

ASSESSMENT ON RESEARCH METHODS OF INCIDENTAL VOCABULARY ACQUISITION

Wang Dongzhi

University of International Business and Economics

Abstract

This paper purports to identify some shared methodological problems by analyzing and evaluating the research methods used in 10 empirical studies about incidental vocabulary acquisition . It also points out that the main reason for these problems is due to the complexity of second language acquisition , which consequently leads to more confounding and uncontrollable factors .

Key words

incidental vocabulary acquisition ; empirical study ; research method

1 . Introduction

In the past dozen years or so , second language vocabulary acquisition has become an increasingly interesting topic of discussion for researchers , teachers , curriculum designers , theorists , and others involved in second language learning . Except for the first few thousand most common words , vocabulary learning predominantly occurs incidentally (Huckin & Coady 1999) . The concept of incidental vocabulary acquisition (IVA) was first proposed by Nagy , Herman & Anderson (1985) and later was developed by Nation (1990) , Laufer (1998 , 2001) and Hulstijn (2001) . Contrary to intentional learning featured by rote memorization and repetitious practice , incidental acquisition means that words are learnt while doing other tasks like reading , listening or writing . Since its inception , some important relevant studies (Pigada & Schmitt 2006 ; Krashen 1989 ; Laufer 1998) have been conducted in the western world . However , IVA study in China started about 10 years later and the number of articles on this topic is relatively small .

The purpose of this paper is not to research IVA , but to examine the research methods used in 10 empirical studies in this field , since empirical study is very fashionable in both natural and social sciences nowadays . It is generally agreed that the methodology employed in such kind of study is so important that it can even determinate validity of the research results . Wen (2001 : 18) addresses empirical research as primary research because its data are derived from the primary source . It includes such methods as introspection , experimentation , case study , solicitation , etc . In this paper , all the cited 10 studies are experimentations , selected from well-established journals published in China within the past 5 years . Centering on the same topic incidental vocabulary acquisition , they can be further sub divided into smaller groups : The first 7 papers deal with vocabulary acquisition through reading , of which Shao (2006) focuses on word frequency's impact on IVA and Duan & Yan (2004) discusses the roles of different ways of glossing in vocabulary learning . The remaining 3 have different focuses : Lei , Wei , Ye & Zhang's paper (2007) studies incidental vocabulary acquisition through writing ; Wang Lei (2006) investigates word learning during listening class ; and Wang (2005) examines the role of movie subtitling in the process of vocabulary learning . Although these studies have different focuses , they adopt a similar pattern of research methodology , such as vocabulary testing , reading comprehension task and statistical analysis . However , due to various reasons , there exist some common but differentiated methodological

loopholes in these studies ;some are remediable and others are not . This paper is trying to find them out in different categories ,so that future researchers will be aware of them and know how to improve their research design .

2 . Basic Information of the Empirical Studies

Ge (2004) examined 18 empirical studies on incidental vocabulary acquisition published in internationally renowned journals , pointing out six types of problems concerned with research methodology ,such as participants ,testing methods ,confounding factors ,statistical errors ,etc . In the similar vein ,this paper makes an effort to examine 10 domestic IVA empirical studies published in the well established journals in China within past 5 years ,addressing the methodological problems by category as well . Here I would like to use the following table to show some key information about these studies .

Table 1 . Basic information about the ten empirical studies on incidental vocabulary acquisition

No .	Author	Research Questions	Variables	Participants	Experimental Design
1	WANG Wen - wei	1 . For the intermediate - level Chinese learners of English ,can extensive reading lead to IVA ? 2 . If it can ,to what extent and at which aspect does it work ?	Independent : extensive reading Dependent :IVA Moderator :language proficiency	2 natural classes of 60 freshmen with the same major	Quasi -experiment with pretest -posttest non equivalent groups
2	XU Hong	1 . How many words can be learnt incidentally ? 2 . Is word frequency related with IVA ? And how ? 3 . Is vocabulary size related to IVA ?	Independent :free reading Dependent :IVA Independent :word frequency Dependent :IVA Independent : vocabulary size Dependent :IVA	2 classes of 55 sophomores with accounting major in a college	Pre -experiment with one group pretest - posttest study
3	GE Shu hua	1 . Is it possible to learn words incidentally in reading ? 2 . Do different reading tasks affect IVA ? And how ? 3 . Do individual factors affect IVA ?	Independent :reading tasks Dependent :IVA Independent : vocabulary size and learning strategies Dependent :IVA	2 parallel classes of 57 sophomores with English major	Quasi -experiment with posttest only nonequivalent groups + questionnaire + quiz for vocabulary size
4	SHAO Yan - chun	1 . Do different reading purposes affect IVA ? 2 . Do different reading purposes have different impact on learners with different reading abilities ?	Independent :reading purpose Dependent :IVA Moderator :reading ability	2 parallel classes of 89 English learners with non -English major	Pre -experiment with one group posttest only study

No	Author	Research Questions	Variables	Participants	Experimental Design
5	DONG Yan - ping	Is direct learning helpful apart from indirect learning in the com municative approach ?	Independent : vocabulary teaching method Dependent : vocabulary acquisition	3 parallel groups from Grade 1 and another 3 from Grade 2 (all majoring in English)	Pre experiment with one group posttest only study
6	DUAN Shi ping & YAN Chen - song	1 . In im mediate vocabulary test , which way of glossing benefits IVA most , single , multiple or none ? 2 . In delayed test how do the 3 ways of glossing differ in word retention ? 3 . Does vocabulary base affect IVA ?	Independent : way of glossing Dependent : IVA and vocabulary retention Independent : vocabulary size Dependent : IVA	3 parallel classes of 89 second year English majors in Henan Institute of Technology	Quasi experiment with posttest only nonequivalent groups + quiz for vocabulary size
7	ZHOU Ling	1 . Is IVA the main word acquisition method for Chinese second year non -English majors ? 2 . Which method is more effective for them , intentional or incidental learning ? 3 . Which method can retain vocabulary longer ?	Independent : way of word acquisition Dependent : vocabulary test result	2 classes of 60 second year non -English majors from School of Material Science and Engineering of Hunan University	Quasi experiment with posttest only nonequivalent groups
8	LEI Lei , WEI Yao yu , YE Lin , ZHANG Mei	1 . Can the two writing tasks help IVA ? 2 . In im mediate vocabulary testing , which one can be most conducive to IVA , two writing tasks or the reading task ? 3 . In delayed testing , how do the 3 tasks affect vocabulary retention differently ?	Independent : reading and writing tasks Dependent : IVA and vocabulary retention	3 parallel natural classes of second year undergraduates of Huazhong Institute of Technology	Pre experiment with one group posttest only study + interview

No	Author	Research Questions	Variables	Participants	Experimental Design
9	WANG Lei	<p>Hypotheses :</p> <p>1 . Having listened to a dialogue ,participants can use vocabulary chunks better and this will be reflected in their writings .</p> <p>2 . Participants can realize such kind of acquisition .</p>	<p>Independent : Listening input</p> <p>Dependent : vocabulary acquisition</p>	2 classes of 50 second year English majors	Pre experiment with one group posttest only study (no delayed test) + 3 open questions
10	WANG Hui	<p>1 . How does subtitle affect IVA ?</p> <p>2 . Do Chinese and English subtitles affect IVA differently ?if yes , how do they differ from each other ? What are the causes ?</p> <p>3 . How do English proficiency affect subtitle s role ?</p>	<p>Independent ;subtitle</p> <p>Dependent :IVA</p> <p>Controlling :Chinese subtitle ,English subtitle ,English proficiency</p>	<p>Low level group :3 natural classes of 92 first year English majors ;</p> <p>high level group :3 natural classes of 90 third - year English majors</p>	<p>Quasi experiment with posttest only nonequivalent groups (no delayed test)</p>

There is one thing which needs to be clarified before discussion ,that is ,how to define the three types of experimental design listed in the above table ,namely ,pre experiment ,quasi experiment and experiment . According to Larsen Freeman & Long (1991 :19 24) ,their fundamental difference can be illustrated as below :

Table 2 . Differences of the three types of experimental design

	Pre experiment	Quasi experiment	Experiment
Control group	-	+	+
Random sampling	-	-	+

This tells us that in a real experiment two criteria must be satisfied :(1) there must be experimental and control groups ,that is to say ,the two groups receive different treatments ,and (2) participants must be randomly assigned to each of these groups . The first criterion can be met easily ,but not the second one ,since often a large number of participants are engaged and random assignment usually interferes regular school arrangement . That is why there is no real sense experiment in these studies . Quasi-experimentation is a more common practice in empirical research .

3 . Some Identified Methodological Problems

In the following section ,factors concerning research methodology are going to be discussed in detail . But it is worthy to note that research results of the 10 studies are beyond our discussion . After a careful assessment of these studies on a same topic ,some methodological problems have been identified ,which fall into 3 categories ,that is ,misuse of key concepts ,loose control of intervening variables ,and weak experimental validity . Firstly ,although published in the major nation-wide journals ,some concepts were still misused or used inappropriately ;as a result the reader might be confused . Secondly ,for experiment ,how to control intervening variables is a critical issue . According to Wen (2001) ,this can be done either artificially by random sampling or statistically by participants stratification . But out of the ten studies only one used random sampling . The last but most important category is internal validity ,which is related to

participants , materials , methodology and experimenter . Wen explains internal validity in the following way :

The internal validity concerns about the question “ To what extent can the claimed cause-effect relation be justified ?” In an experiment , if the researcher succeeds in constructing an environment where only the independent and dependent variable(s) are working while the other variables are controlled , the changes in the dependent variable can thus be attributable to the treatment only . Such an experimental design shows high internal validity . (Wen 2001 :150-151)

In other words , the experiment would lose its foundation if the internal validity were challenged . The following discussion of the 10 cases is just based on the 3 categories specified above .

3.1 Misuse of concept

Concepts , or technical terms , are usually regarded very critical in academic writing ; if used improperly , the reader may feel confused , or the overall research design might be impaired . In Wang Wenwei (2006) , the word “ pretest ” was used to refer to selecting participants by a test before the experiment , but Wang used different contents in the following two so-called posttests . Strictly speaking , this should not be called “ pretest ” , because in common practice , both pretest and posttest must use the same contents so as to compare the results and reach conclusion . In Duan & Yan (2004) , the authors reckoned multiple glossing as a type of “ incidental vocabulary acquisition ” . But according to the definition given by Laufer (1998) , glossing should belong to intentional learning rather than incidental acquisition , just like learning word by consulting a dictionary , which involves much motivated and deep processing activities . The most serious case of misconception is found in Wang Lei's paper . Before the experiment Wang proposed two “ hypotheses ” about vocabulary chunks listening and production and incidental acquisition awareness , but according to Wen (2001) , hypothesis is defined as a declarative statement that describes the hypothetical relationship between two or more variables . In this sense they are not hypotheses , since no variable-bound relations are assumed ; in my view it would be more proper if these two hypotheses were changed into research questions .

3.2 Loose control of intervening variables

Loose control of intervening variables will make research results susceptible for various interpretations , thus threatening both internal and external validity of experiment . Therefore , often the intervening variables are controlled either by deliberately minimizing their impact upon experimentation or by using statistical methods . The former is called artificial control and the latter statistical control , represented by random sampling and participants stratification respectively . However , all studies under discussion except Shao (2006) used natural class in experiments instead of regrouping the participants . This is very risky because the personality , habit , motive , English proficiency , learning strategy and effort of participants of the same class may differ greatly , consequently the validity of results will not be guaranteed . However , it doesn't mean that we can not get valid results by artificial or statistical control . Even the case in Shao (2006) , which studied the relationship between different ways of glossing and incidental vocabulary acquisition and concluded that multiple glossing was the best of the three for word learning and retention , also neglected such a fact : Participants usually spend more effort on multiple-glossing than on the other two , so it's very hard to say whether the effort or the way of glossing contributes to vocabulary learning and word retention . As a result , the conclusion loses its grounds .

3.3 Weak experimental design

Experimental design can be of paramount importance to the internal validity of experiments . In contrast to external validity , which refers to the extent to which the results can be generalized from samples to populations , internal validity refers to the interpretability of research ; in other words , can any differences which are found actually be ascribed to the treatments . Broadly speaking , the factors affecting internal validity are various and could be related with the environment , participants , methodology and even experimental conditions . But , due to the limited number of cases under discussion , this paper just focuses on factors like participant , material , methodology and experimenter .

3.3.1 Selection of participants

In the first study Wang tried to choose the intermediate level English learners as the target group, but she failed to clearly define this level: Does it refer to their English proficiency or just years of English learning? Have they taken any proficiency test? Without answering these questions, her first research question would be meaningless. And this loophole makes the conclusion very shaky. Xu (2006) and Zhou (2006) share a common problem, that is, the participants were underrepresented in the research. Both authors used non-English majors as subjects, whose English proficiency might be low and vary greatly. In order to avoid making a big claim, it would be better if the authors could narrow down the topic by specifying the property of the participants, for instance, "The Empirical Research of Intentional Vocabulary Acquisition by Chinese Non-English Majors". Shao (2006) intended to study the impact of reading purpose on incidental vocabulary acquisition; however, the author divided the participants into two groups: one with strong reading ability and the other with weak ability. As a result, reading ability unexpectedly becomes an intervening variable, which can also have a direct impact upon vocabulary acquisition. Therefore, it would be too rash to conclude that word learning is solely caused by different reading purposes. In this case, reading ability is just the intervening variable that should be under control. Wang (2005) dealt with the role of different ways of subtitling in incidental vocabulary acquisition, however, participants' individual habits and needs were not investigated beforehand when they should be, because some might be good at splitting their attention when listening to the voice and looking at the subtitle at the same time, while others might not. So, under such circumstances, stratification of participants is of paramount importance for research validity in this study. It is true that usually it is not easy to disturb the normal class arrangement and reorganize participants, but sometimes we have to do it to meet the research purpose, or we may consider other alternatives like statistical control so as to guarantee the research results. In whichever cases, selection of participants should not be underestimated.

3.3.2 Materials for experiment

Material as an experimental instrument is an integral part of empirical study. Generally speaking, it is always not easy to find the suitable material for experiment, so it often costs the researcher great efforts to select, test and revise so as to hammer out an effective tool for research purpose. On the contrary, a not well-chosen experimental material could be detrimental. Dong (2001) studied the relationship between different ways of vocabulary teaching methods and incidental vocabulary acquisition. But there were no materials specially designed for the experiment, only the textbook regularly used by participants in class. As we know, material is part of experiment and should be well chosen so as to meet the research purpose. In this sense, any study without specially designed material would potentially invalidate the experiment results. Even worse, Lei et al. (2007) failed to mention what the experimental material was about and how it was selected, only stating that participants were asked to read a 600-word passage. Such an undue attention to material may well make the conclusion questionable. Duan & Yan (2004) tried to study the role of multiple glossing in word learning, so the authors picked up an article from an English newspaper in Taiwan as experimental material. However, the problem is that the language might be relatively simple and not as native as that in the local newspapers like Times or Economist, and thus some participants may possibly guess out the meaning of target words according to the context. Consequently the result of vocabulary testing would be impaired. Additionally, there is one more point which is worth mentioning: To meet the requirements of a well-developed study, it is necessary to test and examine the experimental material objectively for its appropriateness. In other words, it is not solely determined by the experimenter himself, but more importantly by others as well. Wang (2005) just has this problem. The material was chosen just because the author assumed it was appropriate, without testimony of other people nor pilot study, so its suitability is in doubt and as a result the conclusion does not seem convincing enough.

3.3.3 Methodological issues

Most of the studies under discussion have some degree of methodological problems due to confounding factors in experimentation. It's impossible to control everything in the process, but something can be done to improve it. To sum up, there are two major problems involved: one is related

to measurement design and the other to measuring methods. For measurement design, first problem is limited number of target words, for instance, both Ge (2003) and Duan & Yan (2004) only used 11 words for testing. Generally speaking, the sample size can, to some extent, determine reliability of the experiment results. If the size is too small, there would be more probability for statistical errors. Second is incomplete testing. For example, Wang Lei (2006) and Wang (2005) intended to explore the relationship between audio or visual input and incidental vocabulary learning, but neither of them had any delayed tests, so it is very hard for us to know whether the newly learnt words can be retained or not, since vocabulary retention is an integral part of vocabulary learning. The third problem is no testing for the depth of vocabulary learning, as evidenced in Shao (2006), which simply employed multiple choice as a measuring instrument to test whether the words were learnt. The evidence collected in this way is not sufficient enough to prove that the participants have mastered the knowledge of these words. As for the problem with measuring methods or instruments, Duan & Yan (2004) used Wesche & Paribakht's Vocabulary Knowledge Scale for testing, but this instrument is so complicated that participants may lose their patience and take the test not seriously, thus affecting the research result. In Zhou (2006), in order to compare intentional learning with incidental learning, the author asked the control group to memorize 20 words and the experimental group to read a short passage without reminding them to notice the 20 new words. However, the performance of the control group is predictably better than the other, since the memorizing group has a clearer goal and concentrated efforts to achieve it. So it can be said that the tasks are not good means to an end. Lei et al. (2007) intended to study vocabulary acquisition in the writing process by comparing it with reading task. But such kind of comparison does not seem to make sense because reading and writing are two different processes, involving different factors, just like the gap between learning and practicing; the former is passive while the latter is active. In my view it would be more sensible to find out the discrepancy of acquisition percentage when participants perform the reading and writing tasks. Finally, Wang Lei (2006) assumed that listening would contribute to vocabulary acquisition and this could be reflected in writing. But she forgot one thing: After listening to the same clip for 3 times, it is natural for participants to compose a relevant passage using the chunks they have just heard because they are still active and fresh in memory. In this case, it would be better if the writing task was delayed a certain period of time after the listening task.

3.3.4 Experimenter

Experimenter or experiment implementer is usually the last factor which attracts our attention; however, it would play a very critical role if more than one experimenter is involved in an experiment. The experiment in Dong (2001) was administered in real classroom scenarios and there were different teachers (experimenters) for different classes. But Dong did not mention their characteristics: Do they differ greatly in educational background, teaching ability, experience, English proficiency or age? The information about experimenters should be seriously considered, since they may exert much influence on the research results directly.

3.3.5 Other Problems

There is one more problem worth mentioning, which is not covered by the previous categories, but also very important. In Wang Lei (2006), after the writing task Wang carried out an interview, using 3 open-ended questions to solicit opinions on word learning from the participants. And the author gave a detailed description of how it was done. It seems that the interview is one of the key elicitation instruments employed in the experimentation. However, the problem is that two of them are yes-or-no questions, which are not worthwhile answering and do not contribute to the research itself. Apart from that, the question "Do you think you can remember these words and sentences after a certain period of time?" is quite subjective and presumptive, just like a wild guess. In this sense it can be said that the questions are not well structured.

4. Conclusion

Despite of all the problems mentioned above, this paper is not to devaluate the 10 studies contribution to the incidental vocabulary acquisition research, but to report some methodological problems for attention and put them into categories. The facts have shown that empirical study in applied linguistics in China is still nascent and our overall research capacity in this regard needs to be improved.

On the other hand, research methodology is very tricky, so limitations can be found more often than not. As for all the problems listed in this paper, some can be avoided, such as misuse of concept, number of target words, testing materials, interview structure, and so on. But others may not be solved easily, for instance, it is very hard to randomly sample or regroup participants of natural classes, otherwise the regular class arrangement would be disturbed. And also, it is not very much possible to exclude all intervening variables in an experiment, like participants' attitude, learning strategy, motive and effort. This also shows that incidental vocabulary acquisition is a highly complicated issue, which needs our more careful and thoughtful exploration.

References

- Corder, S. P. 1967. The significance of learners' errors. *International Review of Applied Linguistics* 5 :161-9.
- Craik, F. I. M. & R. S. Lockhart. 1972. Depth of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior* 11 /6 :671-684.
- Huckin, T. & J. Coady. 1999. Incidental vocabulary acquisition in a second language. *Studies in Second Language Acquisition* 21 /2 :181-193.
- Krashen, S. 1989. We acquire vocabulary and spelling by reading: additional evidence for the input hypothesis. *The Modern Language Journal* 73 /4 :440-464.
- Larsen-Freeman, D. & M. H. Long. 1991. *An Introduction to Second Language Acquisition Research*. London and New York: Longman.
- Laufer, B. 1997. What's in a word that makes hard or easy? Intralexical factors affecting the difficulty of vocabulary acquisition. In N. Schmitt and M. McCarthy (Eds.), *Vocabulary Description, Acquisition and Pedagogy*. Cambridge: Cambridge University Press.
- Laufer, B. 1998. The development of passive and active vocabulary in second language: Same or different? *Applied Linguistics* 19 /2 :255-271.
- Laufer, B. & J. Hulstijn. 2001. Incidental vocabulary acquisition in a second language: The construct of task-induced involvement load. *Applied Linguistics* 22 /1 :1-26.
- Lewis, B. 1993. *The Lexical Approach*. London: Language Teaching Publications.
- McDonough, J. & S. McDonough. 1997/2000. *Research Methods for English Language Teachers*. Beijing: Foreign Language Teaching and Research Press & Edward Arnold.
- Nagy, W. E., P. A. Herman & R. C. Anderson. 1985. Learning words from context. *Reading Research Quarterly* 20 /2 :233-253.
- Nation, P. 1990. *Teaching and Learning Vocabulary*. New York: Newbury House Publishers.
- Nunan, D. 1992. *Research Methods in Language Learning*. London: Cambridge University Press.
- Pigada, M. & N. Schmitt. 2006. Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language* 18 /4 :1-34.
- Selinker, L. 1972. Interlanguage. *International Review of Applied Linguistics* 10 :209-31.
- Wen, Qiufang. 2001. *Applied Linguistics: Research Methods and Thesis Writing*. Beijing: Foreign Language Teaching and Research Press.
- Wesche, M. & T. Paribakht. 1996. Assessing vocabulary knowledge depth vs. breadth. *Canadian Modern Language Review* 53 /1 :13-40.
- 董燕萍 2001 交际法教学中词汇的直接学习与间接学习《外语教学与研究》第3期。
- 段士平、严辰松 2004 多项选择注释对英语词汇附带习得的作用《外语教学与研究》第3期。
- 盖淑华 2003 英语专业学生词汇附带习得实证研究《外语教学与研究》第4期。
- 盖淑华 2004 词汇附带习得实证研究的质量评估《外国语言文学》第2期。
- 雷蕾、韦瑶瑜、叶琳、张梅 2007 非英语专业大学生通过写作附带习得词汇研究《解放军外国语学院学报》第1期。
- 邵艳春 2006 不同阅读目的对词汇附带习得的影响——一项基于SPSS统计分析的调查研究《外语电化教学》第109期。
- 汪徽 2005 字幕对伴随性词汇习得的影响《外语电化教学》第102期。
- 王蕾 2006 在听力教学中注意提高词汇意外习得《天津外国语学院学报》第5期。
- 王文伟 2006 泛读与词汇附带习得《中美英语教学》第11期。
- 许洪 2006 词频对英语词汇附带习得的影响《长春师范学院学报(人文社会科学版)》第6期。
- 周玲 2006 关于词汇刻意学习的实证研究《湖北教育学院学报》第4期。