GÖTEBORGS UNIVERSITET

The Effects of Participation Structure on Incidental and Intentional L2-English Vocabulary Acquisition During a Reading Task

Benjamin Barry

Ämneslärarprogrammet

Degree essay:	15 hp
Course:	L9EN2A
Level:	Advanced level / Second cycle
Term/year:	HT/2018
Supervisor:	Monika Mondor
Examiner:	Maria Olaussen
Code:	HT18-1160-001-L9EN2A

Keywords: Participation structure, collaboration, L2 vocabulary acquisition, incidental vocabulary learning, intentional vocabulary learning

Abstract

Participation structure can affect how language learners approach different aspects of a language, such as vocabulary acquisition. Collaboration is often regarded as being more beneficial than working alone, as it allows a learner to gain new ideas and knowledge from others. However, previous research has shown mixed results. Additionally, there appears to be a gap in knowledge of both participation structure and L2 vocabulary acquisition from a Swedish school context. This study investigates if participation structure (pairs and alone) can influence Swedish pupils in the acquisition of L2-English in the form of cockney rhyming slang in both incidental and intentional vocabulary learning. Participants were tested on their knowledge of the target expressions after two separate reading tasks, as well as two weeks after the initial testing day. Results found that there were no significant differences between working alone and in pairs in all three tests. Additionally, more vocabulary was acquired with intentional learning than incidental learning for both participation structures. Participants achieved their highest scores in the two-week delayed tests. One possible explanation is a lack of effective communication between the participants who worked in pairs during the reading tasks. Pedagogical implications for both collaborative tasks and L2-English vocabulary are discussed, and suggestions are made for further research with collaborative work.

Table of Contents

1 Introduction

Language learners use a variety of strategies and approaches when they encounter new, unfamiliar vocabulary (Hedge, 2000). While some language learners prefer to process and work with the vocabulary individually, others prefer to communicate and work with the vocabulary in groups. These preferences can also affect an individual's chances of success in vocabulary tasks.

When given a task, an individual will only have a set amount of knowledge about the subject of the task. If the individual lacks sufficient subject knowledge, they would seek help from an external source, for example, a teacher, fellow colleagues or a parent. This would allow the individual to increase their knowledge about the task. According to Vygotsky's socio-cultural theory, the individual would have reached the Zone of Proximal Development (ZPD), which is a knowledge level that can only be reached from interaction with others (Vygotsky, as cited in Mitchell & Myles, 2004). An individual who works alone reduces the number of external factors they can interact with, leading to less opportunities to reach the ZPD, hence less opportunities to increase the current knowledge level. In contrast, a group-work activity can give each member more opportunities to reach their ZPD through member interaction and communication, which are key skills in the learning process, especially in language learning (Cook, 2016). Each member can contribute with their own knowledge and experiences, which would allow different points of views and approaches to the task.

Vocabulary knowledge is an important aspect of any language and when new vocabulary knowledge has been gained by a learner, it is said that the learner has acquired the vocabulary (Cook, 2016). Different theories exist on how learners acquire second language (L2) vocabulary, and indeed language itself, including Krashen's input hypothesis and Swain's output hypothesis. The acquisition of L2 vocabulary can be affected by the choice of the vocabulary learning design. In incidental vocabulary learning, a language learner acquires new vocabulary without specific instruction or awareness of the vocabulary (Laufer, 2010). In contrast, a learner who has acquired new vocabulary with an awareness of specific target words or expressions has been given an exercise with an intentional learning design.

There have been mixed findings in previous research as to whether participation structure is beneficial to an L2 learner. While some research has found it beneficial for learners to work in pairs or small groups (Jones, Levin, Levin & Beitzel, 2000; Kim, 2008), other research has found hardly any differences between working alone or with someone else (Nassaji & Tian, 2010). Additionally, both incidental (Pellicer-Sánchez & Schmitt, 2010;

Reynolds 2015; Vidal 2011) and intentional (Elgort, 2011; Yamamoto, 2014) vocabulary learning has been found to be beneficial to L2 vocabulary acquisition, with intentional vocabulary learning giving greater gains (Barcroft, 2009; Laufer, 2010). While these research papers included participates from many countries with different L1s, for example, Barcroft (2009) investigated with L1-Spanish participates, there is a general lack of research with Swedish participants, especially from a Swedish school context. I wondered if the findings in these research papers could be applied to L1-Swedish pupils learning L2-English.

In Swedish schools, vocabulary teaching is often included in lesson plans for the subject of English. While not explicitly referring to vocabulary teaching, Skolverket (2018) states that pupils between the school years of seven and nine should be taught about word knowledge, such as fixed expressions and intonation (pg.37). Furthermore, helping pupils to develop communication skills is one the aims of the subject English in the Swedish school curriculum, where it states that pupils need to learn "the ability to use different strategies to support communication and solve problems when language skills by themselves are not sufficient" (Skolverket, 2018, pg.34). Therefore, the findings in this research project can help establish if findings on L2 English vocabulary acquisition can be applied to learners of L2 English in Sweden.

The purpose of this research project is to investigate if the organisation of Swedish pupils during teaching (either working individually or in pairs) effects different types of L2-English vocabulary acquisition during a reading task. In order to investigate this, the following research questions for this project were created:

1) Do pupils who work in pairs acquire more English vocabulary in the form of cockney rhyming slang than pupils who work individually in both incidental and intentional vocabulary learning?
2) Will test scores after intentional vocabulary learning be significantly better than test scores after incidental vocabulary learning for both sets of pupils?
3) Will any vocabulary knowledge gained, in the form of word recognition and meaning, be retained after a two-week period?

This research paper will first review previous research in greater detail, before explaining the methodology of the experiment for this study. A results section will then show the outcomes from the experiment, which will be followed by a discussions sections that will suggest possible reasons for the outcomes and possible limitations of the study. Finally, a conclusions sections will summarize the main findings and suggestions for further research.

2 Previous Research

The literature presented in this section has been split into two sections. This first reviews research that investigated participation structure. Vocabulary learning is the focus of the literature in the second section. For this section, only research that involved L2-English vocabulary acquisition was considered for the scope of this research project, even though there might be research that exists involving vocabulary with other L2s. Some of these research papers have been reviewed in Barry (2018), where a deeper review was given for recent research that has been done in L2-English vocabulary.

2.1 Participation structure

According to Swain and Lapkin (1998), collaboration helps to improve the knowledge of language learners. This statement is based on an investigation of pupils from a French immersion class in a collaboration task. Through communication and interaction, a learner can develop a deeper understanding of a topic with the help of support from someone else (also known as scaffolding) and feedback from a fellow learner or even an instructor (Cook, 2016). However, mixed results that have been reported in previous research on whether collaborative work is more effective for language learning than individual work.

One research paper that found that collaboration was more effective in the learning process was Jones et al. (2000). Their research investigated participation structure with two types of vocabulary-learning strategies; semantic-context and mnemonic. American sixth grade pupils (n = 100) were randomly assigned to one of six groups based on both the type of task and organisation of the pupils. The first three groups were assigned the mnemonic instruction condition, which meant the pupils were given tasks that involved association. The pupils in the first group would work alone in both a learning phase and a testing phase, the pupils in the second group would work in pairs for the learning phase but individually in the testing phase and the pupils in the third group would work in pairs for both phases. Groups four to six had the same participation structure as the first three groups, but were given the semantic-vocabulary condition, which involved tasks with synonyms and definitions.

The study investigated from a first language (L1) perspective, as the English-speaking American pupils were tested with 16 target words that were low frequency English nouns. After a learning phase, the pupils were given an immediate definitions post-test, which was followed by a delayed definitions post-test one week later.

Jones et al. (2000) found that pupils who worked individually, group one, were outperformed by pupils who worked in a pair during the learning phase, groups two and three, for the mnemonic strategies condition, although there was no significant difference between groups two and three. With the semantic strategies groups, group six (the pupils who worked in pairs during both phases) produced significantly better results than groups four and five. No significant difference was found between the mean scores of group four (individual participation in both phases) and group five, which meant that these conditions were just as effective as each other. As the pupils that had access to mnemonic strategies had performed better than those with access to semantic strategies, Jones et al. (2000) concluded that pair work is more effective in vocabulary learning than working alone, especially in more effective learning strategies.

The research of Jones et al. (2000) was found in an L1 English context but similar results have also been found in L2 contexts. One such study was Kim (2008), who investigated how participation structure affected L2-Korean learners' vocabulary abilities. Thirty-two participants, who came from six different countries, were split evenly into two groups, individual and collaborative, and worked with dictogloss activities over a period of one week. At the end of the week, the participants in both groups had to reconstruct a text in their given organisation condition from an audio piece, as well as answer a vocabulary test. Participants were recorded during the reconstruction task, where the members of the individual group were asked to state their thought process aloud. The participants were given a similar post-test two weeks after the initial test.

From the analysis of the recordings, Kim (2008) found that both groups had a similar total when identifying lexical Language-Related Episodes (LREs), which was made up of word meaning, pronunciation and spelling. However, in both the immediate and delayed tests, the collaborative group performed significantly better than the individual group. Kim (2008) concluded that the findings of the research agreed with other research papers that found collaboration to be a positive tool for vocabulary learning. This was despite a few limitations that Kim (2008) admitted to, such as too few participants.

Communication is a key component of group work in a language setting because learners have the opportunity to practice with their comprehensible output skills (Mitchell & Myles, 2004). Within a conversation, feedback might be elicited. While investigating the role of feedback on L2-English vocabulary acquisition with thirty Iranian pupils, Nowbakht and Shahnazari (2015) found that comprehensible output and feedback have a positive influence

on a pupil's ability to acquire vocabulary. Furthermore, the more feedback a learner receives during vocabulary learning, the greater the chance that the learner will acquire the vocabulary.

While a collaborative learning approach has been reported to be more beneficial to the learning process than an individual learning approach by some research, such as Kim (2008), other research papers have found different results. Nassaji and Tian (2010) investigated if the learning of English phrasal verbs by L2-English learners is affected by how they are organised. A total of twenty-six adult participants with six different L1s, aged between eighteen and thirty-two, took part in the investigation. During a two-week testing period, participants did cloze reconstruction and editing tasks both individually and collaboratively. Sixteen phrasal verbs were the target words that were the subject of both pre- and post-treatment tests.

Nassaji and Tian (2010) found that there were no significant differences between the mean collaborative scores and the individual mean scores for both the cloze and editing tasks, even though the mean collaborative scores were slightly higher. This meant that participation structure did not affect English phrasal verb learning. Several reasons were suggested for these results by Nassaji and Tian (2010), including lack of effective communication in the collaborative tasks and that phrasal verbs are a difficult part of the English language to learn for many L2-English learners.

Mixed results were found in a study by Storch (1999), who investigated if grammatical accuracy of eight L2-English learners with several different L1s could be influenced by how they were organised during tasks. Overall grammatical accuracy was found to be better in tasks such as cloze exercises and text reconstruction for collaborative organisation than individual. However, for certain parts of grammar, like selecting the correct article in a cloze exercise, the individuals outperformed the collaborative participants by quite a margin. Storch (1999) stated that the results did not include enough participants and may not be representative of reality.

Yi and Sun (2013) also found mixed results when investigating how L2-English vocabulary acquisition is affected by negotiation of meaning, which is the method used by a learner when trying to understand misunderstood parts of communication. With L1-Chinese high school pupils, aged between thirteen and fifteen years old, the pupils who were allowed to negotiate meaning during tasks performed significantly better in vocabulary tests than those who were not allowed. However, the same result was not found with college students aged between eighteen and twenty-one years old, as there was no significant difference between a negotiation of the meaning group and a control group. Although, it should be highlighted that

there was a design limitation in the test for the college participants, as there were only five items in a matching task. The high school pupils received twice as many items in their tests. The items in the college students' tests was too low, which increased the chances of answers being correct due to guessing (Hughes, 2003).

2.2 Vocabulary learning

While there has been a lot of research about L2-English vocabulary since the twentieth century, recent research from 2009 will be the focus for this research paper. The influence of the two types of vocabulary learning, incidental and intentional, on L2-English acquisition have been investigated both separately and comparatively.

According to Pellicer-Sánchez and Schmitt (2010), incidental vocabulary learning can help learners acquire vocabulary. In their study, L1-Spanish students (n = 20) were asked to read an English novel that contained the Nigerian language, Ibo, and some of the novel's Ibo vocabulary formed the target words. Post-reading interviews revealed that the participants had improved their target word knowledge in areas including word class and meaning. A similar outcome was found by Reynolds (2015), although two independent groups of participants were used, rather than the repeated measures design adopted by Pellicer-Sánchez and Schmitt (2010). For the part of the investigation that focused on L2-English vocabulary acquisition, the experimental and control groups in Reynolds (2015) consisted of L1-Mandarin university students (a total number of fifty-nine). The experimental group, who read a novel containing nonce words, which are words that do not exist in the English language, significantly outperformed the control group, who had no knowledge of the book, in a vocabulary test about the nonce words. Both Pellicer-Sánchez and Schmitt (2010) and Reynolds (2015) showed that knowledge of unfamiliar vocabulary could be acquired by an L2 learner during reading tasks, even when the specific vocabulary has not been highlighted during the learning phase (incidental learning).

Incidental vocabulary acquisition is not just possible with reading tasks but listening tasks too. In an investigation with L1-Spanish students (n = 248), Vidal (2011) found that the participants who were placed in a reading treatment group achieved the most improvements from pre-tests in vocabulary to vocabulary tests after a treatment phase. Participants placed in a listening treatment group also improved, but no significant difference was found with scores of participants placed in a control treatment group. Interestingly, results of delayed one-month

tests showed that the participants in the listening group retained more knowledge of the target words than those in the reading group.

Acquisition of L2-English vocabulary from incidental learning in listening tasks was also found by van Zeeland and Schmitt (2013). Unlike Vidal (2011), whose participants all had one L1, which was Spanish, van Zeeland and Schmitt (2013) included participants from seventeen different L1s, including Swedish. It was found that different aspects of vocabulary were acquired in different ways, for example, participants achieved their best scores in a form recognition test immediately after a treatment, but the amount retained significantly reduced after a two-week period. In contrast, immediate test scores for word meaning were lower than form recognition, but the meaning knowledge that was acquired was generally retained after a two-week period.

While incidental vocabulary learning has been found to positively influence the acquisition of L2-English vocabulary, research has also found this to be true for intentional vocabulary learning. In an investigation of forty-eight participants with different L1s, Elgort (2011) found that word form and meaning knowledge of selected pseudowords significantly improved after intentional learning sessions. Although the pseudowords do not exist in the English language, Elgort (2011) showed that it is potentially possible for L2-English learners to acquire new vocabulary through intentional vocabulary learning.

Yamamoto (2014) also found intentional learning to be beneficial to L2-English learners, as it could significantly improve the amount of vocabulary known by the learners. However, other knowledge aspects of vocabulary acquisition, for example, the meaning of a word or how to use a word in context, were found to be not affected by intentional learning. The L1-Japanese participants explained in post-test interviews that they could use memory strategies to remember new vocabulary, as these were commonly used for school exams. This might explain why Yamamoto (2014) obtained these results, as it is easier for most individuals to remember the form of a new word than being able to remember its meaning or to use it in the right context.

A comparison of incidental and intentional L2-English vocabulary learning was done by Barcroft (2009). The participants involved were L1-Spanish students (n = 114), who were given a reading task. Additionally, some of the participants were asked to translate target words from English to Spanish. Barcroft (2009) stated that intentional vocabulary learning was more beneficial in reading tasks than incidental vocabulary learning based on an analysis of two vocabulary tests given to the participants.

According to Laufer (2010), different types of word-focused instruction, which can be used in both incidental and intentional vocabulary learning, can be very beneficial to L2 vocabulary acquisition. One type of instruction is Focus on Form(FonF), where the meaning and context of the vocabulary in tasks is focused on. A common task for FonF in incidental vocabulary learning is to test learners on target words that they were not made aware of in a previous task. In contrast, target words would be known by learners in tasks with intentional vocabulary learning.

Another word-focused instruction is Focus on Forms (FonFs), where the teaching is based on rules and structures of the vocabulary, thus leading learners to learn about the language. Laufer (2010) informs that tasks with FonFs often use non-authentic language because the aim is to know about the structures of the word rather than how it can be used in different contexts. With incidental vocabulary learning, tests would focus on knowledge of the target words not known to the learners during tasks, for example, word class and form.

Laufer-Dvorkin (2006) investigated if there were any differences with L2-English learners between FonF and FonFs instruction in both incidental and intentional vocabulary learning. In the study, one hundred and fifty-eight Israeli participants that had an L1 of either Hebrew or Arabic were split equally into two groups; a FonF group and a FonFs group.

In the first phase of testing, the participants in both groups were given incidental vocabulary learning. The FonF group read a text and answered comprehension tests about the target words, whereas the FonFs were only given information about the target words in English. After the treatment phase, a test was given that was the same for both groups. The results of the test showed that a significantly better mean score was achieved by the FonFs group than the FonF group. In the second phase of testing, both groups received the intentional vocabulary learning treatment, and were given the exact same method of receiving a list of all the target words and information about them. Tests given after the intentional learning found that the scores achieved by both groups were not significant differently from one another. This same test was given again two weeks later to investigate vocabulary retention, yet there were still no differences between the two groups. When the mean scores of the three tests were compared, the scores after the intentional vocabulary learning were significantly better than the scores in the test after the incidental vocabulary learning for both groups. This complements the findings made by Barcroft (2009), where intentional vocabulary learning was found to be more beneficial than incidental. Furthermore, scores had decreased in the two-week delayed test, which held true for both groups.

Laufer-Dvorkin (2006) offered several suggestions for the findings. In the incidental phase, the FonFs group were given an opportunity to know the target words by having access to the translations, whereas the FonF lacked this awareness in their texts. For the results with the intentional vocabulary learning, Laufer-Dvorkin (2006) realised that the methodology had affected the results because both groups were given the same procedures and intentional learning is a type FonFs. This means that the FonF participants were not confined to their given treatment. These findings led to the suggestion that it is highly recommended that FonFs are used in language teaching as they have a major influence on vocabulary acquisition. Additionally, FonF can also lead to gains in vocabulary acquisition, but these gains are generally small.

3 Methodology

This research paper will use both quantitative and qualitative methods. According to McKay (2006), quantitative research analysis measures data that can answer specific questions. In this research paper, the quantitative data was the number of correct answers given in the tests. Qualitative research normally gives deeper knowledge and further insights into an area of interest, with examples including interviews and questionnaires (McKay, 2006). Recordings of the participants who worked in pairs will be the qualitative data in this study, as it could give me knowledge of how the participants approached the tasks.

3.1 Participants

The participants in this study were both male and female pupils who attend a Swedish lower-secondary public school. Initially, there was a total of sixty-one pupils that took part, but five pupils were not present on the final testing day, so their results were not included in the final analysis. The pupils who did complete all testing tasks (n = 56) came from three different classrooms; two classrooms of year seven pupils and a classroom of year eight pupils. This meant that their ages ranged from thirteen to fifteen years old, which was chosen for several reasons. One is that my university programme focuses on this age group and the results of this study could help me in my future career as a year 7-9 English language teacher. Another is that most of the participants had several years of experience in English language learning during their time in school. Furthermore, the participants had at least some degree of second language vocabulary acquisition skills.

As the two year seven classrooms were at a similar English level overall, based on previous English subject grades, both of them were assigned a working condition. One of the classrooms was assigned the alone condition, where they would work individually, whereas the other classroom was assigned the partner condition, where the pupils were placed in pairs. For time and convivence, most pupils worked with their neighbour, while a few opted to work with friends. Assigning each classroom with a specific working condition ensured that the groups could be physically separated due to lesson schedules. This eliminated the possibility that the discussions between the pairs of pupils could not be overheard by those working individually.

The pupils in year eight were also split between the working conditions. The pupils who were assigned the alone condition worked in a group room during the reading phases, while

accompanied by their class teacher. The other pupils who worked with a partner stayed in the main classroom. During the test phases, both sets of pupils worked in the main classroom.

The pupils in years seven and eight who were assigned the alone working condition were called group A. A total of twenty-five participants completed both testing days and were used for the final analysis. In contrast, group B was designated to the pupils who worked with a partner during the reading phases. Only thirty-one participants in group B were present on both testing days. Participants were only accounted for the group they were assigned to; school year was disregarded. As conversations during the experiment were to be recorded for those in group B, the guardians of each pupil received a consent form prior to the start of the experiment.

As this study is investigating vocabulary acquisition from an L2 English perceptive, consideration was made for any L1 English pupils. While none of the pupils themselves has English as their first language, one of the pupils has a parent from the United States. As the focus of the target words was on a specific type of British English, and the chance of these words being familiar to the pupil was small, the pupil's test scores were included in the analysis.

3.2 Materials

An original English text, that was written by me, included eight target expressions that were the focus of the investigation. To minimize the chance that the pupils recognised the target expressions, a type of slang called cockney rhyming slang was chosen. In cockney rhyming slang, which originates from London, a word is replaced with an expression, where the last word of the expression rhymes with the original word, for example, *apples and pears* is used for *stairs* (de Boinod, 2015). The cockney rhyming slang expressions that were used in the text were found on the website cockneyrhyminslang.co.uk ("Cockney Rhyming Slang: London's Famous Secret Language," 2018).

Each pupil received two versions of the same text. The first version was written with the incidental vocabulary learning condition, which means that the target expressions were not highlighted and there was no explanation of cockney rhyming slang in the instructions. This version of the text can be seen in Appendix A. In the second text, under the intentional vocabulary learning condition, cockney rhyming slang was explained and the target expressions were highlighted.

A set of questions were also given after each version of the text, as well as after a two-week period after the initial testing day. To ensure reliability was high, the same set of questions were given to the pupils on all three occasions, i.e. after reading the first and second texts, as well as two weeks after the initial testing day. For the purposes of reporting and analysing the results, each set of questions were given a specific name. *Question paper part 1* was the set of questions given after the incidental condition text, *Question paper part 2* was given after the intentional text and the two-week post-test was called *Question paper part 3*.

A maximum total of ten points could be achieved from a multiple-choice question and two short answer questions. The multiple-choice question was the first question and tested for expression recognition, where there were only two correct items amongst six. One point was given for each correct item selected. The second question tested expression meaning with the items selected in question 1. Short written answers were required, and one point was given for each correct answer with a maximum of two points. The third question also tested expression meaning and was split into six parts. Each part presented a passage from the text that contained a target expression. Pupils needed to write a short answer about the meaning, which elicited one point given per correct answer. The format of the question papers can be seen in Appendix B.

The questions were written in both English and Swedish, so that the participants could understand what was being asked of them. However, they were strongly recommended to give answers in English, but were not penalised if they gave correct answers in Swedish. As the questions were not measuring spelling ability, incorrectly spelt answers were not penalized. This improved the validity of the questions, as they only focused on vocabulary knowledge (McKay, 2006).

3.3 Testing procedure

A pilot test with eight pupils from a different Swedish school was done to find any potential design flaws and problems that had not been thought of while designing the method. Based on what occurred in the pilot test, it was decided that five minutes to read and analyse each text was a sufficient amount of time. The results of these pupils were not included in the reported results.

There was a total of two testing days, which lasted over a two-week period. A general overview of the testing procedure can be seen in figure 1. On the first testing day, the testing procedure started with a brief introduction of what was to happen, so that the participants

were prepared for what to expect and to relieve any anticipation. Participants in both groups received the first version of the text, which was written with the incidental vocabulary learning condition. The text was read aloud by me so that the participants could hear the pronunciation of the words in the text. This was done to reduce the possibility of participants being stuck on difficult words during the reading phase, as they were not allowed to ask me nor their teacher questions about words in the text. All participants were instructed to read and analyse the texts in their given group for five minutes. For participants in group A, this meant that they would read the text on their own, whereas the pupils in group B would read and discuss the text in pairs. Participants in group B who were willing to record their conversations did so with a recording app on their laptops, so that the communication process could be analysed later. These participants could discuss the task in either Swedish or English. After five minutes, the texts were taken away and each pupil was given *Question paper part 1*, which were answered individually. Participants in group B were separated and were not allowed to communicate during the test.

Directly after the conclusion of *Question paper part 1*, the participants in group B would regroup with their partners, and both groups received the second version of the text, written with the intentional condition. The instructions now included an explanation of what cockney rhyming slang was and two examples were given. Participants were asked to read and analyse the text in their assigned group during a 5-minute period. Discussions were once again recorded for participants in group B who were willing to do so. After the 5-minute period, the texts were collected, and each participant individually answered the questions in *Question paper part 2*. When this had been completed by all participants in the classroom, the meanings of the target words were revealed and discussed as one big group in the classroom.

Two weeks after the initial testing day, the second testing day occurred, where the participants were given *Question paper part 3* spontaneously. Like the of question papers, the third question paper was also answered individually. As *Question paper part 3* was measuring the amount of target word recognition and meaning retained after a two-week period, the participants did not have access to either version of the text, neither before nor during *Question paper part 3*. Two pupils in group A and three pupils in group B did not attend school on the final testing day, thus were unable to answer *Question paper part 3*. Therefore, their results in the first two question papers are not reported in this study.

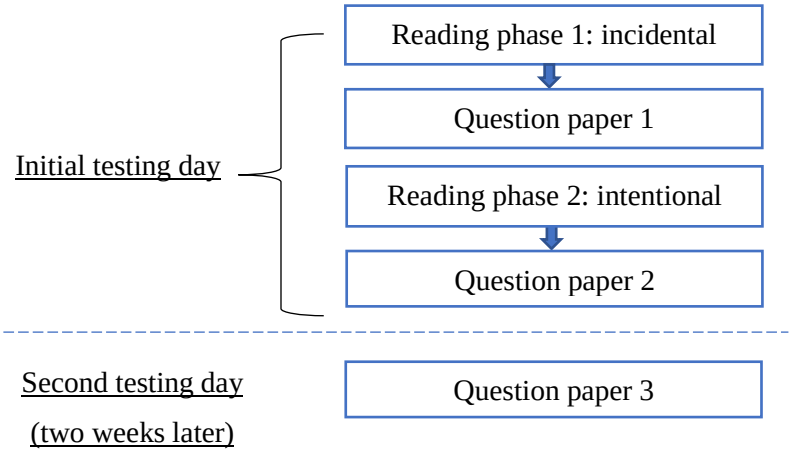

Figure 1: Testing procedure for both groups A and B.

3.4 Data analysis

Three values from each question paper were analysed. The first was the total score achieved, with the possibility of a participant achieving a maximum of 10 points. The second was the score achieved in question 1 (maximum two points), which asked a question about expression recognition. The third was a combination of the scores achieved in question 2 and 3 (maximum eight points), which asked about expression meaning.

Mean scores were found for the three values in each question paper for both groups. These mean scores were compared statistically using the statistical software SPSS to check if any differences were significant. Larson-Hall (2012) helped to inform me of the appropriate statistical tests that needed to be used and how to interpret and report the outcomes. The means scores were compared in the following ways:

1. Comparisons between both groups' mean scores of the three values in each question paper.
2. Comparisons between the mean scores for each of the three values in the three question papers within each group for both groups.

Recordings of the discussions between partners in group B were transcribed and analysed. This was done to see if there were any common approaches and attitudes to pair-work vocabulary exercises. Only a general summary of the discussions will be reported in this study.

3.5 Hypotheses

With regards to the first research question, which asked if there would be any differences in vocabulary acquisition between participants working in pairs and individually in both vocabulary learning conditions, I predicted that paired pupils would elicit better results than those who worked alone in both incidental and intentional learning conditions. This assumption is based on an individual being able to reach their ZPD through communication with someone else, which allows the individual to gain previous unknown knowledge (Vygotsky, as cited in Mitchell & Myles, 2004). The participants in groups A would not have the opportunity to discuss ideas with anyone, so they would be restricted to finding a solution to understanding the text on their own.

Additionally, I predicted that scores would be higher with the intentional vocabulary learning condition for both groups than incidental because the participants would have knowledge of the words that they will be questioned on. This relates to the second research question, which compares both types of vocabulary learning. In the incidental vocabulary learning condition, I predicted that most of the participants would struggle to understand the text due to not being familiar with cockney rhyming slang.

With reference to the third research question, which asked if vocabulary knowledge would be retained, I was unsure whether the scores would have improved, decreased or stay the same after a two-week period. The results would depend on how well each pupil had acquired the expressions.

4 Results

The scores from the three question papers were analysed with descriptive statistics and the results, in the form of mean and standard deviation (SD), are presented in table 1. All statistical outcomes from SPSS can be found in Appendix C. Participants only received scores that were whole numbers, which means that fractions were not given. However, the means and standard deviations have been reported to 2 decimal places to help give a deeper analysis. The maximum score a participant could have achieved in any of the question papers was ten points. This total score consisted of two scores from two different aspects of vocabulary, i.e. recognition and meaning. The maximum score achievable on a question paper for recognition was two points, whereas a maximum of eight points was possible for the meaning part of the question papers.

4.1 Comparisons between groups A and B for each question paper

In each question paper, the mean scores for total, recognition and meaning were higher for group A than group B (see table 1). Independent t-tests between groups A and B for each question paper were used to check if these differences were significant or not, and these results can be seen in table 2. At the 95% confidence level, group A only significantly performed better than group B for recognition in the first question paper. There were no significant differences found between the two groups' mean scores for total and meaning in the first question paper. Similarly, in the other two question papers, there were no significant differences between groups A and B for total, recognition or meaning. In summary, these results show that vocabulary learning acquisition was not generally affected by how the participates were organised, in terms or working individually or in pairs.

Table 1
Mean and standard deviation (SD) of question paper scores for groups A and B

Group	Question paper	Total		Recognition		Meaning	
		Mean	SD	Mean	SD	Mean	SD
A	1	3.16	1.86	1.76	0.52	1.40	1.76
	2	6.16	3.01	1.92	0.28	4.24	2.96
	3	7.68	2.78	1.84	0.37	5.76	2.54
B	1	2.45	1.59	1.42	0.50	1.03	1.38
	2	5.77	2.58	1.81	0.40	3.97	2.56
	3	6.45	3.04	1.65	0.61	4.81	2.80

Table 2
Independent t-tests between group A and B for all three question papers

Question paper	Score	T	df	P
1	Total	1.54	54	0.13*
	Recognition	2.48	54	0.02
	Meaning	0.88	54	0.38*
2	Total	0.52	54	0.61*
	Recognition	1.20	54	0.24*
	Meaning	0.37	54	0.71*
3	Total	1.56	54	0.13*
	Recognition	1.40	54	0.17*
	Meaning	1.32	54	0.19*

Note: t is the t-statistic, *df* is degrees of freedom and *p* is the p-value. The p-value is significant at the $p < 0.05$ level. An asterisk (*) means not significant.

4.2 Comparisons between the three question papers for both groups

4.2.1 Group A

The participants in group A achieved their best total score in the two-week delayed question paper, with a mean score of 7.68 (SD = 2.78) out of a maximum 10 points. In contrast, the mean total score of 3.16 (SD = 1.86) for the question paper after the incidental vocabulary learning reading session was the lowest (see table 1). A repeated measures ANOVA found that there was a significant difference between the mean total scores achieved in all three question papers, $F(2,48) = 65.56$, $p < 0.001$. Further post-hoc tests, in the form of paired-samples t-tests between the three possible combinations, found that the mean total score achieved in the third question paper was significantly better than the second, $t(24)= 4.38$, $p < 0.001$, and first question papers, $t(24) = 10.6.4$, $p < 0.001$. The participants also significantly improved from *Question paper part 1* to *Question paper part 2,* $t(24) = 7.01$, $p < 0.001$.

A similar pattern in mean total scores happened with the mean meaning scores, as the participants performed the best in the third question paper and the worst in the first. Like the mean total scores, a significant difference between the question papers was found with a repeated measures ANOVA, $F(2,48) = 65.46$, $p < 0.001$. Likewise, different combinations of paired-samples t-tests found that there were significant differences between the meaning scores in *Question paper part 3* and the scores in part 2, $t(24) = 4.44$, $p < 0.001$, as well as part 1, $t(24) = 11.40$, $p < 0.001$. The mean score after the intentional vocabulary reading phase

was significantly better than the mean score after the incidental phase for meaning, $t(24) = 6.59$, $p < 0.001$.

Mean scores for recognition showed a different pattern from the other two mean scores. Even though the best mean score was achieved in *Question paper part 2* ($M = 1.92$, $SD = 0.28$), analysis from a repeated measures ANOVA found that there was no significant difference between the three test papers for recognition, $F(2,48) = 2.09$, $p = 0.14$. Post-hoc tests found that only the mean recognition scores between the first and second question papers were significantly different, $t(24) = 2.14$, $p = 0.04$.

In summary, these outcomes show that the participants who worked individually performed better in a two-week delayed test than both immediate tests for general vocabulary and meaning. Those who worked individually significantly improved their recognition scores on average after intentional vocabulary learning compared to incidental vocabulary learning. However, there was no significant difference for recognition scores between scores achieved on an initial testing day and after a two-week period.

4.2.2 Group B

Table 1 shows that the mean total score in *Question paper part 3* was the highest amongst the three question papers and the lowest mean total score was obtained in *Question paper part 1*. A significant difference was found for the total scores by a repeated measures ANOVA, $F(2,60) = 75.76$, $p < 0.001$, and post-hoc test results revealed that there was a significant difference in each combination. This means that the mean total score achieved in the third question paper was significantly better than the mean total score in the second, $t(30) = 2.36$, $p = 0.03$, which in turn was better than the mean score of the first question paper, $t(30) = 9.72$, $p < 0.001$. Furthermore, the participants achieved a significantly better mean total score in the third question paper than the first, $t(30) = 9.89$, $p < 0.001$.

Like group A, the participants achieved their highest mean score for recognition in the second question paper. However, unlike group A, a repeated measures ANOVA found that there was a significant difference between the question papers' recognition scores, $F(2,60) = 8.45$, $p = 0.00$. Post-hoc tests showed that the mean recognition score of *Question paper part 1* was significantly worse than the scores in part 3, $t(30) = 2.24$, $p = 0.03$, and part 2, $t(30) = 4.35$, $p < 0.001$. However, a paired-samples t-test found no significant difference for the mean recognition scores between the second and third question papers, $t(30) = 1.72$, $p = 0.10$.

For the scores in meaning, a repeated measures ANOVA found that there was a significant difference between the question papers, $F(2,60) = 71.27$, $p < 0.001$. In follow up tests, it was found that the scores in the third question paper were significantly better than the first, $t(30) = 9.97$, $p < 0.01$, and second, $t(30) = 3.35$, $p = 0.00$. Furthermore, the participants achieved a significantly better score for meaning in the second question paper than the first, $t(30) = 8.32$, $p < 0.001$.

In conclusion, the pupils who worked together in pairs during the reading phases significantly achieved their worst results in total, as well as in target expression recognition and meaning, after incidental vocabulary learning. Additionally, these participants performed significantly better in total and for meaning two weeks after an initial testing day but achieved similar scores for recognition during the same period.

4.3 Group B conversations

During the incidental reading phase, the recorded conversations revealed that most of the participants were generally confused by the context of the text. An example of a this is when one participant asked their partner, *"What is happening?"*. Most of the confusion stemmed from the participant's interpretation of the target expressions. Instead of knowing that the target expressions represented another meaning, some participants misinterpreted them for their literal meaning, for example, in the passage *A strong smell of uncle Fred came from the kitchen*, some participants thought that *uncle Fred* was an actual person rather than the cockney rhyming slang for *bread*. In general, the pairings tried to support each other with questions and negotiating of meaning, such as *"what do you think that is?"*. However, this rarely helped when they were trying to understand the context of the text.

In contrast to the recordings during the incidental reading phase, most of the pupils seemed to be more motivated during the intentional reading phase. This could be due to the pupils had now understood how cockney rhyming slang works. During one conversation, where a pair of pupils were trying to make sense of the target expression *currant bun*, one of the pupils asked the other, *"what rhymes with bun? Fun?"*, to which the other pupil exclaimed, *"Sun! The sun had risen!"*. Even though the intentional text was written exactly the same as the incidental text except that the target expressions were in a bold font, the conversations after the second text mostly focused on the target expressions and the words surrounding them rather than whole sentences, or even the whole text.

5 Discussion

The main aim of this study was to investigate how both incidental and intentional L2-English vocabulary acquisition was affected by participation structure. To achieve this aim, three research questions were formed, which will now be discussed with the results obtained from the testing phases.

5.1 Research question one

The first research question of this study was to investigate whether participation structure affects vocabulary acquisition with both incidental and intentional vocabulary learning. The results obtained showed that, apart from meaning recognition after incidental learning, there was no statistical differences between participants who worked alone and those who worked with a partner. However, it should be noted that, for each score in each question paper, the participants who worked individually during the reading phases outperformed the participants who worked in pairs for each mean score comparison.

One reason that could explain why there was hardly any differences between the groups was the quality of the communication between the pairs in group B. The pairs might not have had the necessary skills to effectively convey their knowledge to each other. Furthermore, conversations between two participants might have led them down the wrong path. An example of this can be given during the first reading phase, where the pair might have encouraged each other to focus on a part of the text that dominated the conversation that might not have been the focus if they had worked individually. Connected to this, some participants might have felt shy to express their opinions in front of their partner in case they were wrong. Those who worked alone were free to work with the text without interference. Nassaji and Tian (2010) also suggested that underdeveloped communication skills was a reason why there was no significant difference in test performances with English phrasal verbs between individual and paired work.

Another possible reason for this outcome is the level of English proficiency amongst the participants. There could have been differences in English knowledge and abilities between the pupils in the three classrooms that made up the two groups, especially as the pupils from one of the classrooms were in year eight, while the other two were in year seven. Consideration was made before testing whether the participants should have taken a pre-vocabulary knowledge test to see if both groups were at a similar knowledge level. However,

it was ultimately decided that they would not do the pre-vocabulary tests due to time constrictions, as testing occurred during an intense teaching period in school, and convenience for both the pupils and the teacher, who participated during their normal school day. Furthermore, it was felt that the pre-tests in vocabulary would not be an indicator of how well the pupils would perform in this investigation with cockney rhyming slang. Group placement was decided by previous school year grades, where it was deemed that both groups were quite similar.

There could be an argument that that the groups were perhaps on a similar knowledge level due to there being no significant differences between the two. However, the analysis of the total scores achieved shows that the indvivdual participates achieved higher scores in all three question papers, but the paired participants had smaller standard deviations in both the first and second question papers (see table 1). According to McKay (2006), this means the scores of group A in the first and second tests had a greater dispersion than those in group B, where the participant scores were closer to the mean. Therefore, the knowledge gap between proficient English learners and less proficient learners could have been smaller with participants in group B than A. Even though the participants who worked individually achieved a significantly better score for recognition in the first paper than the paired participants, it would be difficult to apply this in a general context as there were only two items that were tested for recognition. Future research should investigate if there is a difference with a higher number of items tested.

5.2 Research question two

The second research question asked if better test scores would be achieved after intentional vocabulary learning than test scores after incidental vocabulary learning for both groups. Participants in both groups achieved significantly better total, recognition and meaning scores in the second question paper than the first. This outcome agrees with Barcroft (2009), who also found that intentional vocabulary learning was more effective than incidental vocabulary learning.

There are several reasons why there were significant differences between the two types of vocabulary learning. In the first reading phase, most of the participants did not realise what the focus of the accompanying follow-up question paper would be. Recordings from participants in group B revealed that these participants had one of three approaches during this phase. The first was to state that they did not understand either parts or the majority of the

text. The second was to focus on other parts in the text but not explicitly on the target expressions, for example, one pair spent some of their communication time discussing why the main character of the text, Toby, went to his window before leaving his bedroom. The final approach, which was the least common in the first reading phase, was to focus on the target words within the contexts of the sentence. In studies that focused on incidental vocabulary acquisition, both Pellicer-Sánchez and Schmitt (2010) and Reynolds (2015) found that learners used inferential knowledge from surrounding structures and words to give them clues about unfamiliar words. This is the same as the final approach made by some of the participants in this study.

In contrast to the first reading phase, the participants in the second reading phase were made aware of the target expressions and were given clues on how to find their meanings, although no actual meanings were given until after the second question paper was completed. Since all of the pupils were aware of the expressions, their approach to the text was narrowed down to finding how the target expressions were used in sentences rather than focusing on unrelated information within the text. Recordings from group B's conversations revealed that the pairs spent most of their time discussing words that rhymed with the target expressions and whether the rhymes would work within the text.

Another possible reason that scores were better with intentional vocabulary learning is test format familiarity. When the participants in both groups received *Question paper part 1*, it was the first time they had seen the questions and they did not know what to expect. The question paper after the second reading phase was their second time with these sets of questions, which meant that the participants were familiar with the test format. According to Hughes (2003), outcomes from a test that is given several times might be due to the test format being learnt from repeated exposure rather than an intended treatment. This could have been rectified with a crossover design, where half of the participants in both groups A and B would have been given intentional vocabulary learning conditions in the first reading and the incidental vocabulary learning condition in the second. However, pupils would have been made aware of the target words in the intentional vocabulary learning condition and it would have been impossible to follow up with incidental vocabulary learning, where the pupils should be unaware of what the target words are.

5.3 Research question three

The third research question asked whether the participants would retain vocabulary knowledge after a two-week period. The analysis of both groups' average question paper scores found mixed results. The pupils in group A, the alone group, significantly improved the mean scores for the total score and meaning in the third question paper than both the first and second question papers. This was the same outcomes for the participants that worked in pairs in group B. On the surface, this means that during this period, the participants gained more vocabulary knowledge. However, a deeper analysis might help explain these results in context. As with the difference between incidental and intentional vocabulary conditions, test familiarity could have been a factor, as the final question paper was the third time the pupils had encountered the questions.

Another factor could have been the revelation of the target meanings after the second question paper. Before the first question paper, the participants were given no clues about what they were reading and in the second question paper, they were merely given the pieces to the puzzle, not the solution. Thus, they knew that the target expressions rhymed with an English word but were not told the exact word needed, for example, the words *lairs*, *bears* and *stairs* all rhyme with *apples and pears* but in the context of the text, the only correct answer is *stairs*. By the third question paper, the participants had effectively been told the answers to the questions two weeks earlier at the end of the initial testing day. Additionally, cockney rhyming slang can be hard to master. The cockney rhyming slang expressions can be viewed as a puzzle that a learner must solve in order to understand it. Once the learner has cracked the code, it could be easier to guess unknown cockney rhyming expressions.

A third factor could have been the impression left on the participants. When entering the classrooms on the second testing day, some of the participants in the various groups who recognised me associated me with cockney rhyming slang. They then proceeded to use some of the cockney rhyming slang expressions with others in the classroom before the third question paper was even given. According to Hedge (2000), if learners have positive feelings towards an activity, then they are more likely to put in an effort. As cockney rhyming slang is part of an English-speaking country's culture, which is one of the core teaching contents in the Swedish curriculum (Skolverket, 2018), some of the participants may have enjoyed working with them from a cultural learning point of view rather than a vocabulary learning one.

One reason for the improved performance in the third question paper that is exclusive to group B was the use of listening skills, which was found by both Vidal (2011) and Van Zeeland and Schmitt (2013). Even though the tasks before the question papers were based on reading, the pair work allowed the group B participants to use their listening skills as well. This could have helped with remembering the target expression, as there were two sources of input, reading and listening, that the participants were receiving, which would increase the chance of retaining the information in the memory (Hedge, 2000).

The findings for recognition retention was different for both groups. The third question paper's average recognition score for the participants who worked individually was not significant different from both the first and second papers' mean scores. This means that there was retention because the third paper's recognition score did not significantly decrease. The participants in group B also did not achieve a significant difference between the second and third question paper, hence retention, but did significantly improve from scores achieved in the first question paper to the third. As noted previously, two items for testing vocabulary recognition knowledge is too small to see if there are any differences from a treatment. Even though a two-week post-test was given to see how much vocabulary was retained, these results give no indications if vocabulary can be retained long-term, especially if the pupils do not come across this type of input again.

5.4 Limitations

The findings of this investigation suggest that participation structure does not have an effect on neither incidental nor intentional L2-English vocabulary acquisition in a Swedish school context. However, more research is needed as this investigation was only done in one school in Sweden. There is a possibility that results may have been different if a variety of schools were involved in the study. Additionally, teachers in the Swedish classroom decide how much of a topic is present during a school year, for example, one teacher might incorporate a lot of vocabulary learning in their lesson plans while another might hardly focus specifically on vocabulary. It is uncertain if the results found would be similar for pupils who have a different teacher or attend a different school.

Furthermore, the results might only be applicable to the specific time-frame that the pupils did the tests, which was late November/ early December. Many of the participants felt stressed due to upcoming term grades and needed reassuring that performance s in this investigation would not contribute to their grades. The results in this study may have been

different if testing was done at other times of the school year, such as the start of a new term or near exam periods, when the pupils may have different attitudes to school and emotions. Additionally, the pupils may develop other techniques and skills in their ability to acquire English vocabulary over the course of their studies, especially as they progress through their education.

The choice of multiple-choice questions for recognition also needs to be considered. According to Hughes (2003), pupils can correctly answer a multiple-choice question by guessing. This means that the results found might not be indicative of their actual knowledge. To reduce the chance that guessing had an effect on the scores, the rest of the question papers had questions that required the participants to write their answers with words.

The paired participants had the opportunity to record their conversations, though few participants chose to do this, even though their parents and guardians had given permission. Additionally, even though the thought processes of the paired participants were recorded, nothing was documented about how the pupils who worked individually approached the tasks. Kim (2008) allowed participants who worked alone to record their thoughts aloud. Due to the way the participants in group B reacted to the recording opportunities, perhaps a written journal or log would have been more appropriate and effective for both groups.

6 Conclusions

Before this study, I wanted to know how different types of vocabulary learning could affect L2-English vocabulary acquisition in a Swedish context. Additionally, I wanted to find out if participate structure had an effect on both types of vocabulary acquisition. The findings of this study have shown that in most cases, there is no difference between learners working individually or with a partner in regards to acquisition of aspects of vocabulary, like L2 vocabulary meaning. However, the study has found that intentional vocabulary learning is more beneficial to acquiring vocabulary than incidental vocabulary learning. Furthermore, not only are learners likely to retain acquire vocabulary after a two-week period, but they can improve their knowledge by applying logical reasoning to understand a difficult language phenomenon, like cockney rhyming slang.

The findings in this study has helped me to consider many different approaches when my pupils come across new vocabulary. While some of them might prefer and acquire more vocabulary through interaction with a partner, others may prefer to learn the new vocabulary on their own and at their own pace. The ability to adapt teaching to the needs of the learners is a pedagogical implication that should be considered by all teachers. Another pedagogical implication is the consideration of more intentional vocabulary learning in the classroom. Outcomes in both this study and Barcroft (2009) show that more vocabulary is acquired with intentional vocabulary learning than incidental vocabulary learning, so it would be in the interest of language learners to include this type of vocabulary learning more in classroom activities.

In this study, only individual and pair work conditions were considered for the participation structure. Future research should investigate this further by adding a group work condition, where there would be three or more participants per group. This would be similar to a study by Dobao (2014), who investigated vocabulary learning with pairs and groups, due to a perceived perception that most previous research has focused on pair work. It was found that L2-Spanish learners did significantly better in vocabulary learning when organised in groups than in pairs. It would be interesting to see if the same outcome could be applied to this study.

Further research also should investigate if participation structure is affected by other types of vocabulary other than cockney rhyming slang. It can be a challenge for L2-English learners to understand how cockney rhyming slang works. This was found in this study, as participants during the first reading phase would often consider the meanings of the words

used in the cockney rhyming slang expressions rather than the actual meanings. The choice of less frequently used English vocabulary, idioms or set English expressions may have given different outcomes.

Reference list

Barcroft, J. (2009). Effects of synonym generation on incidental and intentional L2 vocabulary learning during reading. *TESOL Quarterly*, *43*(1), 79–103. https://doi.org/10.1002/j.1545-7249.2009.tb00228.x

Barry, B. (2018) *Acquisition of L2-English vocabulary: A literature review of research from 2009 to 2018* (Unpublished degree paper). Gothenburg: Department of languages and literatures, University of Gothenburg.

Cockney rhyming slang: London's famous secret language. (2018). Retrieved from http://www.cockneyrhymingslang.co.uk

Cook, V. J. (2016) *Second language learning and language teaching* (5th ed.). New York: Routledge.

De Boinod, A.J. (2015). Cockney. *Encyclopaedia Britannica*. Retrieved from https://www.britannica.com/topic/Cockney

Dobao, A. F. (2014). Vocabulary learning in collaborative tasks: A comparison of pair and small group work. *Language Teaching Research*, *18*(4), 497–520. https://doi.org/10.1177/1362168813519730

Elgort, I. (2011). Deliberate learning and vocabulary acquisition in a second language. *Language Learning*, *61*(2), 367–413. https://doi.org/10.1111/j.1467-9922.2010.00613.x

Hedge, T. (2000). *Teaching and learning in the language classroom: Oxford handbooks for language teachers.* Oxford: Oxford University Press.

Hughes, A. (2003) *Testing for language teachers* (2nd ed.). Cambridge: Cambridge University Press.

Jones, M. S., Levin, M. E., Levin, J. R.& Beitzel, B. D. (2000). Can vocabulary-learning strategies and pair-learning formats be profitably combined? *Journal of Educational Psychology*, *92*(2), 256–262. http://dx.doi.org.ezproxy.ub.gu.se/10.1037/0022-0663.92.2.256

Kim, Y. (2008). The Contribution of Collaborative and Individual Tasks to the Acquisition of L2 Vocabulary. *The Modern Language Journal*, *92*(1), 114–130. https://doi.org/10.1111/j.1540-4781.2008.00690.x

Larson-Hall, J. (2012). How to run statistical analyses. In A. Mackey & S. M. Gass (Eds.), *Research methods in second language acquisition: A practical guide* (pp. 245-274). Chichester: Wiley-Blackwell.

Laufer, B. (2010). Form-focused instruction in second language vocabulary learning. In R. Chacón-Beltrán, C. Abello-Contesse, & M. del Mar Torreblanca-López (Eds.), *Insights into non-native vocabulary teaching and learning* (pp. 15-27). Bristol; Multilingual Matters.

Laufer-Dvorkin, B. (2006). Comparing focus on form and focus on formS in second-language vocabulary learning. *The Canadian Modern Language Review / La Revue Canadienne Des Langues Vivantes*, *63*(1), 149–166. https://doi.org/10.1353/cml.2006.0047

McKay, S. L. (2006). *Researching second language classrooms.* Mahwah, N.J.: Lawrence Erlbaum Associates.

Mitchell, R. & Myles, F. (2004). *Second language learning theories* (2nd ed.). London: Arnold.

Nassaji, H., & Tian, J. (2010). Collaborative and individual output tasks and their effects on learning English phrasal verbs. *Language Teaching Research*, *14*(4), 397–419. https://doi.org/10.1177/1362168810375364

Nowbakht, M., & Shahnazari, M. (2015). The comparative effects of comprehensible input, output and corrective feedback on the receptive acquisition of L2 vocabulary items. *Advances in Language and Literary Studies*, *6*(4), 103–114. http://dx.doi.org/10.7575/aiac.alls.v.6n.4p.103

Pellicer-Sánchez, A., & Schmitt, N. (2010). Incidental vocabulary acquisition from an authentic novel: Do things fall apart? *Reading in a Foreign Language; Honolulu*, *22*(1), 31–55. Retrieved from https://search-proquest-com.ezproxy.ub.gu.se/docview/744447003?rfr_id=info%3Axri%2Fsid%3Aprimo

Reynolds, B. L. (2015). A mixed-methods approach to investigating first- and second-language incidental vocabulary acquisition through the reading of fiction. *Reading Research Quarterly*, *50*(1), 111–127. https://doi.org/10.1002/rrq.88

Skolverket. (2018). *Curriculum for the compulsory school, preschool class and school-age educare: Revised 2018*. Retrieved from https://www.skolverket.se/publikationer?id=3984

Storch, N. (1999). Are two heads better than one? Pair work and grammatical accuracy. *System*, *27*(3), 363–374. https://doi.org/10.1016/S0346-251X(99)00031-7

Swain, M., & Lapkin, S. (1998). Interaction and Second Language Learning: Two Adolescent French Immersion Students Working Together. *The Modern Language Journal*, *82*(3), 320–337. https://doi.org/10.2307/329959

Van Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System, 41*(3), 609–624. https://doi.org/10.1016/j.system.2013.07.012

Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning, 61*(1), 219–258. https://doi.org/10.1111/j.1467-9922.2010.00593.x

Yamamoto, Y. (2014). Multidimensional vocabulary acquisition through deliberate vocabulary list learning. *System, 42*, 232–243. https://doi.org/10.1016/j.system.2013.12.005

Yi, B., & Sun, Z. (2013). An empirical study of the effectiveness of negotiation of meaning in L2 vocabulary acquisition of Chinese learners of English. *English Language Teaching, 6*(10), 120-131. https://doi.org/10.5539/elt.v6n10p120

Appendix A

The original version of the text

A morning in the life of Toby Jones

At first, there was nothing special about the second Tuesday of September. The currant bun had risen, welcoming the start of a new day. Teenagers rode their bikes while delivering the daily wooden pews. Birds fought over uncle Fred on the ground, while children started their journeys to school by going down their apples and pears.

In his bedroom, Toby Jones was soundly asleep. Wrapped in a blanket, he looked like a bug in a cocoon. Suddenly, a loud noise woke Toby up. He jumped out of his bed and reached for his dog and bone. The sound was not coming from it and Toby stood there confused. He slowly realised that the sound was coming from outside the house. Panicking, he dropped his dog and bone, ran to the window and opened the curtains to have a butcher's hook. The glow from the currant bun hurt his eyes and he quickly closed the curtains again.

A voice started shouting Toby's name from one of the rooms downstairs. Toby knew it belonged to his brother, so he quickly ran to his bedroom door and opened it. He saw his brother's trouble and strife outside of the bathroom. "What's going on?" shouted Toby. Toby received no answer, so he ran down the apples and pears.

A strong smell of uncle Fred came from the kitchen, so Toby headed that way. He found his brother, John, dressed in a bowl of fruit. "What is going on?" asked Toby. "Toby, today's the day your life will change forever. Have a butcher's hook at this!" said John, pointing at the wooden pews on TV. They had won a competition and they were to receive their prize money at midday. Toby stood there shocked, then asked if John's trouble and strife knew about it. John nodded his head. Toby wondered if he needed his bowl of fruit.

Note: In terms of format and layout, this is exactly the what the participants received.

Appendix B

A sample of a question paper

Questions

Group B, part 1 **No.**

1. Which two of these following exact expressions were in the text?
   (Vilka två exakta uttryck [ordagrant] fanns i texten?):

- Apples and oranges
- Thunder and lightning
- Apples and pears
- Singing birds
- Lemons and limes
- Trouble and strife

2. What do the expressions mean? Give a meaning for each expression you chose in question 1.
   (Vad betyder uttrycken? Ge en betydelse för varje uttryck du valde i fråga 1).

   - Expression: Meaning:

   - Expression: Meaning:

3. Give a meaning for the expressions that are underlined.
   (Ge en betydelse för de uttryck som är understrukna).

- The glow from the currant bun hurt his eyes.

- Have a butcher's hook at this!

- Teenagers rode their bikes while delivering the daily wooden pews.

- A strong smell of uncle Fred came from the kitchen.

- Toby wondered if he needed his bowl of fruit.

- He jumped out of his bed and reached for his dog and bone.

Appendix C

Statistical outcomes

Repeated measures ANOVA between the question papers for group A's mean total scores						
Source		Type III Sum of Squares	Df	Mean Square	F	Sig.
Part	Sphericity Assumed	264.507	2	132.253	65.562	0.000
Error(Part)	Sphericity Assumed	96.827	48	2.017		

Paired-samples t-tests for Group A's mean total scores

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	A2Total - A1Total	3.000	2.141	.428	2.116	3.884	7.006	24	.000
Pair 2	A3Total - A1Total	4.520	2.124	.425	3.643	5.397	10.642	24	.000
Pair 3	A3Total - A2Total	1.520	1.735	.347	.804	2.236	4.381	24	.000

Repeated measures ANOVA between the question papers for group A's mean recognition scores						
Source		Type III Sum of Squares	Df	Mean Square	F	Sig.
Part	Sphericity Assumed	0.320	2	0.160	2.087	0.135
Error(Part)	Sphericity Assumed	3.680	48	0.077		

Paired-samples t-tests for group A's mean recognition scores

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1 A2Re - A1Re	.160	.374	.075	.006	.314	2.138	24	.043
Pair 2 A3Re - A1Re	.080	.493	.099	-.124	.284	.811	24	.425
Pair 3 A3Re - A2Re	-.080	.277	.055	-.194	.034	-1.445	24	.161

Repeated measures ANOVA between the question papers for group A's mean meaning scores

Source		Type III Sum of Squares	Df	Mean Square	F	Sig.
Part	Sphericity Assumed	244.880	2	122.440	65.456	0.000
Error(Part)	Sphericity Assumed	89.787	48	1.871		

Paired-samples t-tests for group A's mean meaning scores

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1 A2Meaning - A1Meaning	2.840	2.154	.431	1.951	3.729	6.592	24	.000
Pair 2 A3Meaning - A1Meaning	4.360	1.912	.382	3.571	5.149	11.400	24	.000
Pair 3 A3Meaning - A2Meaning	1.520	1.711	.342	.814	2.226	4.442	24	.000

Repeated measures ANOVA between the question papers for group B's mean total scores						
Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Part	Sphericity Assumed	284.151	2	142.075	75.763	0.000
Error(Part)	Sphericity Assumed	112.516	60	1.875		

Paired-samples t-tests for group B's mean total scores

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1 B2Total - B1Total	3.323	1.904	.342	2.624	4.021	9.715	30	.000
Pair 2 B3Total - B1Total	4.000	2.251	.404	3.174	4.826	9.894	30	.000
Pair 3 B3Total - B2Total	.677	1.600	.287	.091	1.264	2.358	30	.025

Repeated measures ANOVA between the question papers for group B's mean recognition scores						
Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Part	Sphericity Assumed	2.344	2	1.172	8.450	0.001
Error(Part)	Sphericity Assumed	8.323	60	0.139		

Paired-samples t-tests for group B's mean recognition scores

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1 B2Re - B1Re	.387	.495	.089	.205	.569	4.353	30	.000
Pair 2 B3Re - B1Re	.226	.560	.101	.020	.431	2.244	30	.032
Pair 3 B3Re - B2Re	-.161	.523	.094	-.353	.030	-1.718	30	.096

Repeated measures ANOVA between the question papers for group B's mean meaning scores

Source		Type III Sum of Squares	df	Mean Square	F	Sig.
Part	Sphericity Assumed	243.505	2	121.753	71.274	0.000
Error(Part)	Sphericity Assumed	102.495	60	1.708		

Paired-samples t-tests for group B's mean meaning scores

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1	B2Meaning - B1Meaning	2.935	1.965	.353	2.215	3.656	8.316	30	.000
Pair 2	B3Meaning - B1Meaning	3.774	2.109	.379	3.001	4.548	9.965	30	.000
Pair 3	B3Meaning - B2Meaning	.839	1.393	.250	.328	1.350	3.353	30	.002

Independent samples t-test between groups A and B								
		t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
							Lower	Upper
Total1	Equal variances assumed	1.536	54	0.130	0.708	0.461	-0.216	1.633
Recog1	Equal variances assumed	2.479	54	0.016	0.341	0.137	0.065	0.616
Meaning1	Equal variances assumed	0.878	54	0.384	0.368	0.419	-0.472	1.207
Total2	Equal variances assumed	0.517	54	0.608	0.386	0.747	-1.111	1.883
Recog2	Equal variances assumed	1.201	54	0.235	0.114	0.095	-0.076	0.303
Meaning2	Equal variances assumed	0.369	54	0.714	0.272	0.738	-1.208	1.753
Total3	Equal variances assumed	1.560	54	0.125	1.228	0.787	-0.350	2.807
Recog3	Equal variances assumed	1.401	54	0.167	0.195	0.139	-0.084	0.474
Meaning3	Equal variances assumed	1.321	54	0.192	0.954	0.722	-0.494	2.401