

On spelling behavio(u)r: A corpus-based study of advanced EFL learners' preferred variety of English

Tove Larsson, Uppsala University

Abstract

The present study investigates variation in spelling based on British versus American English norms in the writing of university students in Sweden, Bulgaria and Italy. It also examines to what extent the students are consistent in their choice of variety. The corpus material on which the study is based allows for investigation of possible changes over time, across student levels and across nations. Contrary to findings of previous studies, the results reveal a clear preference for British English spelling for all the investigated subcorpora. The students are generally consistent in their use of one variety.

1. Introduction

In the European context, English as a Foreign Language (EFL) learners are, for both historical and geographical reasons, more influenced by British English and American English than by any other variety of English (Hoffmann, 2000: 7). However, the general opinion regarding which variety is preferred has changed over the years. While British English traditionally has been perceived to be the standard variety of English in Europe (Trudgill & Hannah, 1994: 1), Europe is now described as being involved in a process of "Americanization" and the acceptance of other varieties of English is gradually increasing (Modiano, 2002: 14). Several studies investigating students' attitudes and use of different varieties of English in a Swedish context have pointed to such a shift from a preference for British English to an increasing preference for American English (e.g. Alftberg, 2009; Mobärg, 1999; Modiano & Söderlund, 2002; Westergren Axelsson, 2002). Previous studies on students' actual linguistic performance have, however, mainly focused on vocabulary or pronunciation, and spelling has been largely neglected. Then, in order to further investigate this alleged shift in foreign language behavior, the present study carries out a large-scale investigation of which variety of English – British English or American English – Swedish university students adhere to with regard to spelling, and also whether the students are consistent in their choice of one variety. The Swedish results are subsequently compared to the results

from two other European countries: Italy and Bulgaria, in order to gain a broader perspective.

1.1. British and American English in an educational setting

Although it is, of course, important to keep in mind that there are more similarities than differences between British English (BrE) and American English (AmE), differences do still exist and are of interest in particular from a teaching perspective since they bring up the question of which variety – or varieties – ought to be the target in EFL teaching. In the Swedish context, this question has been brought to the fore as, although British and American English are now said to have equal status (Altenberg, 2002: 143), several studies have pointed out that teachers still generally seem to prefer British English over American English, (e.g. Modiano, 1993; Westergren Axelsson, 2002), even to the point that it is, according to Modiano & Söderlund (2002: 147), not uncommon for teachers to “let it be known, in one way or another, that AmE [is] less valued in comparison to BrE.” However, considering the number of different Standard varieties of Englishes that are used around the world, maintaining the traditional view of British English as presenting the only model for what is considered to be incorrect and correct use of English might then result in what can be described as a fairly limited view of English in an EFL setting (cf., e.g., Brown, 1995, 2002 for a more detailed discussion).

Despite the fact that Swedish teachers of English have been found to generally prefer British English, studies investigating Swedish students’ attitudes towards and actual preference for different varieties of English have indicated a shift from a preference for British English to an increasing preference for American English. Modiano & Söderlund (2002) found a clear preference for American English (including American English spelling) among the upper secondary school students that took part in their 1999 study. Comparing their results to a previous study conducted at university level in 1992 (Modiano, 1993), in which a majority of students showed a clear preference for British English, the authors concluded that “it would seem that the 1990s will be remembered as the decade when the Swedes abandoned their preference for BrE” (Modiano & Söderlund, 2002: 149). Swedish students’ increasing preference for American English has also been confirmed in

other studies carried out in compulsory and upper secondary school (Alftberg, 2009; Mobärg, 1999) and, according to Modiano & Söderlund (2002: 149), there is “a great likelihood that this Americanization process will continue.” Nonetheless, a preference for British English among university students was still found in Westergren Axelsson’s (2002) study carried out in 2000. The preference was, however, less prominent compared to a similar study carried out in 1992, which was reported to serve as evidence that the preference for the two varieties is increasingly becoming more balanced (Westergren Axelsson, 2002: 143).

There are two main factors that have been reported to influence the preference for one variety of English over another in an EFL context: general norms and traditions of teaching a certain variety on the one hand, and media influence and exposure on the other hand. Since the teachers in Sweden have been found to generally show a higher preference for British English, the increasing preference for American English shown among Swedish students can seemingly not be explained solely by virtue of it being the variety of English that is most commonly taught at school; instead, media influence and exposure appears to be an important factor (cf. Mobärg, 1999). However, the extent to which students get exposed to English through the media differs slightly across different European countries. While people in Sweden receive (mainly American) English input from the media as movies and TV programmes are not dubbed, the Italians generally get less exposure to English as next to all Italian movies and TV programmes are dubbed (cf., e.g., Pulcini, 1997 for more detailed discussion of English in Italy). This, taken together with the fact that British English is reported to remain “the most widespread model among secondary school teachers and students because of geographical proximity and tradition” (Prat Zagrebelsky, 2002: 110), would then be likely to result in a higher preference for British English among Italian students. In Bulgaria, on the other hand, British English and American English have, for political reasons, both been influential in EFL teaching and today, American English is used mainly in the fields of business and technology, while British English is considered to be the primary language of education and literature (cf. e.g., O’Reilly, 1998 for more detailed discussion of English in Bulgaria). Furthermore, movies and TV programmes are not dubbed, which means that people in Bulgaria get exposed to English through the media as well (Blagoeva, 2002; Griffin, 2001).

Another related issue of interest in this context is the extent to which students are consistently using only one variety of English. In previous studies carried out in a Swedish setting, a tendency for inconsistent use of the varieties has been found. Although Modiano & Söderlund (2002) do not draw any explicit conclusions based on their findings in terms of degree of consistency, a relatively high occurrence of inconsistent use of British English and American English spelling conventions can be discerned from their results. One study targeting university students' preferred variety with regard to pronunciation that specifically comments on consistency reports that "many students are aware of their lack of consistency and regret that they cannot reach their ideal, a pure, unmixed national accent" (Westergren Axelsson 2002: 133). Teachers are, furthermore, said to typically prefer their students to be consistent (Westergren Axelsson, 2002: 142).

1.2 Overview of the present study

The present cross-sectional study aims to investigate i) which variety – British or American English – Swedish university students adhere to with regard to spelling, and ii) to what extent the students are consistent in their use of British or American English spelling. Rather than investigating students' attitudinal preference for one variety of English over another, which has been the focus of several previous more small-scale studies, the present corpus-based study allows for a large-scale investigation of their actual preference. The results are analyzed to detect possible variation across the student levels (first through fourth term of studies) and to detect possible changes over time. Finally, although the main focus of the study is on the Swedish context, the results from the two Swedish national subcorpora will also be compared to the national subcorpora from two other European countries – Bulgaria and Italy – in order to contrast the results and gain a broader perspective of the use of English in an EFL context.

2. Spelling differences between British English and American English

Although the vast majority of the words are spelled in the same way in British English and American English, there are certain noteworthy differences between the varieties. Most of these differences are due to

Noah Webster's *Spelling Book* published in 1783, in which he aimed to standardize American spelling and thereby avoid dependence on the linguistic model of the UK (Tottie, 2002: 8-10). Today, there are both rule-bound and irregular differences between the spelling variants of British English and American English; however, the focus of this study is on systematic differences. The following systematic differences are found when comparing British English spelling to American English spelling (Tottie, 2002: 10-11):

Suffixes

- *-our* (BrE) and *-or* (AmE); as in *colour/color* and *humour/humor*
- *-re* (BrE) and *-er* (AmE); as in *centre/center* and *litre/liter*
- *-logue* (BrE) and *-log* (AmE); as in *dialogue/dialog* and *prologue/prolog*
- *-ence* (BrE) and *-ense* (AmE); as in *defence/defense* and *licence/license* (the alteration between *s* and *c* is, however, reversed in certain words such as in BrE *practise*, and AmE *practice*)
- *-amme* (BrE) and *-am* (AmE, sometimes also in BrE); as in *programme/program*
- *-exion* (BrE) and *-ection* (AmE, sometimes also in BrE); as in *connexion/connection*
- *-ise* (BrE) and *-ize* (AmE, sometimes also in BrE); verb-ending, as in *organise/organize*
- *-yse* (BrE) and *-yze* (AmE, sometimes also in BrE); verb-ending, as in *analyse/analyze*

Doubling of -l

- Verb-final *-l* is doubled before the endings *-ed* and *-ing* in BrE, but not in AmE; in words such as *travelled* or *cancelling* (BrE), and *traveled* or *canceling* (AmE)
- In a few other cases, *-l* is doubled at the end of certain words or in the middle of other words in AmE, but not in BrE; in words such as *fulfil* or *skilful* (BrE), and *fulfill* or *skillful* (AmE)

Loanwords

- Greek or Latin loanwords have simplified spellings with *e* instead of *ae* or *oe* in AmE, but usually not in BrE; in words such as *aesthetic* or *foetus* (BrE), and *esthetic* or *fetus* (AmE)

3. *The investigation*

3.1 *The corpus material*

The material used for the present study was culled from two large corpora: the *International Corpus of Learner English* (ICLE) and the *Stockholm University Student English Corpus* (SUSEC). The full ICLE corpus comprises 3,640 essays and over 2.5 million words divided into eleven national subcorpora, three of which – the Swedish, Italian and Bulgarian subcorpora – were chosen for this study. The corpus-material was collected from EFL students at university level and includes mainly argumentative essays. The Swedish subcorpus (SWICLE), comprises 363 essays and 206,015 words; the Italian subcorpus (IT) comprises 397 essays and 229,412 words; and the Bulgarian subcorpus (BG) contains 302 essays and 203,077 words, as is shown in Table 1. All essays were collected during the 1990s from students in their third or fourth term of study.

The complete SUSEC corpus comprises 368 texts and more than one million words of academic essays written by non-native speakers of English studying at the English Department of Stockholm University in Sweden, as well as of native speakers from King's College in the UK. The material was collected in 2007 from university students studying general linguistics, English linguistics and English literature. Since the focus of the present study is on the EFL context, only the Swedish essays were chosen for further investigation. The Swedish subcorpus (henceforth referred to as the SUSEC) comprises a total of 910,324 words and 286 essays from students in their first, second, third and fourth term of study. An overview of the SUSEC subcorpora included in the present study is presented in Table 2 below.

Table 1. Overview of the ICLE subcorpora included in the present study

| Component | Number of essays | Number of word tokens |
|----------------------------|------------------|-----------------------|
| Swedish subcorpus (SWICLE) | 363 | 206,015 |
| Bulgarian subcorpus (BG) | 302 | 203,077 |
| Italian subcorpus (IT) | 397 | 229,412 |
| Total | 1,062 | 638,504 |

Table 2. Overview of the SUSEC subcorpora included in the present study

| Component | Number of essays | Number of word tokens |
|--------------------|------------------|-----------------------|
| First-term essays | 117 | 117,175 |
| Second-term essays | 90 | 182,829 |
| Third-term essays | 62 | 417,772 |
| Fourth-term essays | 17 | 192,548 |
| Total | 286 | 910,324 |

3.2 Method

Since the present study is corpus-based, an investigation of a large number of authentic texts was made possible. A total of 1,348 student essays were investigated and all instances of British or American English

spelling in the present study are found in an authentic context (i.e. as part of an essay rather than a survey), which allows for a focus on actual language use.

The list of spelling variants chosen for investigation in the present study is based on Tottie's (2002: 10-12) overview of different categories of standard spelling that traditionally differs between British English and American English. Non-systematic spelling differences and categories that are not exclusively used in British English and American English were excluded. In addition, both the doubling of the letter *-l* before the endings *-ed* and *-ing* in words used in British English spelling, as well as the doubling of the letter *-l* in certain words, such as *fulfill* and *skillful* used in American English, were excluded from the study. The four categories of spelling differences selected for further investigation were then: *-our/-or*, *-logue/-log*, *-tre/-ter*, and *-ence/-ense*. The next step was to select words to represent each category. For the *-our/-or* category, which contained the most commonly occurring words, as well as for the category including *-logue/-log*, complete coverage of the words spelled in accordance with British English spelling was achieved. These words were subsequently matched with the equivalent words spelled in accordance with American English standards. In order to ensure that this procedure did not result in a list of words that were biased in favor of British English spelling, a random sample of 40 essays (ten from each national subcorpus) were gone through manually to search for any occurrences of words belonging to the *-our/-or* or *-logue/-log* categories spelled in accordance with American English spelling. No additional words were, however, found and the list was therefore concluded to be representative of both varieties. For the remaining two categories in which words including very frequently occurring clusters of letters were contained (*-tre* vs. *-ter* and *-ence* vs. *-ense*), a selection of three sample words for each category was made based on Tottie's list (2002: 10-11). Relevant inflectional and derivational forms of all words were also included; however, words such as *humorous*, where adding a derivational suffix to the word stem entails that the spelling does no longer differ between British English and American English, were excluded.

In total, 30 words were included in the study and are listed below:

- *-our* (BrE) and *-or* (AmE):
Ardour/ardor, armour/armor, behaviour/behavior, colour/color, clamour/clamor, demeanour/demeanor, endeavour/endeavor, favour/favor, flavour/flavor, harbour/harbor, honour/honor, humour/humor, labour/labor, neighbour/neighbor, odour/odor, rumour/rumor, vigour/vigor, vapour/vapor, savour/savor, tumour/tumor
- *-logue* (BrE) and *-log* (AmE):
catalogue/catalog, dialogue/dialog, monologue/monolog, prologue/prologue
- *-re* (BrE) and *-er* (AmE):
theatre/theater, centre/center, metre/meter
- *-ence* (BrE) and *-ense* (AmE):
defence/defense, offence/offense, licence/license

The rate of occurrence of each word was investigated using the AntConc concordancer software (Anthony, 2012). A manual examination of each word was, however, necessary in order to eliminate “noise” hidden within the results such as irrelevant words including the targeted cluster of letters (for example *your, our, of course*, which all include the letters of the targeted category of *-our*). Furthermore, names of places, buildings and parties etc. (for example *World Trade Center, Globe Theatre* and *Labour Party*) were excluded. In addition, all words within quotation marks and all references were excluded in order to be able to detect spelling choices made by the author him/herself. For the SUSEC, the stripped version, i.e. the version in which the list of references and all quotations had been removed, was used; for the remaining subcorpora, references and quotes were, when included, removed manually. The degree of consistency was examined by searching those essays marked for American English spelling in AntConc for any occurrences of the examined words spelled according to British English spelling conventions.

The overall frequency and degree of consistency were calculated for all subcorpora separately. In order to ensure that potentially high frequencies in a small number of essays would not affect the results unduly, both the number of the investigated word tokens and the number

of essays in which British English or American English spelling conventions were used consistently were investigated. Since the subcorpora differ in size, the overall frequency was normalized per 100,000 words, and the number of essays was normalized per 10 essays. The results were furthermore tested for statistical significance¹ using the chi-square contingency test to test differences in relative frequencies, or the chi-square test for goodness-of-fit to test differences in absolute frequencies.

4. Findings

In this section, the results of the present study are presented. In section 4.1, the findings of the investigation of the Swedish data are presented, first with regard to preferred variety and, second, in terms of degree of consistency. The results from the Swedish national subcorpora are subsequently compared to the Bulgarian and Italian national subcorpora in section 4.2. An overview of the results from all the investigated subcorpora showing the distribution of the word tokens spelled in accordance with British or American English standards, as well as the number of essays in which British or American English was used consistently can be found in Appendix A. An overview of the distribution of the inconsistent and consistent essays is presented in Appendix B.

4.1. The Swedish context

4.1.1. Preferred variety: an investigation of possible differences across four student levels

The results show a clear and statistically significant predominance of British English spelling in all of the four subcorpora of the SUSEC. In

¹ Chi-square tests are used to investigate whether the null-hypothesis, which states that there is no difference between the frequencies subjected to the test, can be rejected. If it can be rejected, the frequency differences for a certain set of data is statistically significant (Johannesson, 1986: 91). As is customary for linguistic investigations, the frequency difference was taken to be statistically significant for error probability value (p) lower than 0.05 and chi-square value higher than 3.84 (cf. e.g. Johannesson (1986: 91-99) for a more thorough description of chi-square tests).

total, 78% of the investigated words were spelled according to British English conventions, and 22% according to American English standards, which can be broken down to first-term essays (70% BrE, 30% AmE), second-term essays (63% BrE, 37% AmE), third-term essays (87% BrE, 13% AmE), and fourth-term essays (82% BrE, 18% AmE), as shown in Figure 1. The greatest difference between the spelling varieties was found in the third-term data where almost 90% of the investigated words were spelled in accordance with British English conventions, to be compared to the second-term data, where 63% of the words were spelled using the British English standard. All differences found between the SUSEC subcorpora were statistically significant except for the differences between the first-term and second-term essays, and the third-term and fourth-term essays.

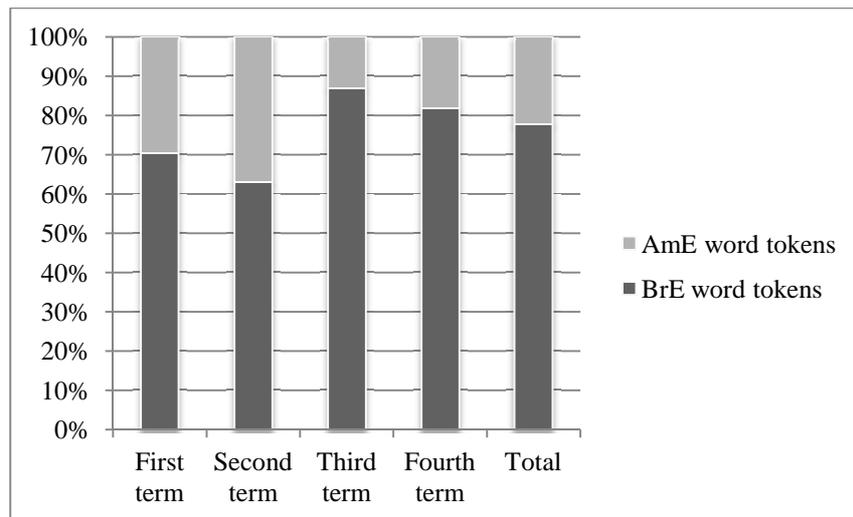


Figure 1. Relative frequency of British English and American English word tokens

The same pattern can be discerned from the investigation of the relative frequency for the number of essays in which British English or American English spelling conventions were used consistently, as shown in Figure 2. In total, out of all the consistent essays, 77% included words spelled in accordance with British English standards, while 23% included words spelled in accordance with American English standards. In terms of usage across the different student levels, the results showed a slightly

more leveled-out pattern. Again, the greatest preference for American English spelling was found in the second-term essays (70% BrE, 30% AmE), whereas the rest of the SUSEC subcorpora showed a comparably stronger preference for British English spelling: 79% BrE and 21% AmE in the first-term essays, 81% BrE and 19% AmE in the third-term essays, and 82% BrE, and 18% AmE in the fourth-term essays. However, although the differences between the preference for British English or American English spelling were statistically significant at all levels, none of the differences between the levels were statistically significant (in some cases, the numbers were too low to test for statistical significance).

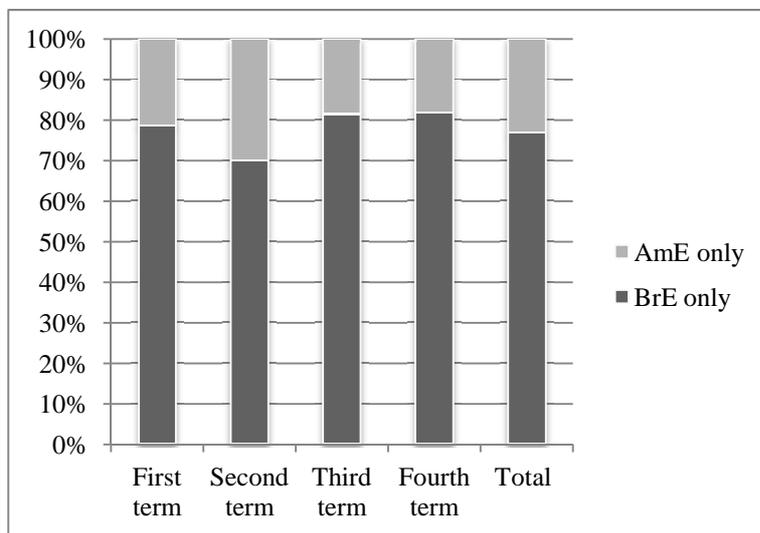


Figure 2. Relative frequency of the essays in which British English or American English spelling is used consistently

There does, thus, not appear to be a gradual increase of preference for either variety from one level to the next, which would suggest that EFL studies at a higher level in a Swedish context do not necessarily entail an increasing preference for either British English or American English. However, the fact that there is a very strong preference for British English across all student levels taken together with a slightly stronger preference among students in their third and fourth term of studies may indicate that British English is perceived to be more formal than

American English (cf. Mobärg, 2002), which could be connected to the history of predominance of British English in Sweden.

4.1.2. Preferred variety: an investigation of possible changes over time

The predominance of British English spelling does not appear to have undergone any considerable change over the years separating SWICLE (compiled in the 1990s) from SUSEC (compiled in 2007), as the slight increase in use of American English spelling was not statistically significant. Out of the investigated words, the results for SWICLE show that 82% of the words were spelled in accordance with British English conventions, while 18% of the words were spelled according to American English standards. This can be compared to SUSEC where 78% used British English spelling and 22% used American English standards, as is shown in Figure 3.

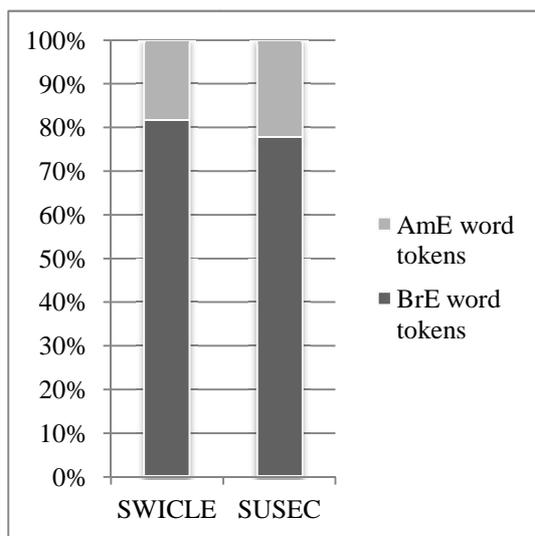


Figure 3. Relative frequency of British English and American English word tokens

A slightly larger, albeit still not statistically significant, difference can be discerned when investigating the relative frequency of the number of essays in which only British English or American English spelling is used. While 83% of the essays only comprised words spelled according

to British English standards and 17% only comprised words spelled in accordance with American English conventions in the SWICLE, 77% of the students used only British English spelling and 23% of the students used only American English spelling in SUSEC, as can be seen in Figure 4.

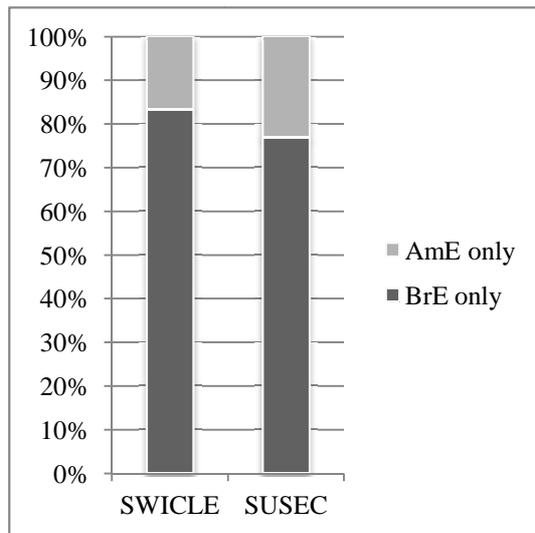


Figure 4. Relative frequency of the essays in which British English or American English spelling is used consistently

The fact that there was no statistically significant difference between the two Swedish national subcorpora suggests that despite the Americanization process allegedly taking place all over Europe (Modiano, 2002: 14), the preference for British English spelling has remained next to unchanged at university level for the years separating SWICLE and SUSEC. However, the slight, albeit not statistically significant, increase in preference for American English spelling over the years might still indicate that the preference for British English and American English is increasingly becoming more balanced, as was also reported in Westergren Axelsson (2002).

Furthermore, since there is a clear and statistically significant preference for British English spelling in both SWICLE and the SUSEC, these findings then seem to contradict findings of previous studies

targeted at younger students where a clear preference for American English spelling was found (e.g. Alftberg, 2009; Modiano & Söderlund, 2002). One possible explanation to the diverging results could be the difference in educational level of the students. Given that British English is still generally perceived to be higher in style compared to American English (cf. Mobärg, 2002), studies at a higher level might then result in a preference for the variety considered to be highest in style.

4.1.3. Degree of consistency: an investigation of possible differences across four student levels

The results for the SUSEC show that the students were generally consistent in their use of one spelling variety, as shown in Figure 3. In total, the results show that 88% of the students were consistent, while 12% were inconsistent. This can be broken down to the first-term essays (95% consistent, 5% inconsistent), the second-term essays (91% consistent, 9% inconsistent), the third-term essays (84% consistent, 16% inconsistent), and the fourth-term essays (65% consistent, 35% inconsistent), as is shown in Figure 5. The differences found between the levels are, however, not statistically significant (in certain cases the numbers were, again, too low to test for statistical significance).

Since the differences found between the different student levels lacked statistical significance, no general conclusions can be drawn from these findings. It is, nevertheless, interesting to note that an unexpectedly high incidence of inconsistent essays can be found among the fourth-term essays despite the fact that the fourth-term students can be expected to have better knowledge of English than the first-term students. As many as 6 essays out of 17 (35%) included inconsistent use of British English and American English spelling, to be compared to the first-term essays where only 3 essays out of 59 (5%) included inconsistent use of the varieties. One possible explanation to these findings could be linked to the average length of the essays, as the fourth-term essays were approximately 11,000 words long, while the first-term essays were approximately 1,000 words long. Given that this is the case, it seems that a longer essay would result in a higher likelihood of inconsistent use of the varieties. Despite thorough revision and the possible use of spell-checkers, these students then appear to be influenced by both varieties of English to the extent that they either are not aware of the differences

between the varieties, or they choose not to focus on them. This observation is, of course, based on a small number of essays and would need to be investigated further; nonetheless, it gives rise to the question of whether EFL teachers in Sweden can – or should – demand consistent use of one variety from their students (cf., e.g., Modiano, 2002).

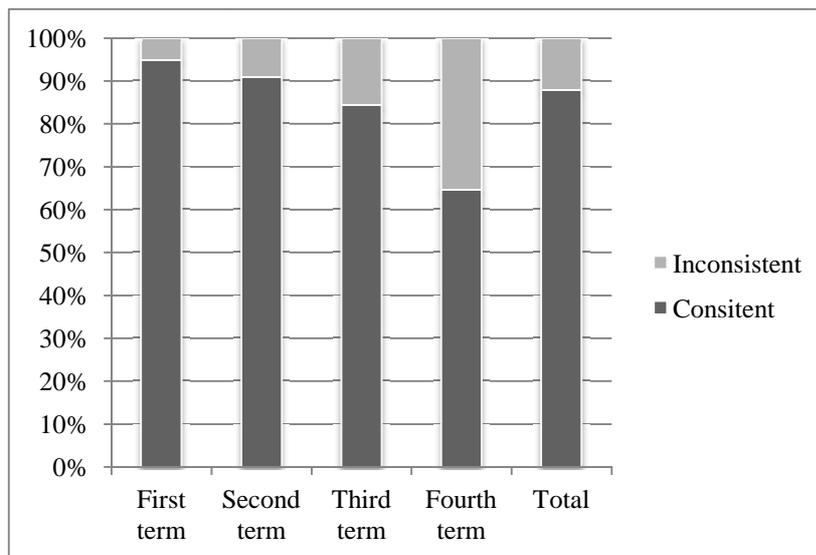


Figure 5. Relative frequency of inconsistent and consistent essays

4.1.4. Degree of consistency: an investigation of possible changes over time

The results show that there has been a slight, albeit not statistically significant, decrease in consistency over the years separating the SWICLE from the SUSEC, as shown in Figure 6. In the SWICLE, 94% of the essays were consistent, while 6% were inconsistent. In the SUSEC, 88% of the students were consistent, while 12% were inconsistent.

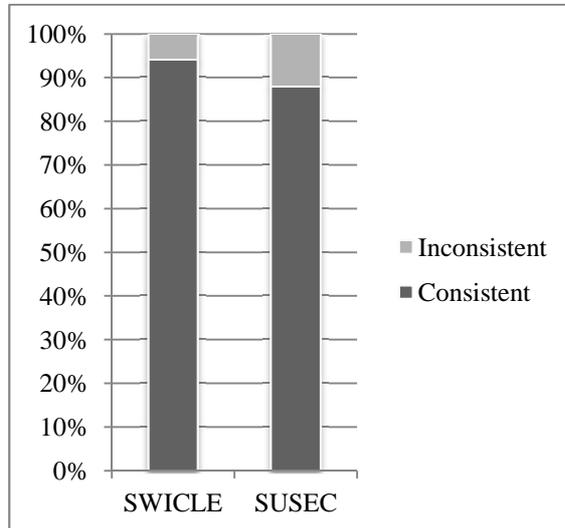


Figure 6. Relative frequency of inconsistent and consistent essays

The students were then generally found to be consistent in their use of one variety and only approximately one student out of ten is inconsistent in his or her use of British English and American English spelling conventions. These findings differ from Modiano & Söderlund's (2002) study, as their results indicated a slightly higher tendency of inconsistent use of spelling conventions. One possible explanation for the lower incidence of inconsistency in this study might be that university students of English are generally likely to be more aware of the differences in spelling between the two varieties than students at upper secondary school level and would therefore be able to consciously be more consistent. The diverging results could also be explained by a possible correlation between an increasing preference for an alternative variety (in this case American English) in a context where British English has a history of predominance on the one hand, and a higher degree of inconsistent use of the varieties on the other hand. Then, since the students in Modiano & Söderlund's (2002) study used American English spelling to a larger extent, these students would be more inclined to be less consistent compared to the present study, which also proved to be the case.

When comparing the results from the two Swedish national subcorpora, we find no statistically significant difference, and the degree of consistency does then not appear to have changed notably in the year separating the SWICLE and the SUSEC. Given that an increase in preference for an alternative variety would result in a decrease in consistency, the slight, albeit not statistically significant, increase in inconsistency found in SUSEC could then be explained by the slight, but again not statistically significant, increase in preference for American English spelling found in SUSEC. The slight differences found when comparing the SWICLE data and the SUSEC data might, however, also be explained by the varying length of the essays in the two corpora. Nevertheless, the alleged Americanization of Swedish society taking place during the last decades does not appear to have notably affected either the preference for British English or the degree of consistency of the university students included in the study.

4.2. The larger European context

4.2.1. Preferred variety: an investigation of possible differences across nations

In the Bulgarian and Italian national subcorpora, as well as in the two Swedish national subcorpora, a statistically significant predominance of British English spelling was found, as shown in Figure 7. The greatest preference for British English spelling can be found in the Italian subcorpus. Of the investigated words chosen for this study, 97% of the words were spelled in accordance with British English conventions, while only 3% of the words were spelled according to American English standards; the difference was highly statistically significant (significance level $p < 0.001$). This can be compared to the Bulgarian subcorpus (87% BrE spelling, 13% AmE spelling), to SWICLE (82% BrE, 18% AmE) and to SUSEC (78% BrE, 22% AmE). However, although the preference for British English spelling was statistically significant within each national subcorpus, the differences found between the four national subcorpora were not all statistically significant. This is the case for the difference between SWICLE and the Bulgarian subcorpus and, as mentioned earlier, between SUSEC and SWICLE; the rest proved to be statistically significant.

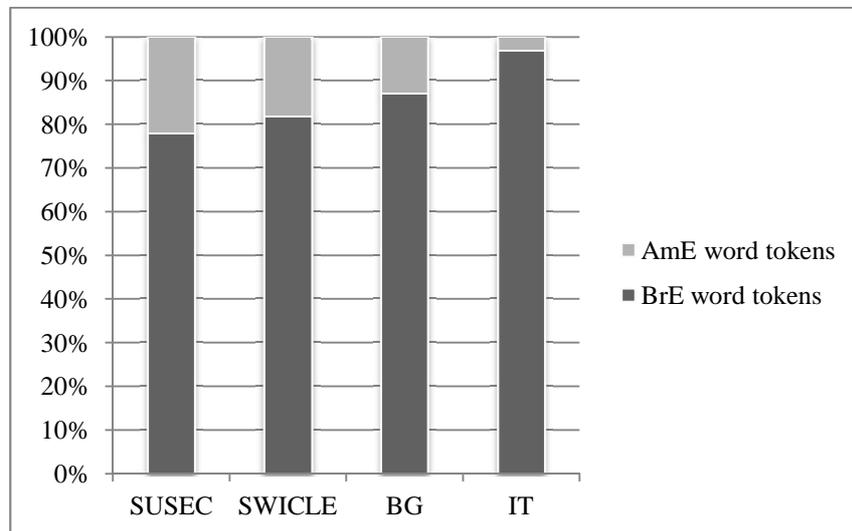


Figure 7. Relative frequency of British English and American English word tokens

In terms of relative frequency of the number of essays in which only British English or American English spelling was used, the Italian subcorpus still showed the largest predominance of British English spelling in percentage as 96% out of the students used British English spelling, whereas 4% of the students used American English spelling. This can be compared to the Bulgarian subcorpus (86% BrE, 14% AmE), to SWICLE (83% BrE, 17% AmE), and to SUSEC (77% BrE, 23% AmE), as shown in Figure 8. There was no statistically significant difference between the Swedish national subcorpora and the Bulgarian subcorpus.

Common for the data from all four national subcorpora is that there is a clear, statistically significant preference for British English spelling, both in terms of number of word tokens and number of essays in which British English or American English spelling is used consistently. When comparing the national subcorpora, we find the most prominent preference for British English in the Italian subcorpus. The connection between media preferences (and thereby exposure to a language variety) and positive attitudes towards that variety reported in the literature (Mobärg, 1999: 68) could serve as one explanation as to why the Italian subcorpus stands out. Since there is a strong predominance of British

English in the Italian school system, in addition to the fact that next to all TV shows are dubbed (Pulcini, 1997: 81), the Italian students are less likely to be exposed to American English compared to the Bulgarian and Swedish students, and British English would then be the obvious choice of variety.

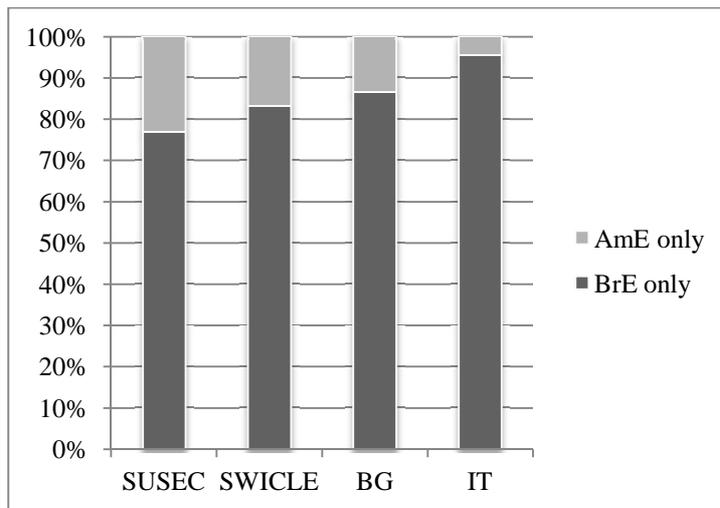


Figure 8. Relative frequency of the essays in which British English or American English spelling is used consistently

Apart from the more frequent occurrence of American English word tokens in SUSEC compared to the Bulgarian subcorpus, no statistically significant differences are found between the Swedish national subcorpora and the Bulgarian national subcorpus. Thus, there is seemingly no considerable difference between the Swedish and Bulgarian national subcorpora, which might be explained by similarities between the countries with respect to media influence and a recent history of British English predominance in education (cf., e.g., Altenberg, 2002 and Blagoeva, 2002; for an overview of EFL teaching in Sweden and Bulgaria respectively).

4.2.2. Degree of consistency: an investigation of possible differences across nations

The only statistically significant difference found was between the subcorpus including the most consistent essays (Italy) and the subcorpus including the highest occurrence of inconsistent essays (SUSEC). For the SUSEC, results show that 88% of the student essays were consistent, while 12% were inconsistent, which can be compared to the Italian subcorpus where 98% were consistent, and 2% inconsistent, as shown in Figure 9. For the SWICLE, 94% of the essays were consistent and 6% were inconsistent and for the Bulgarian subcorpus of ICLE, 95% were consistent, and 5% inconsistent.

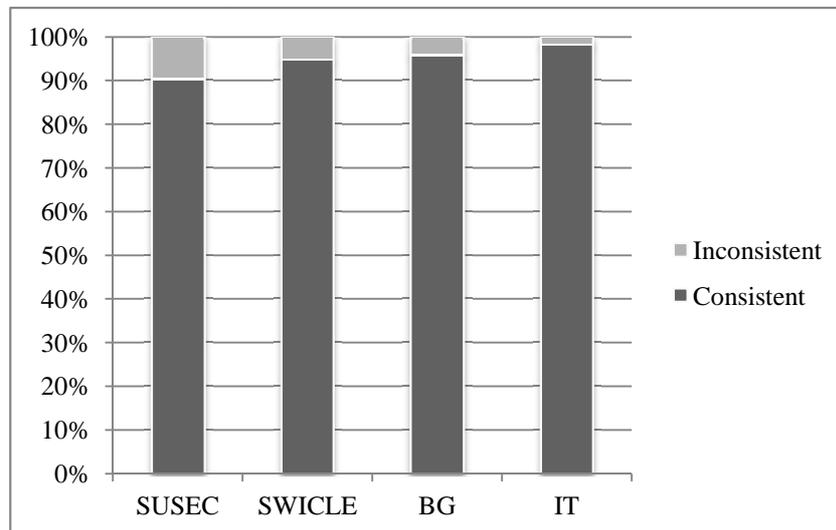


Figure 9. Relative frequency of inconsistent and consistent essays

Again, there appears to be a correlation between a greater preference for an alternative variety (in this case American English) and a higher tendency for inconsistent use of the varieties. Consequently, a lower preference for an alternative spelling variety (American English) would then entail a higher degree of consistency, which was true for the Italian essays. The Swedish essays showed a higher preference for American English and would therefore be expected to be less consistent, as proved to be the case. This correlation would then also serve as an explanation as to why there was no statistically significant difference between the

Swedish national subcorpora and the Bulgarian national subcorpus with regard to the frequency of occurrence of the number of (in)consistent essays.

5. Conclusion

The results of the present study show that British English is the preferred variety in terms of spelling; this was the case for all the investigated student levels and the preference had not changed significantly over time. Hence, although almost 20 years have passed since Modiano (1993) concluded that Swedish university students of English show an attitudinal preference for British English, the actual preference of this variety does not seem to have changed notably over the years separating SWICLE from SUSEC. When comparing the result from the Swedish national subcorpora to the Bulgarian and Italian national subcorpora, an even stronger preference for British English spelling was found in the Italian subcorpus, while the results from the Bulgarian subcorpus did not prove to differ significantly from the Swedish results. The findings could possibly be explained by the traditional preference for British English found in all three countries, and the greater predominance of British English found in the Italian essays might be explained by the lack of American influence from the media in Italy.

In terms of degree of consistency, the students were generally consistent in their choice of one variety. Nonetheless, the lower number of inconsistent essays and the lower preference for American English spelling found in the Italian subcorpus compared to the higher number of inconsistent essays and the higher preference for American English spelling found in SUSEC could point to a possible correlation between an increasing preference for an alternative variety (American English) and an increasing tendency for inconsistent use of the varieties. It is, however, important to bear in mind that the results cannot be generalizable to the full subcorpora of the countries, other than for the words selected to represent the investigated categories; moreover, since the study does not investigate attitudinal preference, the conclusions drawn do not take into account that words could be misspelled or not deliberately chosen.

Apart from adding to our knowledge of how English is used by advanced EFL students, these findings also have implications for EFL

teaching. Since there appears to be a strong preference for British English among university students – many of whom are prospective EFL teachers – in a context where British English is no longer the only accepted variety, it would seem important for EFL teachers and teacher educators to be aware of this strong preference. A conscious effort would then be needed to allow for a more tolerant view with respect to what is considered to be incorrect and correct use of English in an EFL context by also acknowledging the existence of other varieties of English.

In order to gain more knowledge of how non-native speakers of English use the language, both quantitative and qualitative studies are needed. Such studies could, for example, further investigate to what extent EFL students tend to (consistently) adhere to British or American English standards when it comes to grammar and vocabulary, and whether these possible preferences can be found to correlate with students' spelling preferences. Finally, since non-native speakers of English now outnumber native speakers of English (McArthur, 2006: 465), a few questions remain: to whom does the English language really belong, does it matter, and will it matter in the future?

Acknowledgements

I would like to thank Professor Annelie Ädel for her kindness and patience in introducing me to corpus-based research, and for her invaluable feedback during the course of this project. My thanks are also due to Associate Professor Britt Erman for compiling the SUSEC, to Dr. Gabriella Rundblad for providing the native speaker material of the SUSEC, and to the anonymous reviewers for their very useful comments.

References

- Alftberg, A.-K. 2009. *British or American English? Attitudes, Awareness and Usage among Pupils in Secondary School*. Available from Digitala Vetenskapliga Arkivet: URI: urn:nbn:se:hig:diva-5545
- Anthony, L. 2012. AntConc (Version 3.2.4) [Computer Software]. Tokyo: Waseda University. Available from <http://www.antlab.sci.waseda.ac.jp/>
- Altenberg, B. 2002. "The status of English in Sweden". In S. Granger, E. Dagneaux & F. Meunier (eds.), *International Corpus of Learner*

- English* (pp. 140-145). Louvain-la-Neuve, Belgium: Presses Universitaires de Louvain.
- Blagoeva, R. 2002. "The status of English in Bulgaria". In S. Granger, E. Dagneaux & F. Meunier (eds.), *International Corpus of Learner English* (pp. 81-86). Louvain-la-Neuve, Belgium: Presses Universitaires de Louvain.
- Brown, K. 2002. "Ideology and context: world Englishes and EFL teacher training". *World Englishes*, 21 (3), 445-448.
- . 1995. "World Englishes: To teach or not to teach?" *World Englishes*, 14 (2), 233-245.
- Griffin, J. L. 2001. "Global English infiltrates Bulgaria". *English Today*, 17 (4), 51-60.
- Hoffmann, C. 2000. "The spread of English and the growth of multilingualism with English in Europe". In J. Cenoz & U. Jessner (eds.), *English in Europe: the Acquisition of a Third Language* (pp. 1-21). Clevedon, UK: Multilingual Matters.
- Johannesson, N.-L. 1986. *English Language Essays: Investigation Method and Writing strategies*. Stockholm, Sweden: Department of English, Stockholm University.
- McArthur, T. 2006. "The printed word in the English speaking world". In K. Bolton & B.B. Kachru (eds.), *World Englishes: Critical Concepts in Linguistics. Vol 3* (pp. 463-473). London, UK: Routledge.
- Mobärg, M. 2002. "RP or GA? On Swedish school students' choice of English pronunciation". In M. Modiano (ed.), *Studies in Mid-Atlantic English* (pp. 119-131). Gävle, Sweden: Hofors tryckeri.
- . 1999. "School goes to Hollywood: Attitudes towards British and American English among Swedish school students". In C. Paradis (ed.), *Recent Trends in the Pronunciation of English. Social, Regional and Attitudinal Aspects* (pp. 49-70). Stockholm, Sweden: Almqvist & Wiksell International.
- Modiano, M. (ed.) 2002. *Studies in Mid-Atlantic English*. Gävle, Sweden: Hofors tryckeri.
- . 1996. "The Americanization of Euro-English". *World Englishes*, 15 (2), 207-215.
- . 1993. "American English and higher education in Sweden". *American Studies in Scandinavia* 25, 37-47.

- , & Söderlund, M. 2002. “Swedish upper secondary school students and their attitudes towards AmE, BrE, and Mid-Atlantic English”. In M. Modiano (ed.), *Studies in Mid-Atlantic English* (pp. 147-171). Gävle, Sweden: Hofors tryckeri.
- O’Reilly, L. 1998. “English language cultures in Bulgaria: a linguistic sibling rivalry?” *World Englishes*, 17 (1), 71-84.
- Prat Zagrebelsky, M. T. 2002. “The status of English in Italy”. In S. Granger, E. Dagneaux & F. Meunier (eds.), *International Corpus of Learner English* (pp. 106-110). Louvain-la-Neuve, Belgium: Presses Universitaires de Louvain.
- Pulcini, V. 1997. “Attitudes toward the spread of English in Italy”. *World Englishes*, 16 (1), 77-85.
- Tottie, G. 2002. *An Introduction to American English*. Malden, MA: Blackwell Publishers Inc.
- Trudgill, P., & Hannah, J. 1994. *International English. A guide to the varieties of Standard English* (3rd ed.). London, UK: Edward Arnold.
- Westergren Axelsson, M. 2002. “‘Refined’ or ‘relaxed’ English pronunciation: Usage and attitudes among Swedish university students”. In M. Modiano (ed.), *Studies in Mid-Atlantic English* (pp. 132-145). Gävle, Sweden: Hofors tryckeri.

Appendix A

Rate of occurrence of word tokens spelled in accordance with British English or American English conventions for all the investigated subcorpora, followed by rate of occurrence for the essays in which British English or American English spelling conventions are used consistently.

| SUSEC subcorpora | Total number of word tokens | Number of BrE word tokens | BrE word tokens per 100,000 words | Number of AmE word tokens | AmE word tokens per 100,000 words |
|------------------|-----------------------------|---------------------------|-----------------------------------|---------------------------|-----------------------------------|
| First term | 117,175 | 85 | 73 | 36 | 31 |
| Second term | 182,829 | 94 | 51 | 55 | 30 |
| Third term | 417,772 | 245 | 59 | 37 | 9 |
| Fourth term | 192,548 | 103 | 53 | 23 | 12 |
| Total | 910,324 | 527 | 58 | 151 | 17 |

| National subcorpora | Total number of word tokens | Number of BrE word tokens | BrE word tokens per 100,000 words | Number of AmE word tokens | AmE word tokens per 100,000 words |
|---------------------|-----------------------------|---------------------------|-----------------------------------|---------------------------|-----------------------------------|
| SUSEC | 910,324 | 527 | 58 | 151 | 17 |
| SWICLE | 206,015 | 209 | 101 | 47 | 23 |
| BG | 203,077 | 153 | 75 | 23 | 11 |
| IT | 229,412 | 469 | 204 | 16 | 7 |

| Total | 1,548,828 | 1,358 | 88 | 237 | 15 |
|------------------|------------------------|---|--------------------------|---|--------------------------|
| SUSEC subcorpora | Total number of essays | Number of essays with BrE spelling only | BrE essays per 10 essays | Number of essays with AmE spelling only | AmE essays per 10 essays |
| First term | 117 | 44 | 3.8 | 12 | 1.0 |
| Second term | 90 | 35 | 3.9 | 15 | 1.7 |
| Third term | 62 | 35 | 5.6 | 8 | 1.3 |
| Fourth term | 17 | 9 | 5.3 | 2 | 1.2 |
| Total | 286 | 123 | 4.3 | 37 | 1.3 |

| National subcorpora | Total number of essays | Number of essays with BrE spelling only | BrE essays per 10 essays | Number of essays with AmE spelling only | AmE essays per 10 essays |
|---------------------|------------------------|---|--------------------------|---|--------------------------|
| SUSEC | 286 | 123 | 4.3 | 37 | 1.3 |
| SWICLE | 363 | 119 | 3.3 | 24 | 0.7 |
| BG | 302 | 89 | 2.9 | 14 | 0.5 |
| IT | 397 | 189 | 4.8 | 9 | 0.2 |
| Total | 1,348 | 520 | 3.9 | 84 | 0.6 |

Appendix B

Rate of occurrence of inconsistent and consistent essays for all the investigated subcorpora.

| SUSEC subcorpora | Total number of essays | Number of consistent essays | Consistent essays per 10 essays | Number of inconsistent essays | Inconsistent essays per 10 essays |
|------------------|------------------------|-----------------------------|---------------------------------|-------------------------------|-----------------------------------|
| First term | 59 | 56 | 9.5 | 3 | 0.5 |
| Second term | 55 | 50 | 9.1 | 5 | 0.9 |
| Third term | 51 | 43 | 8.4 | 8 | 1.6 |
| Fourth term | 17 | 11 | 6.5 | 6 | 3.5 |
| Total | 182 | 160 | 8.8 | 22 | 1.2 |

| National subcorpora | Total number of essays | Number of consistent essays | Consistent essays per 10 essays | Number of inconsistent essays | Inconsistent essays per 10 essays |
|---------------------|------------------------|-----------------------------|---------------------------------|-------------------------------|-----------------------------------|
| SUSEC | 182 | 160 | 8.8 | 22 | 1.2 |
| SWICLE | 152 | 143 | 9.4 | 9 | 0.6 |
| BG | 108 | 103 | 9.5 | 5 | 0.5 |
| IT | 202 | 198 | 9.8 | 4 | 0.2 |
| Total | 644 | 604 | 9.4 | 40 | 0.6 |