Testing the effectiveness of computer assisted language learning website: WordChamp

Hui Ping Chan

MA in Applied Linguistics
Faculty of Liberal Arts
University of Groningen

Supervisor:
Wander Lowie

June 30th, 2010
Table of Contents

Chapter 1 Abstract .......................................................... 1
Chapter 2 Introduction .................................................... 2
Chapter 3 Background ..................................................... 4
  3.1 Vocabulary learning .................................................. 4
  3.2 Word knowledge ...................................................... 5
  3.3 Vocabulary acquisition process ................................... 8
  3.4 Approaches to vocabulary learning ................................ 9
  3.5 Computer assisted language learning .............................. 11
  3.6 Introduction of WordChamp ......................................... 13
  3.7 Receptive and productive vocabulary test ......................... 20
  3.8 Dynamic System Theory (DST) ..................................... 22
  3.9 Logistic model .......................................................... 22
  3.10 Research questions and predictions ............................... 23
Chapter 4 Methodology: data collection .............................. 25
Chapter 5 Results ............................................................. 29
  5.1 Data description ..................................................... 29
  5.2 Data exploration ..................................................... 30
  5.3 Model specification .................................................. 38
  5.4 Model fitting .......................................................... 39
Chapter 6 Discussion ....................................................... 43
Chapter 7 Conclusion ......................................................... 46
Bibliography .................................................................... 49
Chapter 1

Abstract

Vocabulary learning plays an important role in language acquisition, since learners need a certain coverage of vocabulary to operate a language. WordChamp is a vocabulary learning website offering receptive and productive vocabulary knowledge and incidental and intentional vocabulary learning. Our study tested the effectiveness of WordChamp on the process of vocabulary development, based on three levels of vocabulary knowledge from low to high: active recognition, active recall and controlled production. We analyzed the trend line of each level and the moving correlation of the detrended data between three levels and found the computer learners indeed fill in the gap between receptive and productive knowledge. Besides, the lower vocabulary knowledge supports the higher vocabulary knowledge more for the computer learners than for the paper learner. Next, we used the logistic model to interpret the collected data with initial value, learning rate and learning capacity. The results showed no significant learning difference between the computer learners and the paper learners in the lowest vocabulary knowledge (active recognition). One computer learner showed a higher learning rate in the higher vocabulary knowledge levels (active recall and controlled production), however. We concluded that the effectiveness of WordChamp was feasible but it really relied a lot on learners’ capability, which determined the initial value and the learning rate.
Chapter 2

Introduction

Researchers have widely studied vocabulary learning from receptive to productive, and from incidental to intentional. Second language learners are mostly used to learning vocabulary on paper either by memorizing the translation word lists (intentional learning) or by deducing the word meaning from the context (incidental learning), but each approach has its shortcoming. Only memorizing the translation does not enable learners to use the words correctly in speech or in writing, which is the gap between receptive and productive vocabulary knowledge. That is to say, the inability to produce the words is very likely to happen even when the ability to recognize the words has been acquired. On the other hand, only deducing the word meaning does not make the words consolidated in the brain from short term memory to long term memory. The computer assisted vocabulary learning website: WordChamp, however, combines intentional and incidental learning and also equips learners with receptive and productive vocabulary knowledge. The aim of the study is to test the effectiveness of this promising vocabulary learning website theoretically and empirically to determine whether computer is able to assist vocabulary learning: whether the gap between receptive and productive vocabulary knowledge can be filled by learning with WordChamp, how each vocabulary level interacts with each other, and how the vocabulary development of the computer learners differs from that of the paper learners.

To test the effectiveness theoretically, we linked the functions of WordChamp with vocabulary literature to check whether it is compatible with theories. The findings are significant in word knowledge, vocabulary acquisition process, learning approaches, and criteria of computer assisted language learning, which gave us strong support to our assumptions of the effectiveness in the empirical study. We assume that the empirical study will show that the computer learners can fill in the gap between receptive and productive vocabulary knowledge, and they perform better than the paper learners in productive vocabulary knowledge.

To test the effectiveness empirically, we initiated the on-line testing on one hundred academic words for the inspection of the development of three vocabulary knowledge levels.
Three levels of vocabulary knowledge, which achieve a full coverage of vocabulary knowledge, were tested to see how each level developed over time, and how each level interacted with each other over time. The lowest level of vocabulary knowledge, active recognition, is to recognize the word form among three other options from English definition. The middle level, active recall, is to recall the word form from English definition. The highest level, controlled production, is to link the word form, word meaning and word use. We compiled the data from five intermediate English learners and analyzed the data with three steps corresponding with three research questions of the study. The first step confirmed the existence of the gap between receptive and productive knowledge; the second step explored interactions between three vocabulary knowledge levels through variability analyses; the last step explained the vocabulary development of the paper learner and the computer learners by the logistic model (Verhulst, 1845). This model can assist to interpret three predominant parameters in vocabulary development, which are initial value, learning rate, and carrying capacity. The initial value refers to the first score learners obtain after learning vocabulary; the learning rate refers to the slope of vocabulary development; the carrying capacity refers to the sum of all the resources learners can obtain in the end. Determining the value of these parameters, we might be able to explain the difference of vocabulary development between the paper learner and the computer learners.

The study begins with the background on vocabulary learning in general and computer assisted vocabulary learning in particular, followed by the evaluation of WordChamp, the specification of the research questions, the framework of exploring vocabulary development, and the corresponding methodology. The subsequent results section consists of three parts, in line with three research questions mentioned above. Each part of result is followed by a brief discussion, which is later elaborated and summarized in the final discussion. The subsequent conclusion describes the characteristics of vocabulary development, gives suggestions to WordChamp learners or lecturers and deduces important elements of a computer assisted vocabulary learning (CAVL) website for future research work.
Chapter 3

Background

3.1 Vocabulary learning

Vocabulary learning has been taken as an important component of acquiring a second language (L2) since learners should hold high percentage of vocabulary coverage to speak, to read, to write and to understand L2 (Schmitt, 2008). English learners should have 95% coverage of the vocabulary to sufficiently understand the spoken discourse (Laufer, 1989). If the coverage of the vocabulary is lower than 80%, it would lead to very poor comprehension. (Bonk, 2000). Learners should know around 8000 to 9000 word families, which is 98% coverage of British National Corpus data, to be able to read authentic texts (Nation 2006).

Learners certainly need a large amount of vocabulary to achieve the four aspects of language skills, and each skill requires different aspects of “word knowledge” (Nation, 2001, p. 27), which are word form, word meaning and word use. Learners should know how to produce the spoken word form and link the word form and the word meaning in a short time in order to speak and to comprehend English, which are seen as receptive word knowledge. Learners should be able to recognize and produce the written form and the grammatical word usage in order to read and to write English, which are seen as productive word knowledge. We will first discuss the difference between receptive word knowledge and productive word knowledge and then investigate the subparts of word knowledge to know what word knowledge exactly learners should learn so that they can use the word receptively and productively.

3.2 Word knowledge

3.2.1 Receptive vs Productive knowledge

As we mentioned in the previous paragraph, “word knowledge” can be utilized as “receptive
vocabulary knowledge” or as “productive vocabulary knowledge.” Receptive knowledge is related to reading and listening comprehension while productive knowledge is related to writing and speaking (Laufer, 2004). The receptive knowledge of a learner is naturally larger than the productive knowledge because he has more chances to practice the receptive knowledge, but he needs to spend extra precise learning for productive knowledge. In consequence, receptive knowledge grows faster than productive knowledge in second language (Laufer & Paribakht, 1998) and receptive knowledge does not transfer directly to productive knowledge (Fan, 2000). A gap between receptive knowledge and productive knowledge is significant for most of the learners due to the different growing rate of them. For instance, a person might be able to recognize the word “orientation” in the reading in a sense of receptive vocabulary knowledge, but he/she might not be able to produce “orientation” in his/her writing because he/she might not know either how to spell the word or how to use the word in the right context. Accordingly, it is important to distinguish the receptive and productive word knowledge to evaluate to what extent a learner can receptively or productively use the word. However, to assess receptive and productive knowledge is not sufficient to evaluate vocabulary learning. Important subparts of “word knowledge”, including word form, word meaning and word use, should also be taken into account because each subpart supports the development of four aspects of language skills.

3.2.2 Word knowledge: Word form

Three subparts of word knowledge are spoken form, written form and morphological composition (Nation, 2001, p.27). If learners are able to recognize the word when hearing it or to pronounce the word when expressing meanings, they are able to use the spoken form receptively and productively. Spoken form has an effect on effectiveness of memorizing the word. The more pronounceable the word is, the more likely the learners are to pick up the word (Ellis et al, 1993). Our target website: WordChamp has the function of practicing the spoken form so it would facilitate the learner to memorize the word.

Spoken form is related to written form since the learners need to know the spoken form
and then use it to spell the word. If learners are able to read with high fluency or to spell the words right, they are able to use the written form receptively and productively. When learners cannot manage to use the written form, they tend to use easily-spelt words and avoid the words that are hard to spell, which would lead to very limited vocabulary in writing. WordChamp has the function of practicing the written form by giving the L1 translation or the pronunciation of the target word so we assume that learners could consolidate the written form more easily in the brain.

Written form would be facilitated if the knowledge of “morphological composition” is acquired. Applying morphological compositionality, learners can recognize or spell the words with suffixes and affixes and subsequently are more likely to guess the meaning of the word right and are able to learn the word family with ease. WordChamp does not offer the knowledge morphological composition or put the same type of word families together so learners might not be able to link the word form and the word meaning or retrieve the word in a short time.

Knowing the word form is not sufficient to retrieve the word. Linking the word form and the word meaning is more relevant for word retrieval. For example, learners might have a concept of a meaning but cannot retrieve the word form so they cannot speak or write the word. On the other hand, learners might know the word form but cannot retrieve the word meaning so they might miss out some important information from the reading or the speaking discourse. Connecting the word form and the word meaning is undoubtedly important for vocabulary learning. Moreover, there are words of many different meanings, which will multiply the difficulty of word learning. The subparts of word meaning would help learners memorize different meanings effectively and distinguish which word meaning should be used in what context.

3.2.3 Word knowledge: word meaning

Three subparts of word meaning are forms and meaning, concept and referents, and associations
(Nation, 2001, p.27). The better learners can connect the form and the meaning, the more likely they are to retrieve the words while reading or hearing them. Every time learners retrieve the form or the meaning, they “strengthen” the link between them (Baddeley, 1990). WordChamp applies the usage of word flashcard to vocabulary learning so it would “consolidate” the link between the form and the meaning. The word “consolidate” refers to the second stage of vocabulary acquisition process, which is “consolidation.” In 3.3, more explanation of “consolidation” will be presented.

As we discussed above, some words have many different meanings, especially high-frequency words, and some words are with identical written forms (bridge and bride) or spoken forms (son and sun). These features multiply the difficulty of learning vocabulary. Holding the knowledge of “concept and referents” would mitigate the difficulty. Learners can comprehend a word in a context by “its inherent lexical meaning”, which is the word meaning when it exists “in isolation.” Learners can also comprehend a word by “the inferential meaning”, which is the meaning we can infer from other words or from “the knowledge of the world” (Ruhl, 1989). If a CALL program offers vocabulary-focused contextualized reading, it would assist learners to clarify the confusing word meanings and word forms.

Moreover, learners can also make use of “associations” such as synonym and antonym, to distinguish the word meanings and word forms and to enrich their understanding of the words. If a CALL program encourages learners to find out the synonym and the antonym of the words, it would connect many similarly-functioned words together to accelerate the consolidation of word meaning in the brain.

After being able to recognize, spell the word form and link the word form and the word meaning, learners should learn to possess the knowledge of how to use the words. Only possessing the word form and the word meaning does not guarantee that learners would use the word in the right grammatical use, that they would use the word in an authentic or native-like sense and that they would use the word in the appropriate situation. The word use should be emphasized to enhance the ability to use the words more precisely.
3.2.4 Word knowledge: word use

Three subparts of word use are grammatical functions, collocations, and constraints on use. Learners should use the words with correct “grammatical functions.” They must know what part of speech the words are and how the words are used to grammatically fit into a sentence. In order to produce the words more fluently or to produce more native-like expressions, learners should have knowledge of “collocation”, which is what words the target word typically occurs with. Learners also should have knowledge of “constraints on use,” which means to know where and when certain words can be used. For example, some words are very common in one country but they might be inappropriate in another country. Some words are of very low frequency and they are not very appropriate for productive use. WordChamp offers learners the opportunity for word use such as editing sample sentences and typing the right verb form so learners are more likely to gain accuracy of their writing.

Now we know how word knowledge supports four aspects of language skills and how a CALL program can benefit from word knowledge in vocabulary learning. But shall we present all the word knowledge at the same time? Is there a sequence of acquiring word knowledge? On the top of that, how can learners acquire the word knowledge rapidly and effectively? Are there any superior approaches to word learning? We will present the answers to above questions by discussing “vocabulary acquisition process” (Verspoor & Lowie, 2003) and two approaches to vocabulary learning, which are intentional and incidental vocabulary learning.

3.3 Vocabulary acquisition process

Learners acquire vocabulary by going through two stages: “semantization” and “consolidation” (Verspoor & Lowie, 2003). At the stage of “semantization”, learners notice and link the word form and the word meaning and incorporate the word meaning into the semantic network. Later, at the stage of “consolidation”, they incorporate the word into the long-term memory by inference, repetition, example or image. Learners process the words in the brain by guessing the word meaning in the context, by reading sample sentences, by practicing recalling the vocabulary or by associating the words with images and other words. The more elaborate the
learning paths they use, the better the recall of the word will be.

In consequence, we can assume that learners gain the word knowledge in a sequence of word form, word meaning and then word use since word use needs more elaborate learning paths to achieve and learners can only be ready to learn word use when they already have the knowledge of word form and word meaning. If a CALL program follows the sequence of vocabulary acquisition, learners would feel more ready but still challenging while learning vocabulary.

Consequently, the vocabulary acquisition process should be enhanced by effective and efficient vocabulary learning approaches. For example, incidental learning can make learners link word form and word meaning (semantization) and guess the word meaning from the context (consolidation). Intentional learning would make learners practice words repeatedly and associate word meaning with images or other words (consolidation).

3.4 Approaches to vocabulary learning

Different types of “word knowledge” require different learning approaches. For example, collocations are more likely to be acquired when the learners read articles, but word form is more likely to be acquired when the explicit word teaching occurs. Vocabulary learning approaches play an important role in the acquisition of word knowledge.

Two common vocabulary learning approaches are incidental vocabulary learning and intentional vocabulary learning. Incidental vocabulary learning focuses on meaning-based input while intentional vocabulary learning focuses on the words explicitly. Each of them has its advantages and disadvantages. Our target website “WordChamp” (Daniel Blumenthal, 2006) integrates two vocabulary learning approaches and thus benefits the learners with the advantages from both approaches.

3.4.1 Incidental vocabulary learning

Incidental vocabulary learning exposes the learners to the meaning-focus input where learners can
consolidate and enhance their unfamiliar vocabulary. Learners intend to infer the word meaning from the context so they guess the meaning of the words, which is taken as a useful strategy to deal with the unknown words in reading (Schmitt, 1997). But the effectiveness of “guessing” is questioned because learners of lower proficiency can hardly guess the meanings right and the learners of higher proficiency sometimes solely “guess” without checking whether the “guessing” is right or not. Learners can only strongly enhance the word form when meeting the word but word meaning and word use are less enhanced. Incidental learning only offers part of the word knowledge instead of full level of mastery (Pigada et al, 2006). Moreover, if learners are not exposed to new words frequently enough, the rate of picking up new words is relatively low and the ability to produce words is difficult to gain. The effectiveness of incidental learning still requires some “consolidation” activities such as “glossing” (Nation, 2001) or “explicit follow-up” to enhance and it is still not as time effective as intentional vocabulary learning (Mondria, 2003).

3.4.2 Intentional vocabulary learning

Intentional vocabulary learning makes learners consciously learn new words either in context or in isolation. (Schmitt, 2008) The main task of intentional vocabulary learning is to memorize the words. However, memorization itself is not sufficient for the mastery of full knowledge of vocabulary and for long-term retention. There should be additional activities focusing on “semantization”, as we discussed in 3.3 for vocabulary acquisition process, which is engaging the learners with the target words to have deeper word knowledge and on recycling to pick up the words which might be forgotten quickly.

Accordingly, neither incidental nor intentional vocabulary learning alone is beneficial to vocabulary learning. The combination of incidental learning and intentional learning is proposed to be a better method than only incidental or intentional learning. CALL combines the usages of incidental and intentional learning and thus is considered to be more promising. Some CALL programs have solved the problems caused by incidental learning and intentional learning and they were soundly based on vocabulary learning theory and did improve vocabulary learning. In
the next section, we would take a closer look at how CALL programs can help improve vocabulary learning and in which aspect WordChamp is taken as a promising vocabulary learning website.

3.5 Computer assisted Language learning (CALL)

CALL is to teach a second language by integrating computer into the learning process and has been applied to teaching content-based language, communication skills, speaking, listening, writing, reading, grammar and vocabulary (Chapelle, 2008). Vocabulary learning has been widely investigated in CALL (Stockwell, 2007). CALL has solved the problems of incidental vocabulary learning with “glossing”, which offers first language translation or second language definition to avoid guessing the word meaning wrong. WordChamp has a function of webreader to enable learners to read foreign language articles with a cursor including the first language definition. Glossing makes more difficult texts readable and minimizes the interruption of looking up the words when reading (Nation, 2001). CALL has also solved the problems of intentional learning with “explicit activities”, which automatically listing out all the target words after reading and providing exercises with the target words. What’s more, CALL has improved intentional vocabulary learning by having the learners practice the unfamiliar vocabulary with the electronic flashcards or electronic drills, which fastens the consolidation of vocabulary knowledge. Many research-based CALL programs were developed based on word learning theories. The designers of these CALL programs select a particular theory of vocabulary learning and implement it with computer technology.

Computer Assisted VOCabulary Acquisition (CAVOCA) is an English vocabulary learning program developed by Groot (2000). The background vocabulary learning theory it relies on is “L1 word acquisition process.” There are three stages of L1 word acquisition: noticing, storage and consolidation. CAVOCA takes the learners to go through the sequence of mental operations (Groot, 2000). Learners first deduce the meaning of the word from three sample sentences and judge whether the word is correctly used in a new sentence so they would “notice” the syntax, and the semantics of the word. This stage is similar to the first stage of vocabulary learning.
process: “semantization.” Learners have to link the word form and the word meaning by processing the sample sentences, which will aggregate the consolidation of the words. Next, the explanations of whether the word is correctly used in two new sentences and other additional information of the word are given in descriptive sentences to incorporate the word into the mental lexicon. The target word will be presented with a number of authentic L2 passages to ensure long-term retention and to motivate the learners as they are able to understand the authentic L2 passages. The last two stages are compatible with vocabulary learning process: “consolidation” by giving learners elaborate examples of word. After having the same learning process of 25 target words, the learners have to do vocabulary assessment to retrieve the word.

“CAVOCA” was proved to mitigate retention loss in receptive vocabulary knowledge than the condition in which “bilingual word list” was used, but the retention loss in productive vocabulary knowledge still remained unknown (Groot, 2000). The gap between the receptive and productive knowledge was not filled with assistance of CAVOCA. However, theoretically speaking, CAVOCA still covers three aspects of word knowledge, makes good use of the vocabulary acquisition process by combining semantization and consolidation and applies both incidental (guessing the meaning) and intentional (direct vocabulary teaching) vocabulary learning approaches.

“Lexxica” is another English vocabulary learning website, which was developed by Browne (2005). It first tests the coverage of vocabulary of the learners to decide what high-frequency words they should learn. Later, it makes good use of “spaced repetition process” to assist learners memorize vocabulary rapidly by recalling a certain word at a certain time (Browne, 2007). In contrast with “massed repetition”, spaced repetition would “spread” the word recall with the same amount of time spent in massed repetition. It leads to longer retention than massed repetition because the recall is more effective when it is conducted right after the word is studied since most of the forgetting occurs soon after learning. Plus, the space between each later repetition should be larger than that between each earlier repetition (Pimsleur, 1967). This “spaced repetition process” was computerized by “Lexxica” to help learners memorize the words.

Lexxica covers three aspects of word knowledge, follows the vocabulary acquisition process
by linking word form, word meaning and word use repeatedly, and combines incidental learning (guessing the meaning) and intentional learning (having word list after completing practice). However, it still does not fill in the gap of receptive and productive vocabulary knowledge since most of the practice in Lexxica only focuses on receptive knowledge.

Even though the websites sound very promising, there is no empirical study to prove the effectiveness of “Lexxica” and there is no positive result for the productive knowledge in CAVOCA. Since “CAVOCA” and “Lexxica” are based on ideal vocabulary learning theories and they give explicit learning instructions to the learners, why weren’t they proved to be effective on receptive and productive vocabulary knowledge? It is probably because they both overlook some important perspectives of second language acquisition such as interaction with other learners and the development of autonomy. The target website of our study: “Wordchamp” is more integral than those two vocabulary learning programs, because it includes more elaborate paths of consolidating vocabulary and takes second language acquisition into account.

3.6 Introduction to Wordchamp

“WordChamp” is an open source website which is used to build up learners’ own vocabulary in many languages for self-study or class use. Learners memorize vocabulary with the electronic drills. There are five main features of WordChamp: flashcard, drill, absolute recall, web reader and course management. The learners can build up their own flashcards, share them with other users, effectively practice the drills to memorize the target words and read authentic websites or articles with the assistance of web reader. The teacher can manage the score, learning time, and learning progress of the students with course management.

3.6.1 WordChamp and vocabulary acquisition process

“WordChamp” is mainly based on “behaviorist learning” (Burston, 2005) and is also compatible with vocabulary acquisition process: semantization and consolidation. When creating flashcards, learners start to link the word form and the word meaning (semantization). Different types of flashcards would assist learners to incorporate the words into the semantic network. For example,
learners would assign “optimistic” a Chinese meaning “le guan de” with translation flashcard, they would define “optimistic” as “expecting something good to happen” with definition flashcard, and they would associate “optimistic” with its antonym “pessimistic” with antonym flashcard.

When practicing the electronic drills, learners “consolidate” the words with a great number of retrieval paths. For example, they receive the input of the word “turtle” by the sound, the picture, the definition, the first language (L1) translation, or the sample sentences and they produce the word “turtle” by typing down the words or pronouncing the words. Furthermore, they can read an article by guessing the unknown word first, got the word meaning with a cursor shown immediately and saved it into the flashcard list automatically for repetition practice. They can also edit their own word attributions to enhance their word use. WordChamp offers various paths of consolidating words, which are considered impossible when learners learn vocabulary with paper. Table 1 lists out all the relevant functions of WordChamp which enhance semantization and consolidation. All the flashcards aim at the stage of semantization while all the drills and webreader aim at the stage of consolidation.
<table>
<thead>
<tr>
<th>Function of WordChamp</th>
<th>How it works</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pronunciation drill</td>
<td>Learners can practice the pronunciation of the words and compare it with that of a native speaker.</td>
</tr>
<tr>
<td>Reading practice drill</td>
<td>Learners see the word, try to read it by themselves and compare their pronunciation with that of a native speaker.</td>
</tr>
<tr>
<td>Flashcard Review drill</td>
<td>Learners translate either from English to Chinese or from Chinese to English orally.</td>
</tr>
<tr>
<td>Listening comprehension drill</td>
<td>Learners hear the English words and type down the translation of Chinese or the other way round.</td>
</tr>
<tr>
<td>Dictation drill</td>
<td>Learners hear the word and spell the word.</td>
</tr>
<tr>
<td>Translation drill</td>
<td>Learners translate either from English to Chinese or from Chinese to English by typing the words down.</td>
</tr>
<tr>
<td>Definition drill</td>
<td>Learners see the definition and type down the word or vice versa.</td>
</tr>
<tr>
<td>Antonym drill</td>
<td>Learners see the English word and type down the antonym.</td>
</tr>
<tr>
<td>Synonym drill</td>
<td>Learners see the English word and type down the synonym.</td>
</tr>
<tr>
<td>Verb basis drill</td>
<td>Learners type down the right verb form of present tense, past tense and past perfect tense.</td>
</tr>
<tr>
<td>Web reader</td>
<td>Learners paste an online article or a web link and read the article with an L1 cursor next to L2. Learners can save the new word in their flashcard immediately.</td>
</tr>
<tr>
<td>Course management</td>
<td>Teachers can use course management to see how much time learners spend on practice, how well learners do and what words learners should practice more.</td>
</tr>
</tbody>
</table>

Table 1: How functions of Wordchamp work in practice

3.6.2 Advantages of WordChamp

There are several advantages which “Wordchamp” has over traditional paper-based vocabulary learning. First, “audio flashcard”, “image flashcard” and “audio drill.” provide learners with images and verbal presentations which are proved to be very effective in mnemonic techniques (Pavio et al, 1980). Second, “Wordchamp” offers many kinds of drills to make learners practice the words repeatedly and records the practice to produce “effective recall”, which means the learners will be exposed to unfamiliar words with extra practice. Third, “WordChamp” provides incidental learning with an efficient tool, which is a cursor with L1 translation next to the
unknown L2. Learners can read authentic articles or websites without guessing the words wrong and they are more likely to read more difficult articles with this function. Fourth, learners can edit sentences of the target words to productively use the words so they are likely to have better performance in productive vocabulary knowledge. Fifth, learners can train the fluency of retrieving the vocabulary in speaking, listening, reading and writing with translation, listening comprehension, and dictation drills. Sixth, it gives detailed feedback on how well the learners do on vocabulary learning, what words they still have to work on and how much time they spend on learning so the learners would develop more autonomy of learning vocabulary by themselves in the future. Seventh, it is a language learning community and the learners of WordChamp can communicate through this network, which can bring about more interactions between learners.

3.6.3 WordChamp and Word knowledge

A promising vocabulary learning website should cover wide range of “word knowledge” to equip learners the ability to speak, to read, to write and to comprehend L2. Learners of “WordChamp” can gain knowledge of spoken form and written form but they cannot gain the morphological composition since it does not offer any knowledge of suffixes and prefixes because the flashcards are usually created by learners themselves. They are not likely to put associated words together to facilitate memorization. Learners can be familiar with the word meaning, associations of the word and some knowledge of how words should be used in different context when learners edit their own word sentences but they might not be very aware of editing the sentences with different meanings of the word. Learners can obtain the grammatical functions and collocations but not constraints on use since the overall focus of the website is individual vocabulary learning instead of context-based vocabulary learning. Table 2 summarizes how functions of “WordChamp” enhance “word knowledge.” It covers wide range of “word knowledge” since it includes six perspectives of word knowledge in receptive and productive senses. It is a very promising CALL program for vocabulary learning and seems to fill in the gap between the receptive and productive knowledge.
<table>
<thead>
<tr>
<th>Word Knowledge</th>
<th>Concrete concepts of word knowledge</th>
<th>Wordchamp Function</th>
</tr>
</thead>
</table>
| Form           | Spoken                             | 1. All kinds of flashcard are with audio  
2. Pronunciation drill  
3. Reading practice drill  
4. Flashcard review  
5. Listening comprehension drill  
6. Dictation drill |
| Written        | Learners should be able to recognize the “spoken form” when hearing the word or say the “spoken form”.  
1. Building up the flashcards  
2. Dictation drill  
3. Translation drill  
4. Listening comprehension drill |
| Word part      | Learners should be able to recognize what “word parts” the word is made of and associate “word parts” with word meaning or construct the word with the right “word parts.” | None |
| Meaning        | Form and meaning                   | 1. Definition flashcard  
2. Definition drill  
3. Flashcard translation review |
| Concept and referents | Learners should know the meaning of the word when it is used in different contexts or produce the word in different contexts. | 1. Edit sample sentences |
| Associations   | Learners should know the related words of the target word or produce the synonym or the antonym of the target word. | 1. Synonym flashcard  
2. Antonym flashcard  
3. Antonym drill  
4. Synonym drill |
| Use            | Grammatical functions              | 1. Edit sample sentences in all flashcards  
2. Edit sample reading in all flashcards  
3. Verb basis drill |
| collocations   | Learners should know the collocated words of the target word or use the collocated words or the target word in the writing. | 1. Edit sample sentences  
2. Web reader |
| Constraints on use | Learners should know the property of the word or use the word in the most suitable situation. | None |

Table 2. Word knowledge and functions of WordChamp
However, even though the website can provide sufficient “word knowledge”, we still cannot be sure whether it is an effective website. It is also important to see whether it is compatible with second language acquisition on which the criteria of computer assisted vocabulary learning programs (Chapelle, 2008, pp.37) are based. The criteria consist of “learner fit”, “explicit vocabulary teaching”, “interaction with the computer”, “interaction with other learners”, “evaluation” and “strategy development.” We will take a closer look at whether “WordChamp” fits the six criteria and how it fits or does not fit each criterion.

“Learner fit” is whether learners learn the words which they might use or need and which are of their proficiency level. All the flashcards of “WordChamp” are categorized with different levels and frequency. Learners can easily find levels and frequency which they might fit to practice. Teachers can create their own flashcards for a certain course and share them with the students. In this sense, the website is appropriate for learners with the characteristics of the intended learners. Moreover, “WordChamp” gives different levels of learners a wider range of word learning strategies. Beginners tend to use more “translation flashcard” since they do not know many words to construct a sentence. On the contrary, the intermediate or advanced learners tend to use “definition” “anonym” or “synonym” flashcards, which are more effective when the learners have better proficiency.

“Explicit vocabulary teaching” is whether the program provides explicit instruction to teach vocabulary. “Wordchamp” mostly focuses on intentional vocabulary learning and to some extent turns incidental learning into an intentional one by using “webreader.” Learners can directly save the new words into the flashcards from the online article or website for future practice.

“Interaction with the computer” is whether computer gives help when learners request the word meaning and when they practice vocabulary exercises. “Wordchamp” has webreader to assist learners to read online websites or authentic articles but the meaning of the vocabulary in a certain context will not be recognized by the website but by the learners themselves. Learners have to select the right meaning of the word in the context. Moreover, it gives immediate
feedback on all the practice drills with correct answers so learners are able to correct themselves while learning.

“Interaction with other learners” is whether the program guides learners to work with classmates or have on-line discussions and whether there are any collaborative activities. Learners can exchange flashcards, send messages to other learners and learn in a vocabulary community, which is like a real world learning environment. These would lead to more authentic and collaborative learning.

“Evaluation” has tow categories: whether the website provides feedback to learners about their responses and whether the website provides evaluation of learning outcomes through drills that give learners information about their performance. “WordChamp” gives immediate feedback of the drills and the results of the drills such as how much time is spent, what words are missed for how many times, and what words are overridden for how many times. Teachers can use “course management” to know the time learners spend on the course and the score learners get. The scores are stored and analyzed with bars or graphs so it is easy to track each learner’s progress. Learners can know exactly which word they are still not familiar with and practice them with “absolute recall”, where computer saves the most frequent forgotten answers.

“Strategy development” is whether the website promotes good vocabulary learning strategies and provides guidance for students to develop strategies that will help them continue to learn vocabulary with the website. Learners should learn deeper knowledge of a word and gradually develop a habit of learning the words in depth instead of just knowing the translation. A promising CALL program improves the learners’ ability to “produce” the vocabulary instead of just “receiving” the vocabulary. “WordChamp” promotes various learning strategies of learning vocabulary by giving the learners opportunities to acquire different aspects of word knowledge and to practice with different types of drills. Learners have more interactive activities to “utilize” the words they learn by themselves so they can develop their own learning strategies and their autonomy will be raised up.

However, there is no empirical study on the vocabulary learning effectiveness of “WordChamp.” We will conduct The Longitudinal Academic Vocabulary Test (LAVT) (Caspi &
Lowie, 2000), which incorporates continuous receptive and productive vocabulary tests during the process of vocabulary learning, to determine how learners of “WordChamp” acquire the receptive and productive vocabulary and whether they perform differently from the paper learners.

3.7 Receptive vocabulary test and productive vocabulary test

3.7.1 Computer adaptive test of size and strength (CATSS)

LAVT combines adaptations of two vocabulary testing approaches and consists of three parts, active recognition, active recall and controlled production. The receptive vocabulary tests, which are active recognition and active recall, are from two parts of the Computer Adaptive Test of Size and Strength (CATSS) test (Laufer et al., 2004). CATSS tests the different dimensions of word knowledge for which learners manage to link the word form to the word meaning. There are four levels of word knowledge from the lowest level to the highest level: passive recognition, active recognition, passive recall, and active recall. There are two distinctions that we have to clarify here. The first distinction is "passive" and "active." The passive ability is to retrieve the "word meaning" and the active ability is to retrieve the "word form." The second distinction is "recall" and "recognition." The recall ability is to recall the word form or the word meaning and the recognition ability is to recognize the word form and the word meaning. In consequence, "passive recognition" is to recognize the word meaning. "Active recognition" is to recognize the word form. "Passive recall" is to recall the word meaning. "Active recall" is to recall the word form.

We will use "active recall" and "active recognition" instead of "passive recall" and "passive recognition", which means we want to focus on the retrieval of word form rather than word meaning because the word meanings are always context-dependent so answers to the recall of the meaning would vary and because recognizing the word meaning is so simple for intermediate learners that the results might not present a significant difference between learners of “WordChamp” and those who use paper. By observing the score of "active recall" and "active recognition", we wish to find out whether the learners can connect the word meaning with the word form by spelling out the word or by recognizing the word. However, the test which only
focuses on the connection between the word meaning and the word form is not sufficient. We also want to look at whether learners can comprehend the words in a context and retrieve the word form on which the Lexical Frequency Profile (LFP) (Laufer & Nation, 1995) would test.

### 3.7.2 Lexical Frequency Profile (LFP)

The productive vocabulary test, which is controlled production, is from Lexical Frequency Profile (LFP). LFP has been proved to be reliable and valid measure of lexical use in writing. In our study, we will only focus on the “active version” of the vocabulary levels tests: “controlled production”, instead of analyzing the lexical level of compositions because the testing period only lasts for three weeks and the learners will not show very significant difference in the writing. We cannot expect learners to make evident progress in such a short time because testing writing improvement takes more time than vocabulary acquisition. “Controlled production” elicits the use of the target words in sentences and gives the initial of the target words in order not to have non-target word filled in. This test requires more types of word knowledge such as "collocations" and "grammatical use" instead of only the word form and the word meaning. By testing the participants on “active recall”, “active recognition” and “controlled production,” we will know how well they possess the three main parts of word knowledge: the word form, the word meaning and the word use.

Three vocabulary tests, which are active recognition, active recall and controlled production, were tested on learners for forty two times in twenty one days. We intend to find out, in the first section, whether the gap between receptive and productive knowledge is filled and how each vocabulary level interacts with each other, followed by how the computer learners perform differently from the paper learners to determine the effectiveness of WordChamp. Instead of only looking at the final results of the learners, we adopt Dynamic Systems Theory (DST) to conduct analysis, because it may be appropriate for explaining the receptive and productive gap.
3.8 Dynamic Systems Theory (DST)

In essence, DST (De Bot, Lowie and Verspoor, 2007) is a theory of change. Some changing components not only influence the whole dynamic system but also interact with each other. Moreover, every change over time depends on the previous state of the system so a small change at one stage will lead to a dramatic change at the final stage. Vocabulary learning, like other dynamic systems, develops nonlinearly and is influenced by its changing components such as active recognition, active recall and controlled production and by the interaction between these changing components. The number of acquired vocabulary at one point of time is determined by the previous point of time. The merit of DST with respect to investigating vocabulary learning is its emphasis on the development process itself instead of the deterministic results. It addresses nonlinear growth and changing interaction between changing components over time. Our study is a longitudinal study of vocabulary learning development and interactions between three vocabulary levels on the basis of DST. The general trend line, the detrended data, the moving correlation are used to interpret the vocabulary learning development and interactions between three vocabulary levels. In the final stage, we intend to find out how the paper learner perform differently from the computer learners by fitting the collected data with logistic model. This model will be explored in the section that follows.

3.9 Logistic Model

The logistic model was originally used to describe the population growth over time by Verhulst in 1845. The growth of vocabulary development is similar to that of population, which is slower in the initial stage, fastest in the middle stage and saturates in the final stage. What’s more, the rate of population increase might be limited due to its dependence on population density. The rate of vocabulary development increase might be limited due to its dependence on the resources learners have by themselves or learners can obtain from outside environment. Van Geert (2002) developed the logistic model applied on vocabulary development.

In Verhulst’s logistic model applied on population growth, three parameters were interpreted as maximum possible rate of population growth as the initial number of the population
and as maximum sustainable population. This continuous logistic model can be described by the
differential equation. The number of population varying with the time can be described as:

$$\frac{dN}{dt} = \frac{rN(K-N)}{K}$$

where $N$ represents the number of population, $t$ represents time, $r$ represents the maximum rate of
population growth, $K$ represents the carrying capacity.

Van Geert developed this logistic model and applied it on modeling of vocabulary
developmental processes. In Van Geert’s logistic model, three parameters can be interpreted as
maximum possible rate of vocabulary learning, as the initial value of known words and as
maximum word learning. Quantifying the learning rate, the initial value, and the learning
capacity, we were able to draw a difference that the computer learners have from the paper
learners in different vocabulary knowledge levels.

3.10 Research questions and predictions

After evaluating WordChamp theoretically, we find that it covers a wide range of both receptive
and productive word knowledge to enhance four aspects of language skills and to fill in the gap
between receptive and productive knowledge. It also follows the sequence of vocabulary
acquisition process from semantization to consolidation in case learners should not be ready. It
combines the incidental learning and intentional learning and solves the problems of each
learning approach. Most important of all, it serves vocabulary learning not only as vocabulary
learning theory but also as a second language acquisition, which includes various learning
strategies such as interaction and self-autonomy. WordChamp seems to be effective on
vocabulary learning. There are three main research questions aroused from testing the
effectiveness of WordChamp in this study.

1. The first question is whether WordChamp fills in the gap between receptive and productive
vocabulary knowledge. This question is addressed by observing whether the gap remains stable or
changes over time through the course of development on all vocabulary knowledge levels.
2. The second question is the how three vocabulary knowledge levels of the paper learner and the computer learners interact with each other. This question is addressed by inspecting the shifts of the moving correlations between all vocabulary knowledge levels.

3. The final question is whether the computer learners perform differently from the paper learner. Since the vocabulary development can be accounted for by the logistic model, this question is addressed by fitting the collected data with the logistic model. The values of three parameters, initial value, learning rate, and carrying capacity, can be interpreted as the difference of the paper learner and the computer learners.

Based on the background literature, the predictions concerning these questions would be described as below. Regarding the first question, we expected that the gap between receptive and productive vocabulary knowledge was more possibly filled in by the computer learners than the paper learner. Concerning the second question, we expected that the lower vocabulary knowledge level would support the higher vocabulary knowledge level more strongly for the computer learners than the paper learner. With regard to the final question, we expected the logistic model to fit the collected data well and to prove that the computer learners can perform better than the paper learner on all vocabulary knowledge levels, especially the highest level, controlled production.
Chapter 4
Methodology: Data collection

The framework of the experiment is configured on the basis of the six stages of case study methodology (Singer and Willet, 2003). They are data collection, data description, data exploration, model specification, model fitting and considering extensions. We used data collection to explain “methodology”: how the experiment was conducted in detail so that it could be manipulated in the future and later used data description and data exploration to denote our preliminary results and discussions to see whether they met what we expected: the computer learners can fill in the gap between receptive and productive knowledge and the lower vocabulary knowledge level can support the higher vocabulary knowledge level more for the computer learners than the paper learner. Next, we continued with model specification and model fitting to see how differently the learners perform by looking at the initial value of the score, the learning rate and the learning capacity. After comparing the theoretical evaluation with the empirical study, we will be able to conclude how the computer learners perform differently from the paper learners. In considering extensions, we will suggest what learners or lecturers of WordChamp should use this website and what criterion future computer assisted vocabulary learning websites should take account.

4.1.1 Participants

There are one male and three female participants for experimental group, who used WordChamp to learn vocabulary and one male participant for control group, who used paper to learn vocabulary. They were observed in this longitudinal study for twenty one days consecutively. The age of them ranges from 24 to 38 years old so they are all adult learners, whose L1 is traditional Chinese. They all volunteered to join the study and were very motivated to learn English vocabulary. They were all of intermediate English learners since they were in the same English class six months ago but one of the learners in the experimental group had comparatively lower
proficiency because he used to have problem with listening comprehension and pronunciation.

4.1.2 Material

Since the participants are of intermediate level, we had to choose the vocabulary which they might not know. One hundred words were selected from the Academic Word List (AWL) (Coxhead 2000). Every tenth word was selected from ten sub-lists which range from high frequency to low frequency to avoid the effect of frequency. We did not choose one hundred words only from very high or low frequency because high frequency words can be more rapidly acquired both receptively and productively than low frequency. If learners find it too simple or too difficult, the effect computer has on learners would be reduced. We wish to eliminate the potential effect of word frequency on our results.

4.1.3 Procedure

Both experimental group and control group received similar learning instructions. They both spent a total of fifteen hours learning vocabulary. The learners in experimental group were asked to create flashcards of the target words in translation, definition, synonym, anonym and sample sentences. They consulted the online dictionary to build up the flashcards. They practiced all types of the drills of the target words. They were required to have three one-hour visual classes with the researcher on “Surfgroepen”. The long-distance classes were held for the use of “webreader,” which led learners from reading focusing on meaning to reading focusing on word learning in isolation.

The learner in the control condition made the word list on paper in translation, definition, synonym, anonym and sample sentences. He consulted the same online dictionary to build up the word list. He practiced the target words with translation, definition, antonym, synonym, and sample sentences on paper with his learning style.

We conducted The Longitudinal Academic Vocabulary (LAVT) (Caspi & Lowie, 2000):
“active recognition”, “active recall” and “controlled production,” which range from the lowest vocabulary level to the highest vocabulary level. Active recognition is operationalized as the ability to recognize the word form with English definition among three distracting answer choices and one correct answer as it is operationalized in the CATSS. The three distracting answer choices were shuffled by the computer itself from the database of 200 words which included 100 target words and 100 GRE words. The reason why we chose GRE words was that participants could not guess the answer by excluding other answer choices since they did not know other answer choices as well.

Example of active recognition (correct answer: legislate)

“To make a law”

a) preliminary  b) migrate  c) legislate  d) aspersion

Active recall, which is the second level of vocabulary knowledge, is operationalized as the ability to recall the word form with English definition and spell the word form by being given the initial.

Example of active recall (correct answer: reside)

“To live in a particular place. The word starts with R.”

Controlled production, which is the highest level of vocabulary knowledge, is operationalized as the ability to recall the word meaning, link the word form with the word meaning and spell the word form by being given two sample sentences and the initial of the word. Participants had to do the gap fill exercises according to two sample sentences because words are always used in different contexts and have various meanings. However, they only had to fill in the infinitive verb, singular noun, and the simple form of adjective and the adverb because the participants were not good enough to pay attention to word form, word meaning and word use at the same time. They might not be able to improve fast enough for us to notice the significant
difference in the score.

Example of controlled production (correct answer: norm)

“If you say that a situation is the n____, you mean that it is usual and expected. A n____ is an official standard or level that organizations are expected to reach.”

Both the experimental and the participant in the control condition conducted the same online test. They both did the test twice a day by connecting to our computer-based testing system programmed by PERL for twenty one days. Each test consisted of thirty questions with ten questions for each level of vocabulary knowledge. Each question was shuffled by the computer, and when they tested themselves again, the question would be different from the last time, which would mitigate the learning effect from the test. Moreover, participants did not know whether they answered the questions correctly or wrong or what the score was in case they learned from the test or they were discouraged or motivated by the scores.

After gathering forty-two scores, we plotted all the data points on the graph. We firstly looked at the raw data with the linear trend line to observe the general trend of the learners and then explored more with detrended data and moving correlation to see how each knowledge level interacted with each other. Since it is hard to put the results and the discussion apart to explain, here we will present the results with a brief discussions first and give an extended discussion later.
5.1. Data description

*Raw data and linear trend line*

At the stage of data description, we plotted forty-two scores on the y-axis and time on the x-axis with linear trend lines in figure 1 where CP is controlled production, ARL is active recall and ARN is active recognition. We observed the graph to determine the general trend of development and to visually compare three vocabulary knowledge levels. Both paper learner and computer learners had upward trend lines with all vocabulary knowledge levels. The paper learner had three nearly “paralleled” linear trend lines so all levels of vocabulary grew at the same speed. However, among the computer learners, it was either controlled production or active recognition had the steepest trend line. We observed all the initial values of all participants. They all had highest initial value on active recognition, a bit lower initial value on active recall, and lowest initial value on controlled production.

![Figure 1. Raw data and trend line](image)

Figure 1. Raw data and trend line
However, there was one exception among the computer learners, who did not make any progress after learning for twenty-one days. He reported that he needed help to segment a word to facilitate pronunciation (e.g., Con-tem-po-ra-ry) in order to read and spell the words. He also found that there were many unknown words in the test so he could not understand the meaning of the English definition or sample sentences at all. When he had the assistance of segmenting the word in English class and learned words on paper, he used to get very high score in vocabulary translation test but when the help from the instructor disappeared, he became helpless.

In a dynamic theory system (DST) (De Bot, Lowie and Verspoor, 2007), only how variables change over time can not fully depict a language learning development. Variables are constantly adapting to each other so how they interact with each other is also relevant to learning development. After understanding how each vocabulary knowledge level developed over time, we observed how each of them interacted with each other over time at the stage of data exportation.

5.2. Data Exploration

5.2.1 Detrended data

Since we wanted to look at how knowledge levels interacted with each other as a function of time, we connected all the 42 data points in Figure one to see how they interacted. However, the trends in themselves distracted us from focusing on the variability so we subtracted the linear trends from the raw data to calculate residuals, which facilitates us to observe the variability. Residual is the value subtracting the “model value” from “raw data”. The “model value”, which is the value on the trend line, is determined by the intercept and the slope of the trend line of active recognition

\[
\text{Function 1. model value} = \text{slope} \times \text{time} + \text{intercept}. 
\]

\[
\text{Function 2 Residual} = \text{Raw data} - \text{Model value}. 
\]
These residuals represent developmental variability and are considered a piece of developmental information in a dynamic approach (Van Geert and Van Dijk, 2002). We paired the three vocabulary knowledge levels with each other, which are active recognition with active recall, active recall with controlled production and active recognition with controlled production. In Figure 2, active recall and active recognition of the paper learner exhibit mostly inverse variability. In Figure 3, active recall and controlled production exhibit both parallel variability pattern from 0 to 6 and from 20 to 24 and inverse pattern throughout the rest of the time. In Figure 4, controlled production and active recognition of the control group show mainly parallel variability.

Figure 2: Paper learner: Residuals of ARL and ARN

Figure 3: Paper learner: Residuals of ARL and CP
Figure 4: Paper learner: Residuals of CP and ARN

Figure 5: Computer learner 1: Residuals of ARN and ARL

Figure 6: Computer learner 2: Residuals of ARN and ARL

Figure 7: Computer learner 3: Residuals of ARN and ARL
In Figure 5, 6 and 7, active recall and active recognition of the computer learners show mostly inverse variability to those for the paper learner. In Figure 8, 9 and 10, active recall and controlled production show mostly parallel variability and sometimes inverse variability as those for the paper learner. In Figure 11, 12 and 13, controlled production and active recognition exhibit mainly parallel variability as those of the paper learner. Table 3 summarizes the interaction between three vocabulary knowledge levels.

Figure 8. Computer learner 1: Residuals of ARL and CP

Figure 9. Computer learner 2: Residuals of ARL and CP

Figure 10. Computer learner 3: Residuals of ARL and CP
Figure 11. Computer learner 1: Residuals of ARN and CP

Figure 12. Computer learner 2: Residuals of ARN and CP

Figure 13. Computer learner 3: Residuals of ARN and CP

<table>
<thead>
<tr>
<th>Interaction</th>
<th>Paper learner</th>
<th>Computer learner</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARN &amp; ARL</td>
<td>Negative</td>
<td>Negative</td>
</tr>
<tr>
<td>ARL &amp; CP</td>
<td>Sometimes positive,</td>
<td>Mostly positive</td>
</tr>
<tr>
<td></td>
<td>Sometimes negative.</td>
<td></td>
</tr>
<tr>
<td>ARN &amp; CP</td>
<td>Positive</td>
<td>Positive</td>
</tr>
</tbody>
</table>

Table 3. Interaction between active recognition (ARN), active recall (ARL) and controlled production (CP)
However, a deterministic correlation value cannot disclose how it changes over time as we observed how three variables changed over time. Some information would be overlooked if we solely look at the deterministic correlation value. For example, we can correlate “the amount of food you eat every day” and “the weight” and can probably find out there is strong positive correlation between them. But it is not always true that the less you eat, the less you weigh because our age might slow down the decline of weight or because the food has different calories. Some other variables might interfere the correlation value over time or the variables themselves might have more variables within themselves. In our study, as Table 3 shows, the correlation between the active recall and the controlled production of the paper learner fluctuates between the positive and the negative value. Accordingly, at the second stage of data exploration, we observed the correlation as a “moving window” to compare the difference between the deterministic correlation and the moving correlation to find out other internal or external variables which could have influence on the development of vocabulary learning.

5.2.2 Moving Correlation

We correlated the residuals of three knowledge levels in a moving window of five observations such as [1-5], [2-6], [3-7], [4-8]...[38-42]. The moving correlation window visualizes the correlation between three vocabulary knowledge levels as functions of time. In Figure 14, the yellow line is the moving correlation between active recognition and active recall of the paper learner. It fluctuates dramatically in the first half period from high value to low value while that of the computer learners in Figure 15 and 16 fluctuates dramatically in the first half period from low value to high value and that in Figure 17 does not have much variation. Even though both paper learner and computer learners show negative deterministic correlation values, the trajectories of them differ from each other. The reason why the trajectory of the paper learner is from positive to negative and why that of the computer learners is more from negative to positive is that the growth of active recall of paper learner can not keep pace with the growth of active recognition over the time.
Figure 14. Paper learner: moving correlation between three knowledge levels

Figure 15. Computer learner 1: moving correlation between three knowledge levels

Figure 16. Computer learner 2: moving correlation between three knowledge levels

Figure 17. Computer learner 3: moving correlation between three knowledge levels
In Figure 14, the blue line is the moving correlation between active recall and controlled production of paper learner. It fluctuated dramatically from positive to negative in three weeks while that of the computer learners in Figure 15, 16, and 17 only fluctuated in the first half period from negative to positive and tended to have less variation in the second half period. It again indicated that the growth of controlled production of paper learner could not catch up with the growth of the active recall.

In Figure 14, the pink line is the moving correlation between active recognition and controlled production of paper learner mostly fluctuates within positive values while that of the computer learners in Figure 15 and 17 fluctuates violently from negative to positive. However, in Figure 16, the computer learner 2 presents the same symptom as the control group. The reason why the value of correlation of paper learner mostly stays in the positive area is that the growth of active recognition is the same as the growth of controlled production. The reason why the value of correlation of some computer learners grows from negative to positive is that the variability of both active recognition and controlled production is more violent and sometimes the controlled production grows faster than the active recognition either in the first half period or in the last half period. In Table 4, we summarize what was found from the moving correlation of the paper learner and the computer learners.

<table>
<thead>
<tr>
<th></th>
<th>Paper learner</th>
<th>Moving correlation of paper learner</th>
<th>Computer learner</th>
<th>Moving correlation of computer learners</th>
</tr>
</thead>
<tbody>
<tr>
<td>A R N &amp; A R L</td>
<td>Negative</td>
<td>positive to negative</td>
<td>Negative</td>
<td>negative to positive</td>
</tr>
<tr>
<td>A R L &amp; C P</td>
<td>Sometimes positive, Sometimes negative</td>
<td>positive to negative</td>
<td>Mostly positive</td>
<td>negative to positive</td>
</tr>
<tr>
<td>A R N &amp; C P</td>
<td>Positive</td>
<td>only within positive</td>
<td>Positive</td>
<td>negative to positive</td>
</tr>
</tbody>
</table>

Table 4. Moving correlation between active recognition (ARN), active recall (ARL) and controlled production (CP)
In the stage of data description and exploration, we explore both the general trend and learning rate of each variable and the interaction between variables fitting our assumption: the computer learners tend to fill in the gap between receptive and productive vocabulary knowledge. However, we still need to quantify the observations to interpret how the computer learners and the paper learner differ from each other. Fitting the collected data with the logistic model developed by Van Geert (1994) would give us a more clear idea of how the computer learners and the paper learner perform differently from each other.

5.3 Model Specification

Logistic growth model of vocabulary development (Van Geert, 1994) combines the proportional growth with decelerating effect due to the limited internal and external resources learners have. The vocabulary score varying with time can be described by:

\[ \frac{dL(t)}{dt} = r \cdot L(t) \]

where \( L \) represents the vocabulary score, \( t \) represents the time, \( r \) represents the rate of vocabulary learning.

The learning rate varying with time can be described by:

\[ \frac{dr(t)}{dt} = m \cdot L(t) \cdot r(t) \]

where \( m \) represents the decreasing motivation.

The solution to these two equations can be described by:

\[ L(t) = \frac{1}{1 + e^{-t}} \]

This logistic equation found by Verhulst (1838) is a continuous model of population growth in time. The graph of the logistic equation is shown in figure 18. When the value of \( r \) is lower than zero, the curve has a downward trend. When the value of \( r \) is zero, the curve has no change and remains parallel. When the value of \( r \) is higher than zero, the curve has a upward trend.
We intend to use the logistic equation to fit our collected data. We did not fit the logistic equation with raw data because the variability of the raw data is high and the fitting cannot be optimized when the data is with too much variability. In consequence, we smoothed the data by averaging every five data points so that there would not be too many random variables along the trend line to optimize the model fitting.

5.4 Model Fitting

To modify the curve of logistic model, there is a need to add parameters, which are $A_0$, $A_1$ and $A_2$. The modified logistic equation is:

$$L(t) = \frac{A_0}{1 + e^{-t(\frac{A_1}{A_2})}}$$

where $A_0$ represents carrying capacity, $A_1$ represents learning rate, and $A_2$ represents a value related to initial value. In figure 19, when we change the value of $A_0$ from 1 to 1.5, we will find...
the difference between the black line and the red line is the carrying capacity. When we change the value of A1 from 1 to 2, we will find the difference between the black line and the blue line is the learning rate. When we change the value of A2 from 3 to 2, we will find the difference between the black line and the green line is the initial value.

![Figure 19. modified logistic curve with three parameters](image)

We fit the smoothed data and the modified logistic equation for all learners as in figure 20 and obtained the value of carrying capacity as A0, learning rate as A1, initial value as and correlation coefficient between the smoothed data and the fitting model in table 5 and 6.

![Figure 20. Fitting model of all learners](image)
In figure 20 and table 5 and 6, the lowest vocabulary level (active recognition) of the paper learner and the computer learners shows a similar trend, learning rate and carrying capacity but a different initial value. The paper learner has lower score of initial value than two of the computer learners. This indicates that the computer learners tend to recognize the word form earlier than the paper learner. The middle vocabulary level (active recall) shows similar carrying capacity but different learning rate, trend and initial value. The computer learner one has higher learning rate and shows a steeper curve, which seems to reach the last half of the logistic model while others still remain in the first half of the logistic model. All computer learners have higher initial values.
than the paper learner, which again denotes that the computer learners are more likely to recall the word form earlier than the paper learner. The highest vocabulary level (controlled production) shows different results of carrying capacity, learning rate, initial value and trend. Two computer learners have higher carrying capacity than the paper learner and only computer learner one has higher learning rate than the paper learner and reaches the last half of the logistics model. But all the initial values of all the computer learners are higher than that of the paper learner. It implies that the computer learners can acquire the word use earlier than the paper learner since they also can link word form and word meaning earlier than the paper learner.
Chapter 6

Discussion

The aim of the study is to investigate whether the gap between receptive and productive vocabulary knowledge can be more possibly filled in by the computer learners, to explore whether the lower vocabulary knowledge can support higher vocabulary knowledge more strongly for the computer learners, and to account for the difference the computer learners have from the paper learner during the development of vocabulary. The dynamic system theory was applied to our collected data to observe the general trend, the slope and the interactions between all vocabulary knowledge levels over time. The logistic model was applied to interpret the predominant parameters of vocabulary development by model fitting. In the following section, we will present the answers to the questions by interpreting the results based on the dynamic system theory and the logistic model.

The first question was addressed by observing the general trend lines and the detrended data of the learners. There are three main findings. The paper learner was likely to grow three vocabulary levels with the same rate. Computer learners of higher initial value were likely to have the steepest trend line in controlled production. They could pay attention to controlled production in the early stage since they could memorize most of the words after learning them once on the computer. Computer learners of lower initial value tended to have the steepest trend line in active recognition. They could only focus on active recognition in the early stage and then on active recall and controlled production later on. Therefore, we could assume that the computer learners can benefit from the computer by picking up the words faster in their own starting level than the paper learner. This implication alone cannot sufficiently account for the gap, however.

Second, both paper learner and computer learners acquired vocabulary knowledge in the sequence of their levels from lowest to highest, from active recognition to active recall and to controlled production. The distinction between three levels was ranked in a reasonable sequence. Even though the learning rate varies, the sequence of picking up all vocabulary levels still remains the same.
Third, active recognition is a precursor of controlled production for both computer learners and paper learner. Both computer learners and paper learner, as presented in Table 3, might hold strong competition between active recognition and active recall and some mild competition or mild support between active recall and controlled production, and stronger support between active recognition and controlled production, which might denote that active recognition is a precursor of controlled production.

Combining these findings, we are able to draw the answer to our first research question. The gap between receptive and productive vocabulary knowledge is filled in by the computer learners faster than the paper learner, but the gap still exists among the computer learners. The computer learners pick up the vocabulary levels as the paper learner from active recognition (lowest level), which is the precursor of controlled production (highest level) and they can pick up the words faster. That is why we assume that they can fill in the gap faster than the paper learner.

The second research question was addressed by observing the moving correlation window between three vocabulary levels. There are two major findings deduced from the results of the moving correlation. As for the paper learner, the higher adjacent vocabulary knowledge levels usually cannot keep pace with the former lower adjacent vocabulary knowledge levels. We can assume that the gap between receptive and productive vocabulary knowledge cannot be quickly filled up by the paper learner while the gap of the computer learners can gradually be filled up in the last half period. This finding again proves our implication to first research question that the computer learners do have the gap but are able to fill in the gap faster than the paper learner, but it has not answered our second research question, yet.

Second, the growth of the highest vocabulary knowledge level (Controlled production) is the same as the growth of the lowest vocabulary knowledge level (Active recognition) for the paper learner while the growth of the highest vocabulary knowledge level is sometimes faster than the growth of the lowest vocabulary knowledge level for the computer learners. We can assume that the active recognition supports controlled production more strongly for the computer learners than for the paper learner, which implies that lower vocabulary level supports higher
vocabulary level more strongly by the computer learners and again denotes that the gap is more likely to be filled by the computer learners than by the paper learner.

The third research question is addressed by fitting the collected data with the logistic model and defining the predominant parameters including the initial value, the learning rate and the carrying capacity. Based on these parameters, the computer learners could acquire the word form, the word meaning and the word use earlier than the paper learners, but they do not necessarily perform better on the learning rate and the carrying capacity. The parameters and learning curves among the computer learners, however, are more diverse in higher vocabulary level, which implies the growth of the higher vocabulary level of the computer learners is of instability. The computer learners seem to benefit from the computer to fill in the gap from the beginning but still have difficulty gaining the highest vocabulary level.
Chapter 7
Conclusion

The present study is an attempt to test the effectiveness of the computer assisted vocabulary learning website: WordChamp. We evaluated WordChamp with vocabulary literature, determine whether WordChamp fills in the gap between receptive and productive vocabulary knowledge with DST, and interpret how WordChamp learners perform differently from the paper learner by the logistic model. We will compare the results of the empirical study with the theoretical evaluation to give suggestions to WordChamp learners or lecturers and to deduce some important elements of a computer assisted vocabulary learning (CAVL) website.

First, WordChamp matches the vocabulary acquisition process from semantization to consolidation and the empirical study shows that the computer learners acquire vocabulary knowledge from active recognition, active recall and controlled production as the paper learner. Learning with WordChamp does not change the sequence of acquiring the vocabulary knowledge. Learners still get acquainted with the vocabulary levels from lowest level (word form), middle level (word meaning) to the highest level (word use). Getting learners ready for the next higher level vocabulary learning is essential while designing a CAVL website. If it does not go from lower level to higher level, learners might stay in the same lower level for a long time without making any progress because there is no scaffold to reach the highest level.

Second, WordChamp elaborates various learning paths for consolidation with various drills. In the empirical study, all computer learners showed to be able to consolidate the word form, word meaning and word use earlier than the paper learner, and they all reported that the flashcards with audio were very helpful for memorizing vocabulary and the pronunciation drills helped them activate the sound of the vocabulary, which facilitated memorizing vocabulary again. The “pronunciation” or the “sound” indeed plays an important role while memorizing vocabulary. The function of the audio of the flashcards and the pronunciation drills should be stressed by the lecturers and should be taken into account while designing a CAVL website.
Third, WordChamp covers a wide range of receptive and productive vocabulary knowledge. In the empirical study, the computer learners fill in the gap between receptive and productive knowledge during the development but if we look at the carrying capacity, the computer learners only perform better than the paper learner in receptive vocabulary knowledge. The computer learners do not seem to benefit as much as we expected from WordChamp in productive vocabulary knowledge. The reasons why the computer learners did not perform better in productive knowledge might be that there is not much context-based vocabulary learning in WordChamp, that there is not sufficient time to notice any difference between the computer learners and the paper learner and that the functions of interaction with other learners and webreader are overlooked by our participants. Besides, the computer learners all reported that they felt their fluency of retrieving words was improved by doing the drills but they had difficulty editing the sample sentences in WordChamp. They either edited very short sentences due to limited knowledge of target words or longer sentences without knowing whether the word use was correct or not. The feedback of productive knowledge was insufficient for learners to improve word use. Lecturers should reinforce the accuracy and complexity of the productive vocabulary knowledge by offering more context-based input or interaction between learners. More meaning-focused input or output such as reading graded readers or communication activities with written output should be incorporated into a CAVL website to strengthen productive vocabulary knowledge.

Fourth, WordChamp does not offer knowledge of morphological forms and different meanings of a word since the flashcards were built up by learners themselves. It reflected on our tests because learners tended to learn only one meaning of target word. If they did not pick the same word meaning as the researcher had, they could not correctly answer answers to any vocabulary levels because other meanings never existed in their brain. This again reveals that meaning-focused input or output should be viewed as an important element of a CAVL website.

Last, WordChamp meets the criteria of CALL in perspective of second language acquisition but in the empirical study, one criterion, learner fit is not as complete as we expected in the background. One computer learner could not learn more than three words by himself with
WordChamp because he has difficulty pronouncing vocabulary and reading the sample sentences in on-line dictionary. Accordingly, learners who have lower proficiency or have a problem pronouncing words might not be able to acquire vocabulary by themselves even with the assistance of WordChamp. Plus, if the computer learners are not able to operate the computer with ease, they are reluctant to learn vocabulary with computer because they even have to use “pen” to “write down” the instructions of how to “log in” WordChamp. It is more convenient for them to use conventional vocabulary, which is learning on paper, to learn vocabulary. Accordingly, a CAVL website should consider the learners’ proficiency to give different support for vocabulary learning and the interface of it must be user-friendly to mitigate the difficulty learning vocabulary with assistance of computer.

For future research, more participants should be included to be tested for a longer period of time so that we are more likely to notice the difference between receptive and productive knowledge. Participants should be divided into two groups A and B. To improve the reliability of the effectiveness of the test, group A should be tested both on learning with computer and on learning with paper, and so should group B. The modified logistic model can be used as our basic model to test the effectiveness of vocabulary learning in a DST and can be added more parameters such as strong or weak interaction between each vocabulary knowledge level. What’s more, we can also test on each important element of CAVL website by designing the application of APPLE. Everyone in the world can use the vocabulary learning application. Then we will probably be able to know what is important when it comes to designing a good CAVL website.
Bibliography


Stockwell, G. (2007). A review of technology choice for teaching language skills and areas in the CALL literature. ReCALL, 19, 105-120

