. *International Journal of Lexicography* 23(3): 275-306.

Chen, Y. 2012. Dictionary Use and Vocabulary Learning in the Context of Reading. *International Journal of Lexicography* 25(2): 216-247.

Chi, A. 2009. *Empirical Study of the Efficacy of Integrating the Teaching of Dictionary Use into a Tertiary English Curriculum in Hong Kong.* Hong Kong: Hong Kong University Press.

Craik, F.I.M. and R.S. Lockhart. 1972. Levels of Processing: A Framework for Memory Research. *Journal of Verbal Learning and Verbal Behavior* 11: 671-684.

Duff, P. 2006. Beyond Generalizability: Contextualization, Complexity, and Credibility in Applied Linguistics Research. Chalhoub-Deville, M., C. Chapelle and P. Duff (Eds.). 2006. *Inference and Generalizability in Applied Linguistics: Multiple Research Perspectives*: 65-95. Dordrecht: John Benjamins.

Dziemianko, A. 2006. *User-Friendliness of Verb Syntax in Pedagogical Dictionaries of English*. Tübingen: Max Niemeyer Verlag.

Dziemianko, A. 2010. Paper or Electronic? The Role of Dictionary Form in Language Reception, Production and the Retention of Meaning and Collocations. *International Journal of Lexicography* 23(3): 257-273.

Dziemianko, A. 2011. Does Dictionary Form Really Matter? Akasu, K. and U. Satoru, (Eds.). 2011. *ASIALEX 2011 Proceedings. Lexicography: Theoretical and Practical Perspectives*: 92-101. Kyoto: Asian Association for Lexicography.

Dziemianko, A. In press. On the Use(fulness) of Paper and Electronic Dictionaries. Granger, S. and M. Paquot (Eds.). *Electronic Lexicography*. Oxford: Oxford University Press.

Gass, S., A. Mackey and L. Ross-Feldman. 2011. Task-Based Interactions in Classroom and Laboratory Settings. *Language Learning* 61(Suppl. 1): 189-220.

Gast, D.L. (Ed.). 2009. *Single Subject Research Methodology in Behavioral Sciences*. New York: Routledge.

Gouws, R. 2003. Types of Articles, their Structure and Different Types of Lemmata. Van Sterkenburg, P. (Ed.). 2003. *A Practical Guide to Lexicography*: 34-43. Amsterdam: John Benjamins.

Greenbaum, S., C.F. Meyer and J. Taylor. 1984. The Image of the Dictionary for American College Students. *Dictionaries*: *The Journal of the Dictionary Society of North America* 6: 31-52.

Hartmann, R.R.K. 1987. Four Perspectives on Dictionary Use: A Critical Review of Research Methods. Cowie, A.P. (Ed.). 1987. *The Dictionary and the Language Learner: Papers from the EURALEX Seminar at the University of Leeds, 1-3 April 1985*: 11-28. Tübingen: Max Niemeyer Verlag.

Horst, M.E. 1995. Who Get Summit Mark in the Class? Productive Use of Definitions and Examples in Beginning Learners of English. *Paper presented at TESOL Arabia, Al-Ayn, United Arab Emirates*.

Kobayashi, C. 2007. Comparing Electronic and Printed Dictionaries: Their Effects on Lexical Processing Strategy Use, Word Retention and Reading Comprehension. Bradford-Watts, K. (Ed.).

2007. *JALT2006 Conference Proceedings*: 657-671. Tokyo: JALT.

Koyama, T. and O. Takeuchi. 2003. Printed Dictionaries vs. Electronic Dictionaries: A Pilot Study on How Japanese EFL Learners Differ in Using Dictionaries. *Language Education and Technology* 40: 61-79.

Koyama, T. and O. Takeuchi. 2004. Comparing Electronic and Printed Dictionaries: How the Difference Affected EFL Learning. *JACET Bulletin* 38: 33-46.

Koyama, T. and O. Takeuchi. 2007. Does Look-Up Frequency Help Reading Comprehension of EFL Learners? Two Empirical Studies of Electronic Dictionaries. *CALICO Journal* 25(1): 110-125.

Kwary, D.A. 2012. Adaptive Hypermedia and User-Oriented Data for Online Dictionaries: A Case Study on an English Dictionary of Finance for Indonesian Students. *International Journal of Lexicography* 25(1): 30-49.

Lew, R. 2002. Questionnaires in Dictionary Use Research: A Reexamination. Braasch, A. and C. Povlsen (Eds.). 2002. *Proceedings of the Tenth EURALEX International Congress, EURALEX 2002, Copenhagen, Denmark, August 12-17, 2002, Vol.1*: 267-271. Copenhagen: Centre for Sprogteknologi, Copenhagen University.

Lew, R. 2010a. Multimodal Lexicography: The Representation of Meaning in Electronic Dictionaries. *Lexikos* 20: 290-306.

Lew, R. 2010b. Users Take Shortcuts: Navigating Dictionary Entries. Dykstra, A. and T. Schoonheim (Eds.). 2010. *Proceedings of the XIV Euralex International Congress*: 1121-1132. Ljouwert: Afûk.

Lew, R. 2011. Online Dictionaries of English. Fuertes-Olivera, P.A. and H. Bergenholtz (Eds.). 2011. *E-Lexicography: The Internet, Digital Initiatives and Lexicography*: 230-250. London and New York: Continuum.

Lew, R. In press. How Can We Make Electronic Dictionaries More Effective? Granger, S. and M. Paquot (Eds.). In press. *Electronic Lexicography*. Oxford: Oxford University Press.

Lew, R. and J. Doroszewska. 2009. Electronic Dictionary Entries with Animated Pictures: Lookup Preferences and Word Retention. *International Journal of Lexicography* 22(3): 239-257.

Lew, R. and A. Dziemianko. 2006. A New Type of Folk-Inspired Definition in English Monolingual Learners' Dictionaries and its Usefulness for Conveying Syntactic Information. *International Journal of Lexicography* 19(3): 225-242.

Lew, R. and P. Tokarek. 2010. Entry Menus in Bilingual Electronic Dictionaries. Granger, S. and M. Paquot (Eds.). 2010. *Elexicography in the 21st Century: New Challenges, New Applications*: 193-202. Louvain-la-Neuve: Cahiers du CENTAL.

Lindsay, R.M. 1990. An Examination of the 'Negative Result' Bias: Some Empirical Evidence and a Statistical Redress. *Paper read at Canadian Academic Accounting Conference, Victoria*.

Lindsay, R.M. and A.S.C. Ehrenberg. 1993. The Design of Replicated Studies. *The American Statistician* 47(3): 217-228.

Macaulay, R. 2003. Repeat after Me: The Value of Replication. *International Journal of English Studies* 3(1): 77-92.

Mayor, M. (Ed.). 2009⁵. *Longman Dictionary of Contemporary English*. Harlow: Longman. *http://www.ldoceonline.com*. [LDOCE5]

McCreary, D.R. 2002. American Freshmen and English Dictionaries: I Had Aspersions of Becoming an English Teacher. *International Journal of Lexicography* 15(3): 181-205.

McCreary, D.R. and E. Amacker. 2006. Experimental Research on College Students' Usage of Two Dictionaries: A Comparison of the *Merriam-Webster Collegiate Dictionary* and the *Macmillan English Dictionary for Advanced Learners*. Corino, E., C. Marello and C. Onesti (Eds.). 2006. *Atti del XII Internazionale Congresso di Lessicografia, Torino, 6-9 Settembre 2006*. 871-885. Allessandria: Edizioni dell'Orso.

McCreary, D.R. and F.T. Dolezal. 1999. A Study of Dictionary Use by ESL Students in an American University. *International Journal of Lexicography* 12(2): 107-146.

Müller-Spitzer, C., A. Koplenig and A. Töpel. 2011. What Makes a Good Online Dictionary? Empirical Insights from an Interdisciplinary Research Project. Kosem, I. and K. Kosem (Eds.). 2011. *Electronic Lexicography in the 21st Century: New Applications for New Users. Proceedings of eLex 2011*: 203-208. Ljubljana: Trojina, Institute for Applied Slovene Studies.

Nesi, H. 1999. A User's Guide to Electronic Dictionaries for Language Learners. *International Journal of Lexicography* 12(1): 55-66.

Nesi, H. 2000. On Screen or in Print? Students' Use of a Learner's Dictionary on CD-ROM and in Book Form. Howarth, P. and R. Herington (Eds.). 2000. *EAP Learning Technologies*: 106-114. Leeds: University Press.

Nesi, H. In press. The Effect of e-Dictionary Font on Vocabulary Retention. *Paper read at eLEX2011: Electronic Lexicography in the 21st Century: New Applications for New Users, Bled, Slovenia.*

Nesi, H. and P. Meara. 1991. How Using Dictionaries Affects Performance in Multiple Choice EFL Tests. *Reading in a Foreign Language* 8(1): 631-643.

Nesi, H. and K.H. Tan. 2011. The Effect of Menus and Signposting on the Speed and Accuracy of Sense Selection. *International Journal of Lexicography* 24(1): 79-96.

Nielsen S. and L. Mourier. 2005. Internet Accounting Dictionaries: Present Solutions and Future Opportunities. *Hermes, Journal of Linguistics* 34: 83-116.

Osaki, S. and N. Nakayama. 2004. Denshijisyo vs. Insatujisyo: Yuyosei to Jikkousei no Chigaini-tuiteno Kosatu [Electronic Dictionary vs. Paper Dictionary: A Comparison of the Two Dictionary Studies]. *Bulletin of Tokyo Denki University, Arts and Sciences* 2: 77-83.

Osaki, S., N. Ochiai, T. Iso and K. Aizawa. 2003. Electronic Dictionary vs. Printed Dictionary: Accessing the Appropriate Meaning, Reading Comprehension and Retention. Murata, M., S. Yamada and Y. Tono (Eds.). 2003. *Proceedings of ASIALEX 2003, Dictionaries and Language Learning: How can Dictionaries Help Human and Machine Learning?*: 205-212. Tokyo: Asialex.

Park, C. 2004. What Is the Value of Replicating Other Studies? *Research Evaluation* 13(3): 189-195.

Prinsloo, D., U. Heid, T. Bothma and G. Faaß. 2011. Interactive, Dynamic Electronic Dictionaries for Text Production. Kosem, I. and K. Kosem (Eds.). 2011. *Electronic Lexicography in the 21st Century: New Applications for New Users. Proceedings of eLex 2011*: 215-220. Ljubljana: Trojina, Institute for Applied Slovene Studies.

Quirk, R. 1974. The Image of the Dictionary. Quirk, R. (Ed.). 1974. *The Linguist and the English Language*: 148-163. London: Edward Arnold.

Rogers, M. 1996. Beyond the Dictionary: The Translator, the L2 Learner and the Computer. Anderman, G. and M. Rogers (Eds.). 1996. *Words, Words, Words. The Translator and the Language Learner*: 69-95. Clevedon: Multilingual Matters.

Santos, T. 1989. Replication in Applied Linguistics Research. *TESOL Quarterly* 23(4): 699-702.

Seidlhofer, B. 2003. *Controversies in Applied Linguistics*. Oxford: Oxford University Press.

Sidman, M. 1960. *Tactics of Scientific Research: Evaluating Experimental Data in Psychology*. New York: Basic Books.

Sinclair, J.M. (Ed.). 2008[6]. *Collins COBUILD Advanced English Dictionary*. Boston: Heinle Cengage Learning / Glasgow: HarperCollins Publishers. *http://www.myCOBUILD.com*. [COBUILD6]

Tono, Y. 2011. Application of Eye-Tracking in EFL Learners' Dictionary Look-Up Process Research. *International Journal of Lexicography*, 24(1): 124-153.

Umapathy, S. 1987. Unfavorable Variances in Budgeting: Analysis and Recommendations. Ferris, K.R. and J.L. Livingstone (Eds.). 1987. *Management Planning and Control*: 163-176. Beavercreek: Century VII.

Wehmeier, S. (Ed.). 2005[7]. *Oxford Advanced Learner's Dictionary of Current English*. Oxford: Oxford University Press. [OALDCE7]

Xu, X. 2010. Study on the Effect of Dictionary Use on Second Language Incidental Vocabulary Acquisition: An Empirical Study of College English Vocabulary Learning Strategy. *Journal of Language Teaching and Research* 1(4): 519-523.