Monika Kavalir

University of Ljubljana
Faculty of Arts, Slovenia

Teaching Reference Skills: Does It Really Matter?

Summary

The article discusses the results of a study at the core of which is the question of whether teaching reference skills really matters – i.e. how useful it is to train students to work with dictionaries. To answer this query, two issues are investigated: firstly, how effective the particular methodology developed here is with regard to using thesauri; and secondly, whether it makes a difference if some specific hands-on exercises are included or excluded. By the end of the project, the students felt that thesauri, a previously unknown language tool, would be quite useful for their work in the future; on average, they were planning to use a thesaurus several times a month. The analysis of the student questionnaires and a comparison of lexical variation in the first and second drafts of the essays also showed that the experimental group, which received more practical training, generally found thesauri more useful and, accordingly, used them more than the control group did.

Key words: reference skills, dictionary use, thesaurus, L2 writing

Raba jezikovnih priročnikov: pomen pedagoškega dela

Povzetek

Članek prinaša prve rezultate raziskave, katere cilj je bil preveriti, ali je usposabljanje za delo s priročniki resnično pomembno oziroma kako koristno je za študente urjenje v delu s slovarji. Osrednji del članka obravnava dve vprašanji: kako učinkovita je v projektu uporabljena metodologija za delo s tezavrom in ali vključevanje oziroma izključevanje posameznih praktičnih vaj bistveno vpliva na rezultate. Ob zaključku raziskave so bili študenti mnenja, da jim bodo tezavri – predhodno nepoznano jezikovno orodje – v prihodnosti precej koristili. V svojih odgovorih so izrazili pričakovanje, da bodo tezaver v povprečju uporabljali nekajkrat na mesec. Analiza študentskih vprašalnikov in primerjava raznolikosti besedišča v izvirni in popravljeni različici spisov sta tudi pokazali, da je eksperimentalna skupina, ki je bila deležna več praktičnega urjenja, tezaver na splošno ocenila za koristnejši pripomoček in ga tudi več uporabljala kot kontrolna skupina.

Ključne besede: delo s priročniki, raba slovarjev, tezavri, pisanje v tujem jeziku

Teaching Reference Skills: Does It Really Matter?

1. Introduction

According to Wright (1998, 5), "[d]ictionaries are among the most readily available, widely used, and cheapest learning resources … [but] they are also among the most difficult to use." While dictionaries may not in fact be that widely used in English language teaching (cf. Fraser 1998), special materials for practising reference skills – the ability of the dictionary user to find or select the information being sought – have been developed when it comes to general (typically advanced learner's) dictionaries. The field has benefited from some scientific research, particularly in connection with reading skills and vocabulary acquisition (e.g. Knight 1994), although "in these matters rather more is written on the basis of 'common sense' and general lexicographical or pedagogical experience than from hard empirical research or within any well worked out second language acquisition or applied linguistic theoretical framework (Scholfield 1997, 279)." It is even more exceptional to find a study focusing on the use of a thesaurus, which is generally very little known, as shown for example by Vrbinc and Vrbinc (2005).

2. Methodology

The study involved 29 first year students at the Department of English and American Studies, University of Ljubljana, divided into two groups. In what is labelled as the experimental group there were 14 individuals, 13 female students and one male student, aged 18 to 23. In the control group there were 15 students, of whom approximately half were male, and half female. Their age varied from 19 to 25. All of the students are native speakers of Slovene, except for 3 students from the first group who gave Bosnian, Croatian and Macedonian as their mother tongues. For the most part, they have had approximately 10 years of English instruction.

The decision to use a thesaurus was motivated by both pedagogical and methodological concerns. First, this is a type of reference book that few advanced students of English know and use although it can at their level be a valuable language tool (cf. Gabrovšek 2006).[1] In the broader context of language teaching, focusing on using dictionaries and other reference materials is part of learner training and as such "one of the most useful things which the teacher can do (O'Dell 1997, 275)." Second, since the use of thesauri is far from widespread, the study should test precisely the effectiveness of the methodology and exercises within this project and not some previously acquired skills.

The dictionary used in this study was *Roget's International Thesaurus* (Chapman 1992) – not what is most commonly understood under the term thesaurus, i.e. a dictionary of synonyms (cf. Kilgarriff and Yallop 2000 for the range of language tools the term can denote), but a

---

[1]   Instruction in using thesauri can help students in some crucial areas of L2 writing, and studies have found it to be less complex, less mature and stylistically appropriate, less consistent and academic with regard to language, style, and tone; fewer synonyms are used, the words L2 writers use are shorter, vaguer, and exhibit less lexical variety and sophistication with less lexical control overall (Silva 1993, 666–8). A number of these characteristics might be related to Laufer's (1991) findings that especially advanced students of English, who can function successfully in their environment, adopt the strategy of 'playing it safe' and use a limited stock of words they are most familiar with.

lexicographical or conceptual thesaurus, a non-alphabetical workbook organised thematically (cf. Gabrovšek 1997). As the students participating in the study are quite advanced, the *Longman Language Activator* (1993), which is the only thesaurus-like monolingual EFL dictionary, was not appropriate to their level since it is intended for intermediate learners, and no bilingual dictionaries of the *Cambridge Word Routes* (1994) type exist for Slovenian; hence, a thesaurus originally aimed at native speakers was chosen for the study.

First, the students wrote an essay on a given topic: *The Influence of the Media on One's Self-Image*. This was part of their coursework (both groups were discussing health problems at the time but eating disorders etc. had not yet been mentioned) and they had no way of knowing it formed part of a larger project.

Then the students filled out a questionnaire the purpose of which was to see what kind of dictionaries they used and how often. In class the students later completed a questionnaire that asked them about thesauri – questions such as *Have you ever heard of a thesaurus? What is it? Have you ever used it?* etc. This was followed by a theoretical presentation of what a thesaurus is and how it is used.[2]

The experimental group then carried out two different hands-on exercises. The first exercise, the purpose of which was to practise the transition from the idea to the word, consisted of each student suggesting a keyword related to the topic of anorexia, looking it up in the index, finding the appropriate section in the thesaurus, and reading out the concepts listed there. The class then decided which of those could be helpful when writing a short composition.

The second exercise was introduced to practise going from the word to the idea, effectively using the thesaurus as a dictionary of synonyms. An extract from a student essay was given; each student had to choose a word he or she believed could be improved and look it up in the thesaurus.[3] The student read out all of the suggestions he or she found familiar, and the class decided whether they would be applicable in that specific context or not. These two practical exercises were not carried out in the control group.

In the session that followed, the students worked in groups of three or four and had to produce short written compositions on two different topics (eating disorders and drugs), one with and the other without the help of a thesaurus (the titles were reversed for half of the class). The various versions were then compared. Finally, the students were asked to go back to their original essays and improve them using any techniques available to them, including of course the ones presented in class. Once they had finished, they also filled out the last questionnaire, which dealt with the second draft of their essay and the revisions made.

An overview of the methodology is given in Table 1. While the procedure may seem straightforward, there are some methodological concerns that must be addressed. A major factor

---

[2]   Contrary to the idea of dictionary use being a simple skill, it actually consists of a number of subskills students have to master, which is especially true in the case of Slovenian students of English using thesauri as there are no L1 skills that could be transferred.

[3]   It was explained to the students that this method was used merely for the purpose of practising reference skills and that consulting a thesaurus would typically be an integral part of the writing process rather than post-writing editing; some theories of L2 writing do, however, suggest dictionaries should be used only at a later drafting stage (cf. Scholfield 1997, 287; Silva 1993, 671).

that might significantly skew the results is that neither the essays nor the questionnaires were anonymous. However, while this may play a role in the overall picture, it should not be relevant to the comparison of the two groups. The study has also provided a wealth of data of which only a small portion has been analysed so far. Finally, the number of subjects involved in the study was relatively small and the results might not be readily generalised. In most cases the results do not have the necessary significance but merely exhibit certain tendencies.

Experimental group	Control group
Essay: first draft	Essay: first draft
Questionnaire 1	Questionnaire 1
Questionnaire 2	Questionnaire 2
Theoretical presentation	Theoretical presentation
Hands-on exercise 1	/
Hands-on exercise 2	/
Group writing	Group writing
Essay: second draft	Essay: second draft
Questionnaire 3	Questionnaire 3

Table 1. Methodology.

3. Part 1: Questionnaires

An important issue to address was whether the students had already had some training in the field. Questionnaire 1 asked the students to give a comprehensive list of all the dictionaries they used regularly, together with the frequencies. A thesaurus was listed once. Questionnaire 2 tested how familiar the students were with the concept of a thesaurus. Three students out of 29 had used a thesaurus before and were able to give at least a partial definition of it. However, only 12 students said the word was new to them while as many as 14 reported they had encountered the word before, mainly on the internet and during the library induction session.

Questionnaire 3 focused on the revised drafts of the students' essays. When the students were asked how useful thesauri were when they were trying to improve their essay, one student in each group said they had not used one. The results for the rest of the students show that the group that had received more training found these reference books much more useful (significant at the 10% level). On a scale of 1 to 5, with 3 being *Useful occasionally*, 4 *Quite useful* and 5 *Very useful*, the experimental group scored 4.31 and the control group scored 3.86 (cf. Table 2).

Experimental group	Control group	t-test
4.31	3.86	1.80<2.06

Table 2. Usefulness of thesauri for revision (1-Not useful; 5-Very useful; $p = .05$).

In order to test some more specific hypotheses, the students then had to mark to what extent they agreed with certain statements. For every issue a continuum was established with two converse statements which had a double function: on the one hand, they served as control sentences for

each other, and on the other hand, their being on the opposite sides of the continuum worked to diminish the influence of conformity (i.e. the tendency of respondents to agree with the statements given).

When asked for a holistic judgment of the quality of the second draft, the students, understandably, believed it was better than the first one, but as Table 3 shows, this conviction was stronger with the experimental group (5.32) than with the control group (5.03). The two statements were *I believe the two versions of my essay are similar in quality* (1) and *I feel that the second draft of my essay is better than the first one* (7).

Experimental group	Control group	t-test
5.32	5.03	0.65<2.06

Table 3. General comparison of drafts (1-Similar in quality; 7-Second draft better; p = .05).

Students also believed that they used more appropriate vocabulary in the new version, but again there was some discrepancy between the two groups: on average, the students in the experimental group scored 5.46 and the students in the control group 5.30 (cf. Table 4). For this question, the scale extended from *I think that I used more words inappropriately in the new version of the essay than in the first one* (1) to *I believe that fewer words and expressions are used incorrectly in the second draft* (7). While might seem counter-intuitive to suggest the possibility of using more words inappropriately in the revised essay, the scale was based on previous practical experience and in fact there were two students who agreed more with the former statement than with the latter. This can happen for two reasons: either the students disobey the instructions to use only expressions they know well (in which case their assessment is correct) or they simply feel insecure as they venture out of their comfort zone (but the results are actually good or even very good).

Experimental group	Control group	t-test
5.46	5.30	0.41<2.08

Table 4. Appropriateness of vocabulary (1-Less appropriate; 7-More appropriate; p = .05).

 As expected, the discrepancy was greater (although still not statistically significant) when the students were asked to judge the lexical richness and diversity of their essays. The sentences were as follows: *Despite the changes I have made, I believe the vocabulary has stayed at approximately the same level* (1); *I think that the second draft of my essay has more varied and more interesting words* (7). Table 5 shows the results for the two groups.

Experimental group	Control group	t-test
5.14	4.63	1.11<2.08

Table 5. Lexical richness (1-The same; 7-Greater; p = .05).

The question about the future use of thesauri yielded some surprises, though, as the students from the control group claimed they would use thesauri more often than the students from the experimental group (cf. Table 6). Without exception everybody plans to use thesauri (i.e. no

student chose 0-Never): on a scale of 1 to 6, the average student scored a little above 4, which means they plan to use a thesaurus several times a month.

Experimental group	Control group	t-test
4.14	4.33	- .43<-2.08

Table 6. Frequency of future use (1-Once a year; 6-A couple of times a week; p = .05).

The picture is again clearer when it comes to the usefulness of thesauri in the future: the vast majority of students, 24 out of 29, believe thesauri will prove to be either quite or very useful. As is shown in Table 7, in the experimental group the average score was 4.29 and in the control group 4.07 (significant at the 10% level).

Experimental group	Control group	t-test
4.29	4.07	1.79<2.08

Table 7. Usefulness of thesauri in the future (1-Not useful; 5-Very useful; p = .05).

The students were also asked to evaluate the work done in class. It is worth mentioning again that the questionnaires were not anonymous and the students may have felt this was an opportunity to score some sycophantic points with the teacher; still, this applied to both groups in the same measure – while the absolute values might be suspect, the difference between the two groups should not reflect this in any way. The maximum value is 5, which means the results in both groups are very high, but as is obvious from Table 8, the two practical, hands-on exercises really made an important difference with the experimental group scoring 4.71, compared to 4.07 for the control group. This is also the only variable where the difference between the two groups is statistically significant.

Experimental group	Control group	t-test
4.71	4.07	2.43>2.08

Table 8. Usefulness of training (1-Not useful; 5-Very useful; p = .05).

4. Part 2: Essays

The judgments analysed are extremely subjective, and the students might simply be wrong, which is why it is beneficial to complement the findings with some preliminary results for the essays obtained with the help of the *WordSmith* suite of tools.

Lexical richness of texts can be described in terms of lexical variation, defined as the type/token ratio, i.e. the ratio in percent between the number of different lexemes in the text and the total number of lexemes. When the first and second drafts were compared, the type/token ratio was markedly different. As can be seen from Table 9, the ratio was initially the same for the two groups and then increased in both cases, but it went up twice as much in the experimental group, from 22 to 24, whereas it grew by 1 in the control group.

	Experimental group	Control group
First draft	22	22
Second draft	24	23

Table 9. Type/token ratio.

It is interesting to see how this change came about and have a look at the frequencies of certain word types in Tables 10 and 11. It is characteristic of both groups that the frequency of relatively short words was slightly lower in the second draft than in the first draft (the numbers listed are absolute numbers and cannot be directly compared because the number of subjects in the two groups is not the same).

	Experimental group	Control group
3-letter words, Essay 1	1230	1409
3-letter words, Essay 2	1218	1408
4-letter words, Essay 1	1148	1255
4-letter words, Essay 2	1131	1223
5-letter words, Essay 1	745	852
5-letter words, Essay 2	725	848
6-letter words, Essay 1	514	571
6-letter words, Essay 2	511	552

Table 10. Frequency of short words.

The type/token ratio increased mainly because more long words were used in the revised essays. 7, 8 and 9-letter words are transitional categories, and the frequencies of longer words are consistently higher in the second draft, with the differences being generally more pronounced in the experimental group (cf. Table 11).

	Experimental group	Control group
10-letter words, Essay 1	173	189
10-letter words, Essay 2	213	204
11-letter words, Essay 1	99	77
11-letter words, Essay 2	116	94
12-letter words, Essay 1	41	33
12-letter words, Essay 2	53	41
13+, Essay 1	31	42
13+, Essay 2	35	43

Table 11. Frequency of long words.

This is of course no coincidence and has to do with the use of "more varied and interesting words". What seems to have typically happened is that common, simple, informal, short Anglo-Saxon words were replaced with relatively less frequent, more complex, more formal and longer words, for example of Romance origin. It is important to note here that the students were instructed to only use the thesaurus to remind them of words they already knew, i.e. to activate their passive knowledge. A glance at the word frequency lists reveals some of the lexemes that were used in the second draft for the first time: *achieve, appealing, criteria, frequently, furthermore, impeccable, manner, mental, perfection, permit, stunning, unreachable, accordingly, acquainted, acquire, affect, alternatively, ameliorate, assumption, attain, beyond, bump, contentment* etc.

Two types of further research would be beneficial in providing a more detailed picture of the changes in the essays. On the one hand, more measures of lexical richness could be included, e.g. lexical density, lexical originality, lexical sophistication (cf. Laufer 1991). On the other hand, a qualitative analysis of the essays would yield more information about the appropriateness and success of learners when trying to improve their essays.

Even if the quantitative analysis of the essays seems to confirm the students' estimation and the second drafts were indeed lexically more developed, other factors apart from using thesauri may have contributed to this. To an extent, such an improvement is, of course, to be expected in the process of revision. There are, however, two arguments which support the idea that the use of thesauri was a major influence here.

First, the students did not know they would get to write a second draft; the first draft was to be marked, which is why they would presumably do their best (including revision) work already the first time. They did not at any point receive any feedback or guidance on their essay specifically, which is commonly the case in process writing.

Second, the notable difference between the two groups can arguably be attributed to the two exercises the experimental group did but the control group did not do. If the impact of these two exercises was so great, it would be logically inconsistent to assume all of the other exercises were irrelevant. While there is no straightforward way of determining the exact contribution of dictionary use, it seems likely that thesauri played an important role in the revision.

5. Discussion

The cumulative results (cf. Table 12), although rarely statistically significant at the required level (which is understandable given the nature of the study and the size of the sample), reveal a clear pattern. At the beginning of the study, 26 students had never used a thesaurus, 3 had. By the end of the study, a vast majority found it a useful tool to be consulted quite frequently – and the students who had used a thesaurus before also found more use for it later on in the study. A primary concern in the research conducted here was to test how important practical work on reference skills was and this is the area where interesting differences between the two groups can be observed.

	Experimental group	Control group
Usefulness of thesauri for revision	4.31	3.86
General comparison of drafts	5.32	5.03
Appropriateness of vocabulary	5.46	5.30
Lexical richness	5.14	4.63
Frequency of use in the future	4.14	4.30
Usefulness of thesauri in the future	4.29	4.07
Usefulness of training	4.71	4.07
Type/token ratio	22>24	22>23

Table 12. Overview of results.

Both groups were given a theoretical explanation of what a thesaurus is and how it is used, and students in both groups were handed out thesauri to use for written production in class. This means that all the students were given the opportunity to gain both theoretical and practical familiarity with this particular type of dictionary. The only difference was that the experimental group also did two hands-on exercises where each individual student had to take the thesaurus in his or her hands and use reference skills as demonstrated to provide the class with some information.

A consequence of this drill, which many university teachers would describe as an unintelligent mechanical skill best practised at home, was that the experimental group found thesauri more useful when trying to improve their essays (4.31 vs. 3.86 on a scale of 1 to 5). The students in the experimental group felt more strongly that the revised versions were generally better than the first drafts (5.32 vs. 5.03), that the vocabulary used was more appropriate (5.46 vs. 5.30), and that they used more varied and interesting words in the improved essay (5.14 vs. 4.63, all on a scale of 1 to 7).

As might be expected given these results, the students in the experimental group believed thesauri would prove to be more useful in the future than the students in the control group (4.29 vs. 4.07), but surprisingly it was the latter who felt they would use thesauri in the future more frequently (4.14 vs. 4.30, both on a scale of 1 to 5). This might be explained by the possibility that the control group felt more could be done with the thesaurus than they knew. When asked whether they felt they needed further training in using dictionaries, 5 students in the control group gave a positive answer and 4 of them were open to further practice on using thesauri while of the 6 students in the experimental group who answered in the affirmative, only 1 would welcome more training in this specific field and all of the others would appreciate work on other kinds of specialised dictionaries. This suggestion is, however, purely speculative, especially given the number of respondents is small and the difference between the two groups slim.

The disparateness of the two groups is greatest when it comes to the evaluation of the training the students received in class (4.71 vs. 4.07, on a scale of 1 to 5). It is clear from these data that the students appreciated the opportunity to practically apply the gained theoretical knowledge.

The change in the type/token ratio shows that these results do not stem from a subconscious conviction that more time and effort invested automatically equals more knowledge. As would be expected of two comparable groups, the values for the first draft were the same, but the final versions showed a greater increase in the experimental group. While it may be assumed that the change was for the better, this can by no means be verified without a qualitative analysis of the essays.

6. Conclusion

The presented study has shown that practical exercises are indeed important when developing reference skills. While a larger sample enabling more sophisticated statistical analysis and a qualitative appraisal of the essays might shed more light on the issues discussed, the basic quantitative analysis of some answers the students provided in the questionnaires and an initial assessment of lexical variation in the first and second drafts of the essays make it obvious that students not only appreciate such hands-on training but that it yields palpable results. The results of the study – while limited in their applicability by the small number of subjects involved – thus suggest that this kind of methodology is an effective way of turning students into skilful users of dictionaries and that training really does matter.

References

Cambridge Word Routes Anglais-Français. 1994. Cambridge: Cambridge UP.

Chapman, R.L., ed. 1992. *Roget's International Thesaurus*. New York: Harper Collins.

Fraser, C.A. 1998. The Role of Consulting a Dictionary in Reading and Vocabulary Learning. *Canadian Journal of Applied Linguistics* 2, no. 1: 73–89. *http://www.aclacaal.org/Revue/vol-2-nos1-2-art-fraser.pdf* (accessed 16 October 2009).

Gabrovšek, D. 1997. Vocabulary Reference: The Non-Alphabetical Mode in Action. *Vestnik* 31, no. 1–2: 187–208.

---. 2006. In a Phrase, the Non-Alphabetical Wordbooks. *Mostovi* 40, no. 1–2: 144–67.

Kilgarriff, A., and C. Yallop. 2000. What's in a Thesaurus? In *Proceedings of the Second International Conference on Language Resources and Evaluation*. Greece, Athens, 1371–9. *http://www.kilgarriff.co.uk/Publications/2000-KilgYallop-LREC-thesauruses.pdf* (accessed 23 October 2009).

Knight, S. 1994. Dictionary Use While Reading: The Effects on Comprehension and Vocabulary Acquisition for Students of Different Verbal Abilities. *Modern Language Journal* 78, no. 3: 285–99.

Laufer, B. 1991. The Development of L2 Lexis in the Expression of the Advanced Learner. *Modern Language Journal* 75, no. 4: 440–8.

Longman Language Activator. 1993. London: Longman.

O'Dell, F. 1997. Incorporating Vocabulary into the Syllabus. In *Vocabulary: Description, Acquisition and Pedagogy*, ed. N. Schmitt and M. McCarthy, 258–78. Cambridge: Cambridge UP.

Scholfield, P. 1997. Vocabulary Reference Works in Foreign Language Learning. In *Vocabulary: Description, Acquisition and Pedagogy*, ed. N. Schmitt and M. McCarthy, 279–302. Cambridge: Cambridge UP.

Silva, T. 1993. Toward an Understanding of the Distinct Nature of L2 Writing: The ESL Research and Its Implications. *TESOL Quarterly* 27, no. 4: 657–77.

Vrbinc, M., and A. Vrbinc. 2005. Future Professional Dictionary Users and Their Use of Dictionaries. *Acta Neophilologica* 39, no. 1–2: 139–52.

Wright, J. 1998. *Dictionaries*. Oxford: Oxford UP.