English spelling errors in Swedish high school

An investigation of English spelling errors among Swedish high school students and their possible causes

Gule Akman

2019

Student thesis, Bachelor degree, 15 HE
English Studies
ENG802

Supervisor: Henrik Kaatari
Examiner: Marko Modiano
Abstract

This study investigates spelling errors among Swedish students of English. The purpose of this research is to develop a better understanding of the causes of spelling errors and to investigate whether there are any differences between program orientation and gender. The participants are English A students from Swedish high schools in year 1. The data analysed in this study consists of essays written by the participants which have been gathered into the Uppsala Learner English Corpus (ULEC).

The results showed that the academic program has a lower error rate when compared to the vocational program and that female students display a lower error frequency than male students. These results reflect the same pattern found by previous studies on achievement between the programs and across gender. However, when comparing all variables, both gender and program, the results demonstrate that academic male students had the lowest error rate of all groups, which could partially be explained by the fact that the academic male students have a higher level of engagement with video- and online games when compared to the other groups.

The language processes that the participants found the most challenging were letter omission, letter insertion and letter substitution. These results follow the same pattern highlighted in previous studies on spelling among both native speakers and learners. Finally, the exchange between the vowels <a>, <e> and <i> was found to be a frequent error among the students. Experiencing difficulties with the correct usage of vowels can be explained as a transfer error.

Keywords: Spelling error, Swedish English learners, Second Language Acquisition (SLA), Transfer, Language & Gender, Gender & spelling mistakes, Program orientation, Uppsala Learner English Corpus (ULEC)
When the English tongue we speak.
Why is break not rhymed with freak?
   Will you tell me why it's true
We say sew but likewise few?
   And the maker of the verse,
Cannot rhyme his horse with worse?
   Beard is not the same as heard
Cord is different from word.
   Cow is cow but low is low
Shoe is never rhymed with foe.
   Think of hose, dose, and lose
And think of goose and yet with choose
   Think of comb, tomb and bomb,
Doll and roll or home and some.
   Since pay is rhymed with say
Why not paid with said I pray?
   Think of blood, food and good.
Mould is not pronounced like could.
Wherefore done, but gone and lone -
   Is there any reason known?
To sum up all, it seems to me
   Sound and letters don't agree

Our Strange Lingo by Lord Cromer (1902)
Contents

1. Introduction ..................................................................................................................... 1

1.1 Research Questions ...................................................................................................... 2
  1.2 Hypotheses .................................................................................................................. 2

2. Background ....................................................................................................................... 3
  2.1 English Orthography .................................................................................................. 3
  2.2 Second Language Acquisition & Transfer .................................................................. 4
  2.3 Gender Differences within Swedish high schools & SLA ....................................... 5
  2.4 Academic & Vocational Programs in Swedish high schools ..................................... 6

3. Material & Method ........................................................................................................... 7
  3.1 Material - Uppsala Learner English Corpus ............................................................. 7
  3.2 Method ....................................................................................................................... 8

4. Results ............................................................................................................................. 11
  4.1 General Results ......................................................................................................... 11
  4.2 Results of Error Category ......................................................................................... 13

5. Discussion ......................................................................................................................... 16
  5.2 Discussion of Error Categories ............................................................................... 19
  5.3 Method Discussion .................................................................................................... 21

6. Conclusion ......................................................................................................................... 23

References ............................................................................................................................ 24
1. Introduction
With increasing globalisation, it is becoming more important for people across the world to communicate directly with each other, both for business and social purposes. English is the lingua franca of today; and, thanks to globalisation, has taken on an even more dominant role for communication purposes (Gimenez et al. 2008: 32-36). According to the British Council, as of 2014 there were approximately 1.75 billion English speakers across the globe, and by the year 2020, 2 billion people will be speaking English (British Council 2014: 2).

In Sweden, the English language is becoming more of a second language than a foreign one. English is taught in school from an early age and Swedes are well known for their high level of English proficiency (Gee 2013). The fact that the English language plays an important role in Sweden can also be seen by the English curriculum in Swedish compulsory school and high school. In high school, English language learning is built on five foundations to give the students the best opportunity to develop their language learning. These five foundations are:

1. Understanding of spoken and written English, and also the ability to interpret content.
2. The ability to express oneself and communicate in English in speech and writing.
3. The ability to use different language strategies in different contexts.
4. The ability to adapt language to different purposes, recipients and situations.
5. The ability to discuss and reflect on living conditions, social issues and cultural features in different contexts and parts of the world where English is used. (Skolverket 2019)

Despite the popularity of the English language and the objectives of the Swedish National Agency for Education, it can be a challenge for Swedish students to learn English. Among the challenges, we find a writing system that causes difficulties even for native speakers. Thus, it is a challenge for second language learners to master the language, both in speaking and writing. Generally, the focus for English language learning is more on speaking abilities; however, as writing is a means of communication and correct spelling is needed to convey messages accurately, it is essential to master it (Cook 1997). This study will focus on the writing challenges, and more precisely, the
spelling mistakes common among Swedish high school students. The spelling error categories this study will investigate are letter insertion, letter omission, letter substitution, letter transposition, apostrophe and compound. The students’ spelling mistakes, gender and high school orientation will be analysed in order to better understand where the mistakes arise, and the underlying reason for them.

1.1 Research Questions
This study aims to investigate spelling errors among Swedish high school learners of English and to understand the challenges associated with correct spelling in English. The following research approaches the subject from the point of view of four main questions which help to break down the main themes. Firstly, is there a difference in spelling errors depending on the program orientation? Secondly, is there a difference between male and female students in L2 spelling proficiency? Thirdly, which of the spelling error categories have the highest frequency of unique errors? Lastly, what causes these errors?

By identifying these areas and in particular by addressing the last two questions, this essay hopes to provide further insight into the area of spelling errors in English for both students and teachers alike. By developing knowledge and insight in the field, we may be able to provide guidance and advice or suggest practices that may help to reduce the occurrence of these errors, and as a result, students’ English spelling will strengthen.

1.2 Hypotheses
For the academic programs, the English curriculum compromises 200 hours of English education, while the vocational programs compromise 100 hours (Skolverket 2019a). Furthermore, these academic programs are good preparation for certain further studies courses, some of which are delivered in English. This may be a driving factor in the motivation of students to learn English. According to research, female students perform better than male students, and there is a higher percentage of female students that graduate from high school (Preutz 2015). Therefore, the implication of this is that the female student participants will produce fewer spelling errors than male students.

Studies conducted on both L2 and L1 users have shown that the most common type of spelling error is letter omission, followed by letter substitution and letter insertion (Cook 1997). Drawing on this finding, two of the six spelling error categories
used in this thesis have been chosen based on the difference in frequency of occurrence between English and Swedish. Thus, the hypotheses for this study are:

1. Academic students will make fewer spelling errors than vocational students.
2. Female students will make fewer spelling errors than male students.
3. The most frequently occurring spelling error category will be letter omission, followed by letter substitution and letter insertion.
4. Language transfer will be a factor affecting the spelling errors.

2. Background
2.1 English Orthography
Understanding the orthography of a language can have a significant impact on the learning process. Orthography can signify two levels of meaning: on one level, it refers to the spelling of a language, and on another, it refers to the way the letters are written. There are various categories of orthography, from very shallow to deep orthography. A very shallow orthography, such as that of Turkish and Serbo-Croatian, consists of “…highly regular, one-to-one grapheme-phoneme correspondence” (Miller 2019: 2). Deeper orthography contains a less direct correspondence between letters and sounds, such as with French and Danish. There are also certain languages such as German and Swedish which fall somewhere between shallow and deep orthographies; these are known as semi-transparent (Miller 2019).

English has an extremely deep orthography, with a complex syllabic structure which creates an inconsistent pattern between spelling and pronunciation. A single letter can be represented by several phonemes, and a single phoneme can be represented by several letters. For example, the letter <a> has six different pronunciation patterns: /a:/ in ask, /ɔ/ in was, /ɔː/ in tall, /ə/ in any and /ɜ/ in village. In addition, the sound /e/ can have different letter combinations, such as <a> in many, <ai> in said, <e> in end, <ea> in dead, <eo> in leopard, <ie> in friend, and <ue> in guess (Umera-Okeke, N. 2008: 66-70).

Because of this inconsistency across spelling and phonetics and overall the extremely deep orthography of the English language, it is important for the students to be able to understand and use these “…groups of letters, morphemes, and lexical information that is unique to each word” (Miller 2019: 3). According to Davis (1999: 98), one explanation of the deep orthography and the difficulties associated with it is
that the modern alphabet does not contain enough letters. Instead of the present 26 letters, 44 letters would be required to encompass all the phonetic sounds in English. Venezky (1970: 127) concluded that because of this complexity it is impossible to correctly read most English words if one was to “…scan left to right, letter by letter, pronouncing as he goes…”.

Researchers have also discovered that to read and write within deep orthographic systems requires a larger memory capacity, and more parts of the brain are activated compared to reading and writing in a language with shallow orthography (Spencer 2001: 218). Evidence that the orthographic spelling and syllabic complexity of the English language creates difficulties in reading, spelling and pronunciation has also been demonstrated in studies by Landerl et al. (1997), Spencer (2000) and Seymore et al. (2003).

2.2 Second Language Acquisition & Transfer

Second language acquisition (SLA) is a research field focused on the process of learning a different language (L2) from one’s first language (L1). An L2 is any other language that a person learns after their L1, whether in a natural or tutored environment (VanPatten & Beanati 2010: 1). While learning a new language, one challenge that learners can encounter is language transfer, which can affect all areas of language learning, including spelling. Language transfer is defined by Wardhaugh & Fuller (2015: 120) as follows: “the concept of transfer in second language acquisition is that learners use features of their first language in the language they are learning”. Thus, the influence of an L1 can negatively affect a student’s L2 learning.

However, the extent to which language transfer affects the errors in SLA is highly debated. In the 1950s and 1960s, language transfer was considered as having a large impact on the errors and interference in L2 learning. In the 1970s, the influence of L1 was not seen as being as important as it was previously. Rather, the opposite view was common, and the influence of an L1 was minimised using the argument that L2 errors could not be traced to the interference of the L1 (Mitchell et al. 2013: 15). According to Mitchell et al. (2013: 16), theorists today acknowledge a transfer influence from the L1, although there are different opinions about the extent to which language transfer occurs.

Several studies relating to the influence of negative transfer show different results. A study by Dulay and Burt (1973) linked only 3 per cent of errors to L1, while a
study by Tran-Chi-Chau (1975) found 51 per cent of errors to be linked to the L1. However, most studies showed that around a third of all errors could be linked to the interference of the L1 (Mitchell et al. 2013: 35-36). Finally, to understand transfer error in the writing system, Cook (2004: 139) emphasised that further discussion highlighting the actual letter shapes, the pen movements, and direction, is required for more in-depth understanding.

2.3 Gender Differences within Swedish high schools & SLA

Globally, as well as specifically in Sweden, girls have consistently obtained better results and grades than boys. The results across year 9 in the school year 2016/2017 showed that 86.4 per cent of female students and 79.2 per cent of male students were eligible for high school (Skolverket 2017).

The trend of females performing better than males has been observed since the 1960s, when the academic results of Swedish students were first collected and measured (Olsson 1981: 56). Löfström (2015: 15) emphasises that while this gap is nothing new, the discussion surrounding the reasons for this achievement gap is constantly evolving. There is a fear that males are increasingly lacking interest in school – a theory which is backed up by a growing body of research. A study by Öhrn & Holm (2014) showed that males, to a higher degree than females, do not want school to occupy too much of their time. Instead, it was important to them to engage with interests or work outside of school. Other explanations for the performance gap include that school has become too ‘girl-friendly’, and that, as a result, it does not suit boys anymore; and that girls are more diligent and adaptable than boys (Löfstrom 2012: 15).

Many SLA studies analyse age, motivation, attitude, aptitude, personality, learning styles, intelligence, personal belief, and identity. However, language and gender research so far has mainly consisted of L1 language usage, and not L2 learning (Feery 2008: 32). The few studies on gender differences in SLA address mainly three areas: “…the popularity of language learning among males and females, the learning strategies used by each gender and their differences in attitude towards second or foreign language learning” (Rivas 2017: 6). Furthermore, these studies are mainly conducted on adult middle-class western women within a classroom environment, thus covering only a small portion of L2 learners.

According to Rivas (2017: 4), females tend to be more motivated, and their motivation in language learning tends to be more integrative, which is known to
produce “…better language learning and even native-like proficiency in L2”. Iwaniec (2019: 131) points to the fact that many studies on language learning motivation have shown that female learners are more motivated than male learners; however, there are a few which show no difference in motivation levels (Iwaniec 2019: 131). As for learning strategies females’ usage of memory, cognitive compensation and affective strategies and overall a higher proficiency in L2 acquisition, can be explained by the fact that females seek more “social approval and recognition” compared to men (Rivas 2017: 8). According to Oxford (1994: 146), females not only use more strategies, but their strategies are also overall qualitatively better.

In conclusion, on the subject of whether females are more proficient in L2 acquisition, Rivas (2017: 8) explains that “…although there is no consistent conclusion on this, there seems to be a strong connection between gender and second language acquisition and performance favouring females”. Several studies come to the same conclusion that female achievement is higher than that of male students. For example, Chavez (2014) on accuracy in oral production, Gtowka (2014) in a general proficiency test, Polat (2011) on judgement of accents in terms of likeness to native speakers, Fernandez Fontecha (2010) on lexical availability and Jimenez Catalan (2010) on vocabulary production tests all show that females score higher than their male peers (Iwaniec 2019: 131).

2.4 Academic & Vocational Programs in Swedish high schools

Figure 1 shows the national average of the lowest, average and highest numbers of admission points for the academic and vocational programs in the school year 2016/17. The average number of admission points for the academic program was 57 points higher than the vocational program; 246 versus 189. The academic Natural Science programme required the highest number of admission points with 272.3, while the highest vocational program, the Natural Resource Use programme, required only 205.6. The lowest accepted number of admission points for an academic program was 229.7, which is still higher than any of the vocational programs (Gymnasium website 2017).
3. Material & Method

This section will present the material used, and the method and process applied while investigating the research questions of this essay.

3.1 Material - Uppsala Learner English Corpus

The data used for this paper was gathered from the Uppsala Learner English Corpus (ULEC). The ULEC consists of essays written by Swedish junior and senior high school learners of English and contains approximately 275,000 words. The collected essays are in .txt-format and coded into different categories showing the date of composition, essay genre (argumentative, narrative or descriptive), school year, level of English course (A, B or C), type of educational program (academic or vocational), gender of the writer and age of the writer. The students have been given various topics and written the essays on a computer with “a simple web interface” (Johansson and Geisler 2009: 1).

As one area of interest covered by this essay is the difference between both program and gender, it would also have been of interest to investigate possible differences between the first, second and third-year students. However, in year three and year two, there were only a handful of digitized essays produced within the vocational program. Thus, all essays analysed are from year one students, which contained less data in the vocational than in the academic programs. Therefore, all the vocational data was used, while only part of the academic data was used for this essay. Nearly as many

Figure 1. Admission points for Vocational and Academic program, school year 2016/17.
words both across gender and program were used for this study. See table 1 for the exact number of words used for the respective program and gender.

Table 1. Number of words used per program orientation and gender.

<table>
<thead>
<tr>
<th></th>
<th>Female</th>
<th>Male</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Academic</strong></td>
<td>11,203</td>
<td>19,627</td>
<td>30,830</td>
</tr>
<tr>
<td><strong>Vocational</strong></td>
<td>11,317</td>
<td>19,517</td>
<td>30,834</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>22,520</td>
<td>39,144</td>
<td></td>
</tr>
</tbody>
</table>

The essay genres used, from both the academic and vocational programs, was a mix of argumentative, narrative, and descriptive, with most of the essays being descriptive. The topics for the essays varied, and most topics were covered in both the vocational and academic programs.

3.2 Method

The selected essays, which were saved in a plain text format\(^1\), were copied into Microsoft Word, with the built-in spell checker set to the U.K variety of English. The U.S variety of English was equally accepted for this study. After transferring the essays into Microsoft Word, they were read one at a time. The errors discovered by the spell checker were collected first, and then the essays were thoroughly read so as to make sure no errors were missed. The spelling errors that were collected were subsequently categorised into six groups, each group representing one type of spelling error. Four of the error categories, letter insertion, letter omission, letter substitution and letter transposition have been based on the study by Cook (1997). In addition, two categories named apostrophe and compound have been included, as these are common areas of difficulty for Swedish English learners (Swan & Smith 2001: 26). See table 2 for the categories and an example of each spelling error.

Table 2. Error categories with examples.

---

\(^1\) .txt.format
The collected errors were placed into tables which were categorised according to error category and organised in alphabetical order according to the intended spelling. Each table consists of the intended spelling, the misspelt word, the number of times the misspelling occurred, as well as the number of females and the number of males that misspelt the word. The categories of omission and insertion both have an extra row for the missed and the inserted letter. Letter substitution has two extra rows in order to display which letter was replaced by which. See Table 3 for an example of the tables.

Table 3. Example of letter substitution table.

<table>
<thead>
<tr>
<th>Intended word</th>
<th>Misspelt word</th>
<th>Letter used</th>
<th>Instead of</th>
<th>Total</th>
<th>Female</th>
<th>Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>because</td>
<td>because</td>
<td>o</td>
<td>a</td>
<td>7</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>become</td>
<td>bacome</td>
<td>a</td>
<td>e</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

The final step was to count the total number of words for each essay and mark it in a separate table; thus, this also made it possible to keep track of the number of essays used. After all the errors in one essay were collected and put into the correct table, the same process was started for the next essay, and so on.

Once all the vocational data was collected, the collection of the academic data started. The academic data consists of more than double the amount of data and folders collected from almost twice as many teacher students. Thus, to obtain the most
objective result possible, only a certain selection of essays from each folder were collected and used for this study. All other steps followed the same method used for the collection of the vocational data.

The following step was to analyse the collected data. The aim of the analysis was to determine if there were any differences between the vocational and academic programs, as well as if there were any differences across gender. Subsequently, an analysis of all data was conducted, taking into account gender both within and across the programs. Finally, a thorough analysis of the data was conducted, looking in depth at the various categories of errors and individual errors made.

In Cook’s study (1997) certain misspelt words were excluded: these were “…omission of grammatical inflections’, such as the plural ‘s’ atom (atoms), and verb endings, such as find (finding) and form (formed), which are better considered in grammatical terms”. This paper followed Cook’s (1997) structure, and also refrained from using other clear past tense errors such as ‘*happend’ for ‘happened’ or ‘*buyed’ for ‘bought’, and clear plural mistakes, such as, ‘*familys’ for ‘families’ or ‘*phenomenons’ for ‘phenomena’. Then, following Kusuran (2016) the decision was made not to include the incorrect usage of the definite/indefinite articles, as when writing an instead of a and vice versa, and capitalisation as in ‘*sweden’ for Sweden.

Further mistakes which were discounted included words with obvious typing errors which occurred only due to the medium of writing used2, for example ‘*bur’ and ‘*byt’ for ‘but’, or ‘*it¥s’ for ‘it’s’. Slang lexis has also been discounted; as in ‘*kinda’ for ‘kind of’.

A key interest of this paper is in understanding the frequency of unique spelling errors rather than the total frequency of spelling errors. Therefore, regardless of how many times a student has misspelt a word, it has only been counted once. Then, once again following Cook’s (1997) study, as some words contained more than one mistake, they were also counted accordingly into the relevant error categories, for example ‘*all

---

2 Typing on a computer keyboard.
ready’ for ‘already’; this error is counted both as a compound error and as insertion error of the letter <l>. If a word has several errors within the same category, it has also been counted as such. For example, ‘*acses’ for ‘access’ belongs to the omission category. As both the letter <c> and the letter <s> are missing, it has been counted twice. There are two reasons for this; the first is to understand as thoroughly as possible what students are struggling with within the various error categories, including the various letters. Then, if a letter would have been inserted in the same word, for example, ‘*acsees’, it would also have to count as an insertion error. Thus, the logical step is to count both letters as errors.

Microsoft Word 2018 was used to help gather and detect spelling errors, and The Oxford English Dictionary (OED)\(^3\) has been used in the event of uncertainty around the spelling of a word. It has also been used in situations where it has been necessary to first determine whether a word is slang or a proper word (The Oxford Dictionary 2019a).

4. Results

In this section, the results will be presented. First, the overall results will be presented and then the results for each error category. Please see table 1 in section 3.1 for the exact number of words. Throughout this paper, whenever the phrase ‘error rate’ is used, it refers to the frequency of unique spelling errors, and not the total frequency of spelling errors.

4.1 General Results

As can be seen in figure 2, the error rate per 100 words is 3.19 for the academic students in comparison to 4.76 for vocational students. This difference is quite significant, as the difference between error rate across academic and vocational students amounts to more than 1.5 errors per 100 words. Therefore, the academic students outperform the vocational students.

---

\(^3\) The OED include both U.K and U.S varieties of English, and also the other major varieties of world English. Thus, the OED consults many other dictionaries in order to check currency and find new examples (The Oxford Dictionary 2019b).
Figure 2. Error rate per 100 words for the academic and vocational programs.

Figure 3 shows that the error rate of male students per 100 words is 4.03, in comparison to 3.88 for female students. Thus, the female students are more proficient at spelling than the male students.

Figure 3. Error rate per 100 words for female and male students.
Finally, figure 4 shows the overall results across gender and program. As can be seen from the chart, academic male students showed the best results, followed by academic females, then vocational females, and finally the vocational male students.

![Graph showing error rate per 100 words for gender in academic and vocational programs]

Figure 4. Error rate per 100 words for gender in the academic and vocational programs.

### 4.2 Results of Error Category

As can be seen in figure 5, for both programs, omission comprised the largest group of errors, and transposition comprised the smallest group. The most significant difference between the groups was displayed within the category of letter insertion, where the vocational program participants exhibited more than 70 per cent more mistakes than the academic students. The academic students performed better than the vocational students in all error categories.

The most frequently omitted letter for both groups was the letter <e>, followed by the letter <a>. The difference in error rate was quite high in both groups. In the vocational group, omission of the letter <e> was more than three times higher than that of the letter <a>, and in the academic group, it was nearly three times higher. The vocational students struggled most with the words ‘really’, ‘some’ and ‘because’, while the academic students struggled most with the words ‘because’, ‘believe’ and ‘interest’.

The apostrophe errors had the second-lowest frequency for both programs. The omitting of an apostrophe was by far the largest error; there were only occasional errors of an added apostrophe in both programs. Omission or intrusion of an apostrophe
denoting possession equally consisted of a low number. The words that both programs struggled with the most were ‘don’t’, ‘it’s’, ‘that’s’ and ‘I’m’.

Figure 5. Error rate per 1,000 words in the various categories for the academic and vocational program.

The results for females and males are shown in figure 6. The largest error category for both genders was letter omission, followed by insertion and substitution. However, with respect to differences between the genders, the male students showed an error rate that was 1.5 points higher per 1,000 words. The most significant difference was seen in the category of compound errors, with over 30 per cent more errors produced by males, followed by transposition errors where they have nearly 20 per cent more errors than females.

The most common compound error among both females and males was two words written as one. The most common misspelt words for females and males were ‘a lot’ written as ‘*alot’, and ‘of course’ written as ‘*ofcourse’. Transposition counted for the least number of errors in total. The most common error in this category for both males and females were with the words ‘because’ and secondly ‘with’. For males, this was followed by the words ‘people’ and ‘cannot’, and for females, the next most commonly misspelt words were ‘maybe’ and ‘doesn’t’.
Both figure 7 and table 4 show the total error rate in each category across gender and program. The academic male students performed better in four of the six error categories, followed by the academic females, who performed better in two error categories. Third place in overall performance were the vocational female students. Lastly, the vocational male students had the lowest performance in four error categories and made the most errors in total.

Letter insertion was the second-largest category across gender and program and accounted for about 21 per cent of all errors. The most frequently inserted letter was <e> followed by <l>, <t> and <o>. The most common words with inserted letter errors were ‘*where’ for ‘were’, ‘*also’ or ‘*allsow’ for ‘also’, ‘*hotell’ for ‘hotel’, and ‘*allways’ for ‘always’.

The third-largest group of errors across gender and program was letter substitution. The students struggled most with the word ‘because’, replacing the letter <a> with <o>. The words that followed were ‘than’, ‘since’, ‘many’, and ‘probably’. Vowels were particularly challenging for the students. The letter <a> was overall the most substituted and was most frequently replaced by the letter <e>. The letter <e> was the third most substituted letter and was replaced by various letters, however, it was most frequently replaced by the letter <a>. The letter <i> replaced by the letter <e> was
also a common error. The letter \(<c>\) was the second most substituted consonant and was replaced almost as many times by the letter \(<k>\) as by the letter \(<s>\).

Figure 7. Error rate per 1,000 words in the error categories, across gender and program.

Table 4. Error rate per 1,000 words in the error categories, across gender and program.

<table>
<thead>
<tr>
<th></th>
<th>Academic Female</th>
<th>Academic Male</th>
<th>Vocational Female</th>
<th>Vocational Male</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apostrophe</td>
<td>2,05</td>
<td>3,26</td>
<td>3,97</td>
<td>3,48</td>
</tr>
<tr>
<td>Compound</td>
<td>2,76</td>
<td>3,51</td>
<td>4,24</td>
<td>5,89</td>
</tr>
<tr>
<td>Insertion</td>
<td>6,33</td>
<td>6,01</td>
<td>10,6</td>
<td>10,7</td>
</tr>
<tr>
<td>Omission</td>
<td>11,8</td>
<td>10,4</td>
<td>14,3</td>
<td>16,9</td>
</tr>
<tr>
<td>Substitution</td>
<td>7,67</td>
<td>5,55</td>
<td>7,68</td>
<td>9,73</td>
</tr>
<tr>
<td>Transposition</td>
<td>2,94</td>
<td>2,13</td>
<td>3,09</td>
<td>2,92</td>
</tr>
</tbody>
</table>

5. Discussion
This section will provide a discussion of the results of this study as well as the method.
5.1 General Discussion

In total, the errors of the academic students accounted for 40.1 per cent of all errors, with the errors made by vocational students accounting for the remaining 59.1 per cent. Thus, the first hypothesis made at the outset of this study\(^4\) has been supported by the results of this study. There are a few potential reasons for this, which will now be discussed.

As stated in section 2.4, the difference in admission points is much higher for the academic program, as well as the curriculum for the academic program have many more hours of English per school year. Then, a research study by Statistics Sweden (SCB) (2018: 5) on how many students planned to continue with further studies showed that 80 per cent from the academic programs planned on undertaking further studies, while only 25 per cent of students from the vocational program planned on further studies. Among the academic students, 10 per cent had no plans for further studies in comparison to 50 per cent of the vocational students. These factors can provide an explanation for the results of this study. For example, students who start high school with better grades and have the ambition to continue on to further studies are surely also more inclined to perform better and become more proficient in the various subjects, including English.

The second hypothesis, that female students would display fewer spelling errors than male students, is simultaneously both supported and rejected by this study. When comparing across all females and males, the results seen in figure 3 demonstrate that females performed better than males; thus, the second hypothesis is supported. The many studies on the gender difference in school achievement, both in Sweden and internationally, provide support for this finding. Sunderland (2000: 3) points to international studies by Burstall (1975), Boyle (1987), and Arnot (1996), made on gender differences in language learning which have shown that females outperform males. In addition, Statistics Sweden’s report shows that 65 per cent of females and 50 per cent of males planned to continue their studies at university. These results are not exclusive for this school year, rather the opposite, and the same difference has been found between the genders for several years (Statistic Sweden 2018: 8/9).

\(^{4}\) The academic students will display fewer spelling errors than the vocational students.
Furthermore, as found in a study by Löfström (2014: 29), the results shown from the school year 1996/97 for students in the 9th grade shows that female students achieve better final grades than male students in English. Skolverket, in their report (2017b: 12), shows the gender difference between graduating students in year 9. A total of 92.4 per cent of females, in comparison to 87.2 per cent of males, accomplished a graduation grade (A-E). Thus, it seems that the overall grades and statistics are also represented in English spelling.

However, the second hypothesis was rejected when comparing the general results of all factors across gender and program. As can be seen in figure 4, the academic male students achieved the best results, rather than the academic female students as was expected. As noted previously in section 2.3, most studies on SLA show that females are more proficient learners than males. However, some studies support the contrary. In the study by Arnot et al. (1996), boys were more proficient language learners at A-level in the UK. Furthermore, a study by Cross (1983) on boys in two British mixed-sex comprehensive schools also showed higher results among boys (Sunderland 2000: 3).

One explanation for this may be that males have a more positive approach to digital technology (Insight Intelligence 2016). As technology has become an increasingly integral part of our lives, it can provide a foundation for gaining knowledge. Today’s video games are often played online with participants from all around the world, and much of the communication with one another is in English, the lingua franca. The modern online games that are played are frequently in English which can spark an interest in learning the language.

In 2009 Sundqvist analysed essays among 9th graders in order to measure the correlation between video games and English vocabulary. The results showed that the students, mainly boys, who played more video games had a larger vocabulary than their peers. Sundqvist and Wikström (2015) carried out a more meticulous study to better understand the correlation between vocabulary proficiency and video games. The results showed that those students who played more than 5 hours of video games per week had better results and used more advanced vocabulary in their essays, in comparison to the other groups who played between 1 to 5 hours per week, or the third group who played for less than 1 hour per week. The study also concluded that the group with frequent game players also had the highest average final grades in English. According to Sundqvist, the higher level of English proficiency seen among the boys in this study
was due to the frequency of playing video games. However, she also points out that it
cannot be taken as an absolute certainty, as more studies need to be conducted on the
subject (Sundqvist & Wikström 2015).

Sundqvist and Wikström (2015: 66/67) also point to a couple of studies made on
specific video games which measured the players’ increased English vocabulary, which
showed positive results (Miller & Hegelheimer, 2006, Ranalli 2008 and Rankin et alt
2010). There are also a few other studies measuring the impact of video games or digital
teaching games on English learning which also have shown positive results (Aghlara &
Swedish National Agency of Education (2012) have also shown the correlation between
students’ higher English proficiency and out of school digital activities.

These studies provide support for the results of the higher performance shown
by academic male students; however, it does not provide any insight as to why
vocational male students did not perform better than vocational female students. One
reason may be that academic male students play more video- and online games than
vocational male students. However, this speculation is outside the scope of this essay,
and further investigation must be done in order to clarify this issue.

5.2 Discussion of Error Categories
As the third hypothesis predicted, letter omission was the largest category of errors
across all groups. The results of testing the second hypothesis, however, were
unexpected. The original hypothesis was that letter substitution would follow next after
letter omission; however, overall, letter insertion had more errors than letter
substitution. It was only in the context of the academic program where substitution
followed letter omission.

The most frequently omitted letters were <e>, <a>, <l>, <o> and <h>. That three
of the most omitted letters among the Swedish high school students were vowels can be
best explained by the fact that double vowels are rare in the Swedish language.
Beckman & Ekhall (2006) made a statistic study with the 32 most common bigrams
with its per centage for both the Swedish and English language. There was only one
vowel bigram, 0,3 per cent, among the 32 most common Swedish bigrams. On the other
hand, there were a total of nine vowel bigrams with a total of 4.5 per cent among the 32
most common English bigrams. Instead, the Swedish language has nine vowels, and a
clear distinction is made between long and short vowels (Zetterholm 2019).
Another difference with the English language is that when double vowels occur in the Swedish language, they consist mainly of the same letter. Common nouns create double vowels as in ‘leende’ and ‘seende’. Compound words also create double vowels, as in ‘extraavgift’ and ‘dataavdelning’. Last, negations create double vowels, and it is mainly in this category where two different vowels will create a bigram as in ‘oacceptabel’, ‘oätbar’ and ‘oönskad’. Instead, in comparison to English the Swedish language has a higher number of double consonant letters (Beckman & Ekhall 2006: 3 & 6). Thus, the students’ struggle with vowel omission might have its explanation in these factors. However, a similar result with vowel omission being the largest error was also found in a study on Arab English learners by Subhi & Yasin (2015).

As for insertion, the letter <l> also showed a high error number in the insertion’s category among the Swedish students. Thus, there is confusion over whether to use a single or double <l>. However, Cook’s study (1997) showed that adult L1 and L2 also had problems with the double <ll>, thus not making it an exclusive error made by the Swedish students. The insertion of the letter <t> was the third most common, with most of the mistakes made being either a doubling of the letter <t> or an insertion of the letter <t> before <ch> sounds as in ‘*mutch’ and ‘*beatch’. That the doubling of <t> is common among the Swedish students might very well be a transfer error as it is more common in the Swedish language. In a study by the National Centre for Mathematics Education, the bigram <tt> was among the 7th most common bigrams in the Swedish language, accounting for 1.1 per cent, in comparison to English, where it accounts only for 0.48 per cent (Beckman & Ekhall 2006: 3-6).

Another discrepancy between the results of this study and Cook’s findings (1997) is that the L2 users of Cook’s study displayed more consonant substitution than vowel substitution in contrast to the Swedish high school students of this study where vowel substitution constituted a much larger number of errors. It seems that for all English users, learner or native, the exchange between vowels <a>, <e> and <i> is the most common. This result was also found in the study by Bebou (1985), as well as in this study.

In Cook’s study (1997), most consonant errors were an exchange between <s> and <c>, then <z> and <t>. The Swedish students did make the exchange between <s> and <c>, though there was one key difference. It was the letter <c> that was most

---

5 See section 5.3 for further details on L2 users of Cook’s study and the Swedish high school students of this study.
frequently replaced and almost the same amount by <k> and <s>. Then it was followed by the exchange between the letters <v> and <w>. According to Swan & Smith (2001), the replacement of <c> with <k> is due to the fact that the letter <k> is much more frequently used than the letter <c> in the Swedish language. Thus, it can be classified as a transfer error. Furthermore, the exchange between the letters <v> and <w> can be explained as being due to the Swedish English learners tending to replace <v> with <w> in those cases where uncertainty exists, as they assume that this spelling is more “English” (Swan & Smith 2001: 26).

The apostrophe is rarely used in the Swedish language, especially in comparison to English. The genitive apostrophe is used, however, only in occasions to avoid confusion with the letters <s>, <x> or <z>, as in ‘Anders’ cykel’ (Anders’ bicycle). Then, the apostrophe is used to mark quotes within quotes as in English. Contractions, on the other hand, are highly rare in Swedish and are mainly used to convey phonetically transcribed slang or slur. Thus, this might indicate a reason as to the challenges with the correct usage of the apostrophe, especially contraction apostrophes. Further indications that this might be a transfer error can be found when comparing errors with L1. The studies show that English speakers struggle with apostrophe errors as well; however, it seems that the genitive apostrophe error is more common than contractions among the native speakers (Bryant et al. 1997: 106 & Little, 1986: 16).

5.3 Method Discussion
There are a few limitations to this study with regards to the ULEC and Microsoft Word. The data used from the ULEC has been extracted from essays of the same grade and level. However, there was no information available as to whether the students had carried out their studies entirely in the Swedish educational system, nor whether Swedish is the students’ L1, or if they have a different L1. Though, as the interest of the ULEC is to have data of Swedish learners of English, thus an assumption can be made that as broad as possible the compiling data are from students with Swedish as their L1. Therefore, whenever this study refers to Swedish learners of English it is making a general assumption that the students have Swedish as their L1, and that English is their L2. In Cook’s study (1997) the L2 are of various nationalities as well as the data are from two different collections, an EFL test and students’ work, from the University of Essex, England.
In certain classes, many students made the same mistakes, while other classes did not. Thus, this raises the question of how much of a difference there is in English education between the various schools across the country. Of course, regardless of the fact that the curriculum is the same, it is impossible for the same lessons and knowledge to be given to each student across the various Swedish schools. However, it does raise several important questions, for example of how much of the difference between the genders and between the programs have to do with those variables, and how much has to do with the difference in a given education system.

Another issue that was raised during the process of this study is the question of the actual knowledge of the students. For example, student A might not use advanced words; thus, the student will equally not display any spelling mistakes, whereas student B may use advanced words but also misspell them. The question is who knows more, the student who uses more frequent words and is less advanced in their spelling, or the student who uses less frequent words with a more advanced spelling. Hence, this leads to the concept that the ideal situation for a study of this topic would be to have students do a spelling test consisting of a list of specific words.

Another interesting aspect to look into would have been to investigate the difference in achievement between the first year and the last year of high school, and to analyse whether these results change throughout the school years or remain the same. A study into this might help shed further light onto the differences between the genders and orientation, and, in turn, used as an aid to improve the students’ spelling, and thus their English proficiency. It was not possible for this study due to the lack of vocational data in the third year. However, it could be possible to execute this research with another corpus or a specially constructed study.

Microsoft Word 2018 was used in the process of gathering data for the study mainly due to its built-in spell checker. However, the program has its limitations and cannot detect all errors. For example, a single word compound which is written as two words such as ‘a way’ instead of ‘away’, and misspelt words that form actual words such as ‘dear’ instead of ‘dare’ or ‘tow’ instead of ‘two’. The essays were, therefore, read one by one; however, there is a possibility that some subtle errors may have been missed.
6. Conclusion

The initial questions posed by this essay were as follows: is there a difference in spelling errors depending on the program orientation, is there a difference between male and female students in L2 spelling acquisition, which of the spelling error categories have the highest frequency of unique errors, and what causes these errors?

In terms of the difference between the academic and vocational program, this study showed the same pattern as previous studies on achievement levels between the programs. The academic students displayed far better results than the vocational students. Factors which can explain the reasons behind the results of this study are that students of the academic program have far higher admission points, a stronger desire and ambition for further studies, and the fact that the academic programs contain twice as many hours of English per school year than the vocational programs.

In total, females performed better than males which also is supported by previous studies on scholar achievement both internationally and in Sweden. However, narrowing the results down and comparing both by gender and program, the results showed that academic male students have the lowest error rate. While investigating the reasons for this, interesting research was found on the correlation between technology and language proficiency. In particular video and online games have been shown in previous studies to have a positive impact on students’ achievement in English. Most of those studies have also shown that males are more open to technology and play more video and online games, and thus, they achieve better results. Even though no conclusion can be made, it would be interesting to see further studies conducted in this area.

As for which error categories the students struggled most with, the results did not show any major irregularities. The three largest groups of error, omission, insertion and substitution, with omission being the largest, followed the previous results of studies made on English spelling errors. The only difference found by this study was that insertion errors were more frequent than substitution errors. That the irregularity and inconsistency of the English writing system creates difficulties between spelling and pronunciation is also evident from this study.

This study showed large numbers of vowel mistakes in the error categories omission, insertion, substitution and transposition. Previous studies also show that L1 students and other L2 students struggle with the English vowels, although, not to the
same extent as this study. The factor for the errors shows signs of transfer as there is a difference in the number of vowels and its usage between Swedish and English; Swedish has nine vowels and makes a distinction between short and long vowels as well as double vowels which are very rarely used. Then, the mistakes of the consonant letters <t>, <c>, <k>, <s>, <v> and <w> also indicated that the Swedish language affected the students’ English spelling. However, due to the limitations of this study, as well as the fact that not enough research has been conducted on Swedish English learners’ spelling mistakes, it is impossible to draw any conclusions.

Apostrophe and compound errors were used as categories for this study because of the significant difference in their usage between Swedish and English. The apostrophe has minimal usage in Swedish, very rarely as a genitive, and hardly ever as a contraction. Whereas in English, it is more frequent, especially as a genitive or contraction. As this study showed, most errors were due to contraction errors; thus, it is also a possibility that transfer is a causal factor. The compound also has a difference in usage between the two languages; it is more frequently used in Swedish, and most errors of this study consisted of two words being written as one, as it would have been written in Swedish. These findings are interesting, however, because of the lack of research in this area and the limitations of this study, no conclusions can be made.

In conclusion, this paper has shown indications of error patterns within the spelling categories. However, there is not enough scientific research conducted in the area to draw any firm conclusions. Further scientific studies on Swedish English learners’ spelling mistakes are required to understand the area and the errors in more depth. The results can be a great help for Swedish English teachers. By understanding the area of errors and thus, the difficulties, this can provide a basis to the teachers for what to put more focus on, and, in turn, to raise the knowledge level of the students.

References


