Automaticity in L2 learning: Correlation between vocabulary proficiency and response time in word recognition

Signe Jonsson
Automaticity in L2 learning: Correlation between vocabulary proficiency and response time in word recognition

Signe Jonsson

Abstract

Automaticity (in this essay defined as short response time) and fluency in language use are closely connected to each other and some research has been conducted regarding some of the aspects involved. In fact, the notion of automaticity is still debated and many definitions and opinions on what automaticity is have been suggested (Andersson, 1987, 1992, 1993, Logan, 1988, Segalowitz, 2010). One aspect that still needs more research is the correlation between vocabulary proficiency (a person’s knowledge about words and ability to use them correctly) and response time in word recognition. Therefore, the aim of this study has been to investigate this correlation using two different tests; one vocabulary size test (Paul Nation) and one lexical decision task (SuperLab) that measures both response time and accuracy. 23 Swedish students partaking in the English 7 course in upper secondary Swedish school were tested. The data were analyzed using a quantitative method where the average values and correlations from the test were used to compare the results. The correlations were calculated using Pearson’s Coefficient Correlations Calculator. The empirical study indicates that vocabulary proficiency is not strongly correlated with shorter response times in word recognition. Rather, the data indicate that L2 learners instead are sensitive to the frequency levels of the vocabulary. The accuracy (number of correct recognized words) and response times correlate with the frequency level of the tested words. This indicates that factors other than vocabulary proficiency are important for the ability to recognize words quickly.

Keywords

Vocabulary proficiency, Response time, Automaticity, English, Second Language Acquisition, Paul Nation, Frequency levels, Coefficient of Variation.
Contents

1. Introduction .............................................................................................................. 1
2. Background and previous research ........................................................................ 2
   2.1 Vocabulary proficiency and word recognition ............................................... 2
   2.2 Frequency levels ............................................................................................... 3
   2.3 Automaticity and the process of automatization ............................................. 4
3. Methodology ............................................................................................................. 5
   3.1 Participants ......................................................................................................... 6
   3.2 Data collection .................................................................................................... 6
   3.3 The questionnaire ............................................................................................... 6
   3.4 The vocabulary size test ..................................................................................... 6
   3.5 The word recognition test (LDT) ...................................................................... 7
   3.6 Analysis of the results ......................................................................................... 7
   3.7 Pearson’s correlation coefficient test ................................................................. 7
4. Results ...................................................................................................................... 8
5. Discussion ................................................................................................................ 11
   5.1 Implications and limitations of the study ......................................................... 13
6. Conclusion ................................................................................................................ 13
References ................................................................................................................... 15
Appendix A Questionnaire .......................................................................................... 16
Appendix B Information sheet ..................................................................................... 17
Appendix C Sample of vocabulary size test ............................................................... 18
1. Introduction

Learning a second language is a process involving many different factors. One aspect of language knowledge is vocabulary proficiency (a person’s knowledge about words and ability to use them correctly), which is also an important part of the automatization process (Dekeyser, 2001) for second language learners. For instance, previous research shows that in order to comprehend a text during reading, the reader needs to know approximately 95% of the words in the text (Nation, 2001) and if a person wants to be able to express herself/himself in another language she/he needs to know at least 2000 words (Schmitt, 2000). Furthermore, Schmitt (2000) claims that vocabulary knowledge also helps learners to see grammatical patterns and understand the meaning of the discourse. Thus, vocabulary proficiency is clearly helpful in several ways when learning a new language. Furthermore, vocabulary is acquired both explicitly and implicitly. Explicit vocabulary acquisition, which is to consciously learn vocabulary, occurs for instance in school, with the help of dictionaries. Implicit vocabulary acquisition occurs when a learner focuses on something other than learning vocabulary, for instance when reading, and understands new words from a context. Today, learners in Sweden are exposed to English vocabulary in almost all aspects of their lives – through TV, music, Internet, etc. This means that implicit vocabulary acquisition happens all the time. However, learners are probably mostly exposed to the more frequently used vocabulary (Nation, 2001). Therefore, it is possible to assume that their level of automaticity is higher in this spectrum and lower in the less frequent vocabulary spectrum.

Automaticity is the result of a process called automatization (DeKeyser, 2001). Automaticity is how fast a learner can access information or knowledge cognitively when speaking or reading. This information could, for instance, be vocabulary. DeKeyser’s (2001, p.125) description of the process emphasizes the complexity and the time aspect of achieving automaticity in a language:

Through a complex chain of mental operations, carried out in a fraction of a second, we can convert complex thoughts and feelings into soundwaves; and our interlocutor can convert them back into thoughts and feelings with the same amazing speed. Given the complexity of this skill and the speed with which it is used, it is not surprising that it takes years to acquire, and that learning a new language in adulthood is a slow and frustrating process.

In regards to this, it is of course interesting to investigate what helps the process of automatization and which constituents are the most important. Furthermore, it is a general conception that a higher level of automatization equals a higher level of fluency (Rodgers, 2011). Since vocabulary knowledge is clearly helpful when learning a language, an assumption would be that a high level of automatization should include a high level of vocabulary proficiency. Moreover, the main goal of language learning has always been for the learner to become as fluent as possible and, therefore, the abstract notion of automaticity have been the focus of many researchers. However, the notion of automaticity is still debated and many opinions and definitions of what automaticity is
and how it is best measured has been suggested (Andersson, 1987, 1992, 1993, Logan, 1988, Segalowitz, 2010). However, even though there has been some previous research regarding automaticity within second language acquisition (SLA) and English language teaching (ELT) (DeKeyser, 2001, Nation, 2001), more research is needed regarding the correlation between vocabulary proficiency and response time (level of automatization) for second language learners in Sweden. Also, a majority of the studies regarding automaticity has been focused on non-linguistic areas such as mathematics and computer programming (DeKeyser, 1996). More research within this field would be beneficial for the ELT community in Sweden and would provide a foundation for further research. Therefore, the main aim of this study is to investigate whether greater vocabulary proficiency and shorter response time (level of automaticity) in word recognition are correlated and to what extent. A second aim is to investigate whether frequency levels are important for word recognition. In this study, I have chosen to use a quantitative method to investigate the correlation between vocabulary proficiency, response time and accuracy in word recognition.

2. Background and previous research

2.1 Vocabulary proficiency and word recognition

There are different definitions of what vocabulary proficiency is and how it is best measured. One widely accepted way of defining a person’s vocabulary proficiency is to look at the size/breadth and depth of the vocabulary (Meara, 2005, Schmitt, 2010, Qian, 2002, Stæhr, 2009). The size is a measurement of how many words that a person knows and the depth is how much knowledge the person has about each word. Another aspect of this is also how fast the person can access this information (the level of automatization) (Meara, 2005).

Furthermore, according to Nation (2001), there are two ways a learner can know a word. Firstly, a person can have receptive knowledge about a word, which means that he/she can identify a heard or seen word, and recognize the parts that it is made up of. He/she can also understand the meaning based on the context that the word appears in and has knowledge about collocations and related words. Secondly, there is productive knowledge, which means that a learner not only has the receptive knowledge of a word but can write it down and pronounce it correctly as well as produce related words such as synonyms and antonyms. Moreover, learners who know a certain word are able to use it in the correct way, which involves knowing or recognizing the grammatical patterns the word functions in (Nation, 2001).

Word recognition can be measured in several ways and in this study I have chosen to use a lexical decision task (LDT) (Gough, 1984). In tasks like these, the participants are presented with a string of letters that either are real words or nonsense-words. The participants are then asked to identify the real words. The nonsense-words should be
pseudo-words; very word alike and pronounceable. The nonsense-words are then omitted in the analysis since they do not trigger word recognition.

2.2 Frequency levels

Different words are used with higher or lower frequency. Nation (2001) has divided words into frequency levels depending on how often they occur in various corpora. The two major categories are high-frequency words and low-frequency words. Two other categories are academic words and technical words. High frequency word are the 2000 most frequently used words in all kinds of uses of the language. Academic words are common in all kinds of academic texts and technical words are closely connected to the topic of the text. All other words are low-frequency words. This means that the low frequency category by far is the largest category of words and therefore, nearly impossible to learn. However, Nation (2001) also emphasizes that the boundary between the high-and-low frequencies is “arbitrary” and that the difference in frequency can be ever so slightly between a high-frequent and low-frequent word. For instance, since this method of division is strictly quantitative, a word can be just outside the high-frequency limit but still be a commonly used word.

The vocabulary size test used in this study measures the participants’ receptive knowledge (see appendix C for a sample of the test). There are 14 levels in the test and each level represents a frequency level and a set of word families. Each one of these levels consists of 10 words that each represent a word family within the frequency level. A word family can be described as a group of words that share the same root to which affixes can be added to form new words (Bauer & Nation, 1993). For instance, if the root is the word ‘man’ then possible affixes could be ‘ly’ or ‘un’ which would form the words ‘manly’ and ‘unmanly’. The test provides an estimation of the participants total vocabulary size. The most frequently used words are in the lower levels of the test, therefore, a decrease in test score is expected in the higher levels. However, there are only a limited number of words representing each level, which means that the test cannot be a reliable source for measuring the knowledge of each level. It is the total score of the test that estimates the vocabulary size. The first two levels represent the high-frequency word families. The third to ninth levels represent the mid-frequency word families and the tenth and above represent the less-frequent word families. The participants’ scores are to be multiplied by 100 to get an estimation of the vocabulary size. E.g. a test score of 40 would mean that a participant has an estimated vocabulary size of 4000 word families. According to Nation (2012), this would indicate that the participant knows about half of the word families necessary to understand the majority of text he/she encounters in his/her everyday life. A score of about 8000 word families would indicate that the person can understand the majority of text he/she encounters in his/her everyday life and a score of 9000 word families is necessary to understand most novels.
2.3 Automaticity and the process of automatization

Automaticity can be described as the end state from the process of automatization. There are many different views of what automaticity is and how it is measured, however, there are two major, widely accepted theories. Both theories come from cognitive psychology. The first theory is Anderson’s (1987, 1992, 1993 as cited in DeKeyser, 1996) ACT model (Adaptive Control of Thought of the human cognitive architecture). In this theory, a distinction is made between declarative/explicit memory structures, procedural memory structures and working memory. This model describes the process of automatization as knowledge starting out as explicit and declarative and through “specialized procedural rules” (DeKeyser, 1996, p.350) the knowledge becomes more and more refined. This is seen through a drop-off in response time and error rate which indicates that the knowledge has become more automatized. The ACT model focuses on rule-proceduralization (Rodgers, 2011).

The second major theory about automaticity is by Logan (1988 as cited in Rodgers, 2011). According to Logan (1988), automaticity is reached when a cognitive task has been performed enough times so that retrieving information from the memory is a faster process than applying a rule to solve a problem (Rodgers, 2011). Logan’s theory focuses on memory-retrieval (Rodgers, 2011).

Although both theories views automaticity as a decrease in cognitive processing, there are slight differences. In the ACT model, the process does not change but becomes more refined, whereas in the memory-retrieval theory the process changes when the user starts to omit applying rules and instead is accessing the information needed directly. One question that remains is of course which one of these theories that is most valid when it comes to studying automaticity. Rodgers (2011) suggests that a combination of the two theories is preferable, since neither rule-proceduralization nor memory-retrieval has been proved to work on their own according to previous research. However, they can be more or less involved depending on the knowledge that is being automatized. Another approach suggested by Segalowitz (2010, p.5) is a “cognitive science perspective”. According to Segalowitz (2010), automaticity can be divided into three main characteristics: processing speed, processing stability, and processing flexibility. Processing speed is of course important but not sufficient alone for automaticity. Instead, automatization involves being able to process information in a flexible way and rearrange parts of the cognitive task to improve the processing stability and efficiency. This means that Segalowitz (2010) views automaticity as an ever changing process rather than an end goal, which contradicts both the ACT model and the memory retrieval theory.

Furthermore, according to DeKeyser (1996) there is an agreement in research that “drop-offs in reaction time and error rates […] are reliable criteria of automaticity” (p.350). However, in recent years, there has been a change in the way automaticity is viewed. It has now become practice to focus on the process (the automatization) rather
than the end product (automaticity), since it is hard to define or know when someone has reached full automaticity (Rodgers, 2011). This is also supported by Segalowitz et al. (1998, 2003 as cited in Rodgers, 2011) who claim that most skills are somewhere on the scale between not automatized at all and fully automatized, thus being a blend between automatized and controlled processes (Rodgers, 2011). Furthermore, Segalowitz (2010) describes a process as automated if it has improved in some way: “A process is automatic, even if one cannot specify the exact nature of that automaticity, as long as one can say that the process is functioning more efficiently in some meaningful way compared to how it would have functioned if it were non-automatic” (Segalowitz, 2010, p. 79). This further supports Rodgers (2011) suggestion of using both theories about automatization in research but also to include Segalowitz’s theories.

Coefficient of variation (CV) is a way of measuring automaticity. CV measures intra-individual processing stability and speed (Segalowitz, 2010). CV is calculated by dividing the standard deviation by the mean response time (RT). The CV value indicates a participant’s processing stability; a smaller CV indicates a higher level of processing stability and a larger value indicates a lower level. According to Segalowitz (2010), only a qualitative change in processing can change the CV value because there is a linear correlation between the standard deviation and RT; if the RT change, the standard deviation will change proportionally. Thus, if there is only a speed up in processing, there will be no change in the CV. For instance, if there has been changes in the processing of a linguistic task (e.g. such as cognitive steps being skipped like translating a word before using it), the standard deviation would exceed the reduction in response time and, therefore, the CV will decrease and the process can be viewed as more automated than before. However, if the process would only become faster but the same cognitive steps are used then the CV will stay the same.

It is important to note that “automatization displays skill specificity” (Rodgers, 2011, p.297), which means that automatization within one area does not necessarily mean high level of automaticity in another. For instance, the level of automatization regarding comprehension does not transfer to the productive aspect of the language like writing.

3. Methodology

In this study, I have chosen to use a quantitative method to investigate the correlation between vocabulary proficiency, response time and accuracy in word recognition. However, this study is explorative in that the relationships between recognition speed and accuracy and general vocabulary knowledge is not well understood. A quantitative method was chosen because the average values generated by the different tests provided a good foundation for a statistical analysis.
3.1 Participants
23 participants were given both the vocabulary levels test and the lexical decision task. They were all between the age of 17-19 and students at an upper-secondary school in Sweden. In order to have the group as homogenous as possible an English 7 group was selected. English 7 is a voluntary course which means that students who are partaking in this course usually have an interest in doing so. It also means that they have passed English 5 and 6, which indicates that they should be approximately at the same proficiency level.

3.2 Data collection
The data was collected in three steps. First, the participants filled in a questionnaire (appendix A). Second, the participants took a vocabulary size test (sample in appendix C). Third, the participants took a computerized word recognition test (LDT). Before starting the test all participants were informed of the study and the different parts of the tests orally. They were also provided with an information sheet (appendix B). No personal information was collected that could connect any participant to a specific questionnaire or test result to ensure full confidentiality. Since no personal information was collected, consent forms were not needed. All three parts (questionnaire, vocabulary size test and word recognition test) were marked with a participant code (1-23). Codes were chosen depending on the order the participants did the LDT. The first participant was given the number 1, the second participants was given code 2 etc.

3.3 The questionnaire
The participants filled in a questionnaire about their English knowledge and background. The questionnaire was used primarily to ensure that no participant has English as their first language or has lived in an English speaking country for a longer period of time. That would have made the results from the study less reliable. They were also asked questions to determine their own perception of their proficiency and if they are learning vocabulary explicitly. Also, the questionnaire further ensured the homogeneity of the tested group. The questions was chosen in regards to the research questions. The questionnaire was piloted on a group of people to ensure that the questionnaire functioned as intended and that no changes was needed. Questionnaire available in appendix A.

3.4 The vocabulary size test
I have used Paul Nation's vocabulary size test in this study (retrieved at http://www.victoria.ac.nz/lals/about/staff/paul-nation). This test was used to estimate the average, receptive vocabulary size of the participants. The test is based on the British National Corpus word family list and has been piloted in several ways: I) Applied linguistics with English as their mother tongue has critiqued the test. II) The answers were replaced with nonsense words and checked by native speakers to check if
the choices themselves were indicating the correct answer. III) The test was checked through range programs to test the frequency levels. IIII) A Rasch-based analysis was conducted with 200 non-English speaking students. Therefore, the test is considered to be a very reliable source for estimation of vocabulary size.

3.5 The word recognition test (LDT)

The second test was a response time test where the participants saw a word on the computer screen and had to decide whether the word was English or not using software called SuperLab. This program measures the response times for the different words and the accuracy of identifying English words.

All participants were given the same instruction before starting the test; first a brief instruction when they were informed of the requirements of the study before the study started, and then once again individually before they started the computerized test. The instruction were as follows: This is a word recognition test. You will see words on the screen and you have to decide if the word is English or not. If you believe that the word is English, press the left mouse button. If you do not believe that the word is English, press the right mouse button. It is equally important to be accurate as it is to be fast. Do you understand everything? Press + when you are ready and the test will start.

3.6 Analysis of the results

In order to analyze the results, the nonsense words were removed from the word recognition test so that measures from only the real words were left. Nonsense words do not trigger word recognition. The number of correct recognized real words were used as the accuracy value for each participant. Subsequently, the average vocabulary test score, response time and accuracy for the entire tested group was calculated. Next, the real words were divided into three frequency levels according to Nation’s (2001) division; highly frequent words, mid-frequent words and low-frequent words. Within these three groups, the average response time and average accuracy were calculated. Subsequently, the correlation between the vocabulary test scores and response times, the correlation between vocabulary test scores and accuracy and the correlation between response times and accuracy were calculated using Microsoft’s Excel program. In addition, the same correlations were done with the “all words”-group. This group functions to see whether or not frequency levels are relevant. The correlations were calculated using Pearson’s correlation coefficient test. The median of the vocabulary size test was used to divide the participants into a ‘good’ group and ‘not-so-good’ group. The groups’ results were compared to the low-frequency and mid-frequency levels.

3.7 Pearson’s correlation coefficient test

The Pearson’s correlation coefficient was used in this study to determine the correlation between the observed data values. This test generates a result on a scale that goes from -1 to 1. A value close or on 0 indicates that there is no statistical significant correlation.
A value close to or on -1 indicates a negative correlation. A value close to or on 1 indicates a positive correlation. The correlations are calculated using this formula:

\[ r = \frac{n(\Sigma xy) - (\Sigma x)(\Sigma y)}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2][n\Sigma y^2 - (\Sigma y)^2]}} \]

4. Results

The results show that there is a weak to moderate correlation between vocabulary proficiency and LDT response time. However, when the frequency levels of the LDT words was taken into consideration, the strongest correlation was with the mid-frequency words, both for vocabulary proficiency and LDT response time on the one hand, and between response time and accuracy on the other. This suggests that L2 learners are sensitive to frequency levels of words when responding to the LDT, but, unexpectedly, not for the high frequency words (see Table 1 and 2 below).

Table 1. Correlations between vocabulary test scores and response time and vocabulary test scores and accuracy in different frequency levels.

<table>
<thead>
<tr>
<th>Frequency levels</th>
<th>High frequency</th>
<th>Mid-frequency</th>
<th>Low frequency</th>
<th>All Frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response time and vocabulary test score</td>
<td>-0.184</td>
<td>-0.343</td>
<td>0.0165</td>
<td>-0.181</td>
</tr>
<tr>
<td>Accuracy and vocabulary test score</td>
<td>0.272</td>
<td>0.553</td>
<td>0.174</td>
<td>0.435</td>
</tr>
</tbody>
</table>

Table 2. Correlations between accuracy and response time in different frequency levels.

<table>
<thead>
<tr>
<th>Frequency levels</th>
<th>High frequency</th>
<th>Mid-frequency</th>
<th>Low frequency</th>
<th>All Frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy and response time</td>
<td>-0.244</td>
<td>-0.517</td>
<td>-0.218</td>
<td>-0.457</td>
</tr>
</tbody>
</table>

Furthermore, even when the participants were divided into a ‘good’ and ‘not-so-good group’ (depending on their vocabulary test score) the correlations to vocabulary proficiency were weak. The expectations would be that the good group would have
stronger correlations, at least for the low frequency words. However, there is only a slight difference in correlation between the groups (see table 3 and 4 below).

Table 3. Correlations in the ‘good’ vocabulary score group at different frequency levels.

<table>
<thead>
<tr>
<th>Correlations between:</th>
<th>All frequencies</th>
<th>Mid-frequency</th>
<th>Low frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary test score and accuracy</td>
<td>0.112</td>
<td>0.226</td>
<td>0.042</td>
</tr>
<tr>
<td>Vocabulary test score and response time</td>
<td>0.084</td>
<td>-0.248</td>
<td>0.179</td>
</tr>
</tbody>
</table>

Table 4. Correlations in the ‘not-so-good’ vocabulary score group at different frequency levels.

<table>
<thead>
<tr>
<th>Correlations between:</th>
<th>All frequencies</th>
<th>Mid-frequency</th>
<th>Low frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocabulary test score and accuracy</td>
<td>-0.081</td>
<td>0.118</td>
<td>-0.248</td>
</tr>
<tr>
<td>Vocabulary test score and response time</td>
<td>0.085</td>
<td>0.046</td>
<td>0.225</td>
</tr>
</tbody>
</table>

Furthermore, the CVs in the groups do not support the idea that the good group should be more automated than the not-so-good group. A lower CV would indicate a higher level of automaticity. However, the average CV in the not-so-good group was lower than the average CV in the good group. Nevertheless, as expected, both groups had a higher CV in the low frequency word range compared to the mid-frequency range and all words combined (see table 5 and 6 below). The CV should decrease with either a smaller standard deviation or a larger RT (for the same standard deviation), but if the RTs are roughly the same, then the smaller standard deviation is what suggests greater automaticity. Therefore, since the good group overall has a lower standard deviation they could be considered more automated than the not-so-good group.
Table 5. Average coefficient of variation (CV), average response time and average standard deviation in the ‘good’ vocabulary score group at different frequency levels.

<table>
<thead>
<tr>
<th>Frequency levels</th>
<th>All Frequencies</th>
<th>Mid-frequency</th>
<th>Low frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient of variation (CV)</td>
<td>0,542</td>
<td>0,361</td>
<td>0,499</td>
</tr>
<tr>
<td>Response time</td>
<td>1095,017</td>
<td>1026,993</td>
<td>1388,093</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>652,854</td>
<td>419,910</td>
<td>782,253</td>
</tr>
</tbody>
</table>

Table 6. Average coefficient of variation (CV), average response time and average standard deviation in the ‘not-so-good’ vocabulary score group at different frequency levels.

<table>
<thead>
<tr>
<th>Frequency levels</th>
<th>All Frequencies</th>
<th>Mid-frequency</th>
<th>Low frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coefficient of variation (CV)</td>
<td>0,045</td>
<td>0,045</td>
<td>0,038</td>
</tr>
<tr>
<td>Response time</td>
<td>1272,674</td>
<td>1310,767</td>
<td>1507,822</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>738,720</td>
<td>654,476</td>
<td>778,095</td>
</tr>
</tbody>
</table>

The response times, as well as the accuracy for the lexical decision task, seem to be dependent on the frequency level of the tested word. The response times increased and the accuracy decreased at the mid-frequency and low-frequency levels. The response times were lower (faster) and the accuracy better at the high-frequency levels. Table 7 illustrates the differences between frequency levels:
Table 7. Accuracy and response time at different frequency levels for all participants.

<table>
<thead>
<tr>
<th>Frequency levels</th>
<th>High frequency</th>
<th>Mid-frequency</th>
<th>Low frequency</th>
<th>All Frequencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average response time (milliseconds)</td>
<td>916.7</td>
<td>1138</td>
<td>1439</td>
<td>1164.53</td>
</tr>
<tr>
<td>Average accuracy</td>
<td>9.565 (95%)</td>
<td>8 (80%)</td>
<td>6.174 (62%)</td>
<td>23.74 (79%)</td>
</tr>
</tbody>
</table>

As table 7 shows, there is a fairly large difference in accuracy between the low frequency word level and the high frequency word level (ca. 54% of the low frequency average). There is also a difference in the response time of 522.3 milliseconds between the low and high frequency levels (ca. 57% of the high frequency average). The mid frequency level roughly corresponds to the results from all words combined. Generally, the higher the frequency, the faster and more accurate the test scores.

5. Discussion

Overall, that there is only a moderate correlation between vocabulary proficiency and accuracy on the LDT is not surprising. The quite weak correlation between vocabulary proficiency and response time is, however, unexpected. The moderate correlation between accuracy and response time on the LDT is also not unexpected. The weak correlation between vocabulary proficiency and response time could be due to other factors than automaticity in vocabulary access, such as test taking strategy (different priorities for accuracy and speed among the informants). This could be connected to the fact that automaticity is skill-specific and does not transfer to other skills (DeKeyser, 2001), e.g. lexical access during context-based reading may not transfer to the uncontextualized LDT. Additionally, it may be that some informants may find LDTs somewhat stressful and consequently perform erratically. Faster response times are usually accepted as a measure of a higher level of automaticity (DeKeyser, 2001), and higher levels of automaticity are gained through exposure time (reading and listening). Therefore, the expectation would be that there should be at least a moderate correlation, if not a strong one, between response time and vocabulary proficiency.

Furthermore, in a comparison between the vocabulary test scores and the response times (in all frequency level groups) the data shows that there is only a slight difference between the participants who had better vocabulary scores (on or above the average score of 93) compared to the participants who had a lower vocabulary test score. 58% of the participants with a higher vocabulary test score also had shorter response times than average (on or below the average value of 1164.53). The corresponding value for the
participants with lower vocabulary test scores was 46%, which is not a major difference in such a small group of participants. However, a comparison between the good and the not-so-good group in the low-frequency and mid-frequency shows that the average vocabulary test score in the good group was significantly better than in the not-so-good group: 102.6 (73%) in the good group and 78.55 (56%) in the not-so-good group, based on the median 88.5. The response times were also faster in the good group in both the low frequency and the mid-frequency, which would indicate a correlation. Interestingly, the CVs are lower in the not-so-good group, which according to Segalowitz (2010) would indicate that the participants in that group were more automated than the participants in the good group. However, the difference in RT between the groups is not major but the standard deviation in the good group is lower in the mid-frequency and in the all words group, which could explain the CVs. The CV should decrease with either a smaller standard deviation or a larger RT, but if the RTs are roughly the same, then the smaller standard deviation is what suggest greater automaticity. Therefore, since the good group overall has lower standard deviation they could be considered more automated than the other group. The standard deviation is roughly the same in the low frequency, which is not unexpected.

Also connected to automaticity is the correlation between vocabulary test score and accuracy. The data indicate a moderate correlation across all frequency groups, but weaker in the low and high frequency levels. The expectation is that a broad and deep vocabulary should enable a reader, in theory, to be better (faster) at recognizing known words, especially if we apply Nation’s (2001) definition of knowing a word. However, the data collected in this study do not seem to support this assumption. Continuing on using Nation’s (2001) definition of knowing a word, which includes several aspects, one plausible reason for this result could be that the participants do not know enough words, or the right words, to be fully successful in this word recognition test. This could be due to several reasons, such as their age (they are young and cannot be expected to know the full meaning of the words required for this test), luck (the participants are highly proficient in vocabulary but no one can know all words and they did not know the words on the test) or they do not understand the grammatical patterns where these words could appear. Finally, the participants may be lacking in their receptive or productive aspects of knowing a word. This must remain speculation, however. We cannot know what goes on inside someone else’s head – we can only ask them to perform a task and then analyze the data (Gough, 1984, p.226). Clearly, other factors than word knowledge and automaticity must be involved in order to explain why the correlations are weaker in the low and high frequency levels than in the mid-frequency range.

The collected data do indicate, however, that L2 learners are sensitive to frequency levels. As expected, the results are better in the high frequency group. Both the accuracy and response times are better in this group and as the frequency level decreases so does the results from the tests. The words represented in the high frequency group are the
words that the participants are expected to encounter in their everyday lives the most. Moreover, Nation (2001) suggests that these should be the words most taught in school. This is probably the reason for the result. However, even though the data does not support a strong correlation between vocabulary proficiency and accuracy, participants who performed better at the vocabulary size test also had better accuracy in the LDT low frequency level.

5.1 Implications and limitations of the study

The size of the tested group should be discussed. 23 participants were tested, which could be viewed as a rather small group of participants. This could affect the outcome of the data. A larger group or a different group may result in different data. However, the method for collecting and analyzing the data is very time consuming, which made a larger group impossible for this study. Also, the same group cannot be tested more than once since this would mean that the participants would know how the test is constructed and which words it contains. Therefore, the data would not be reliable if the students took the vocabulary size test several times, since it would no longer represent the knowledge of the word families if the students only know the words from the test. However, I suggest that more research should be carried out in larger groups to verify the results. I would also suggest more research regarding the correlation between accuracy and vocabulary proficiency.

There is reason to question the results of the data since automaticity should be seen as a continuum rather than a black and white dichotomy. Students can be on different levels even in an as homogenous group as the one I have examined. The students can also perceive their proficiency levels/automaticity levels differently. However, the data from this study can be considered to be correct with even with these minor differences. Especially since the students are partaking in the English 7, which is a voluntary course for more accomplished students.

Another thing to consider is the motivation level of the students. Even if they are voluntarily participating in the study, some students might not take the test as seriously as would be desirable. However, the results from the vocabulary size tests indicates that all participants did take it seriously and Nation (2012) claims that we have to assume that the test is taken seriously and that we assume that the participants do not guess.

6. Conclusion

The aim of this study has been to investigate the correlation between vocabulary proficiency and response time in word recognition (lexical decision task). The collected data shows that there is not a strong correlation between vocabulary proficiency and response time. However, there is a moderate correlation between vocabulary test scores and accuracy and between accuracy and response times. Furthermore, the correlations
were weaker in the lower and higher frequency levels than for all the words as a whole. This indicates that even though there are moderate correlations, other aspects might be more important and that L2 learners are sensitive to frequency levels. More research regarding frequency levels in L2 learning would have to be conducted in order to verify the results.
References


Appendix A Questionnaire

Questionnaire

Automaticity in L2 learning: Correlation between vocabulary proficiency and response time in word recognition.

Researcher: Signe Jonsson
Stockholm University

Age: __________

Gender: Female Male Do not wish to answer

What is your first language?
________________________________________________________

What is your second language?
________________________________________________________

What is your third language?
________________________________________________________

Have you ever lived in an English speaking country? Yes No

Do you speak English on an everyday basis? Yes No

Would you describe yourself as fluent in English? Yes No

Do you actively look up words you do not understand when you read? Yes No

Are you getting opportunities in school to work with vocabulary? Yes No
Appendix B Information sheet

Automaticity in L2 learning: Correlation between vocabulary proficiency and response time in word recognition

Researcher: Signe Jonsson
Supervisor: Alan McMillion
Stockholm University
Mail: signe.c.jonsson@gmail.com

In this study you will firstly be asked to answer questions about your English knowledge and background. Secondly, you will be asked to take a vocabulary test. Thirdly, you will be asked to choose between two words and decide which one you believe is English. The last test is conducted through a computer software and will test both accuracy and response time, not one being more important than the other.

Participants must be attending upper secondary school.

The study focuses on vocabulary proficiency and response time. Participating persons’ names will be encoded to ensure full confidentiality.

Any questions will be answered at the time of the survey. Questions can also be emailed to the researcher.
Appendix C Sample of vocabulary size test

1. soldier: he is a soldier
   a. person in a business
   b. student
   c. person who uses metal
   d. person in the army

2. restore: it has been restored
   a. said again
   b. given to a different person
   c. given a lower price
   d. made like new again

3. jug: he was holding a jug
   a. a container for pouring liquids
   b. an informal discussion
   c. a soft cap
   d. a weapon that explodes

4. scrub: he is scrubbing it
   a. cutting shallow lines into it
   b. repairing it
   c. rubbing it hard to clean it
   d. drawing simple pictures of it

5. dinosaur: the children were pretending to be dinosaurs
   a. robbers who work at sea
   b. very small creatures with human form but with wings
   c. large creatures with wings that breathe fire
   d. animals that lived a long time ago

6. strap: he broke the strap
   a. promise
   b. top cover
   c. shallow dish for food
   d. strip of material for holding things together

7. pave: it was paved
   a. prevented from going through
   b. divided
   c. given gold edges
   d. covered with a hard surface

8. dash: they dashed over it
   a. moved quickly
   b. moved slowly
   c. fought
   d. looked quickly

9. rove: he couldn’t stop roving
   a. getting drunk
   b. travelling around
   c. making a musical sound through closed lips
   d. working hard

10. lonesome: he felt lonesome.
    a. ungrateful
    b. very tired
    c. lonely
    d. full of energy