The handle http://hdl.handle.net/1887/54849 holds various files of this Leiden University dissertation.

**Author:** Ebner, C.
**Title:** Proper English Usage: a sociolinguistic investigation of attitudes towards usage problems in British English
**Issue Date:** 2017-09-05
7. Current Usage Attitudes in England: the Online Questionnaire (Part One)

7.1. Introduction

Having outlined my methodological approach in Chapter 5 and having described the usage problems investigated in this study in Chapter 6, I will now turn to the data analysis and present my findings in the following three chapters. In this chapter, I will discuss the results of the first part of the online questionnaire in detail. The second part, which includes the language statements and the open question, will be discussed in Chapter 8, while the data collected in the interview sessions is presented in Chapter 9. Since I have chosen a mixed-methods approach consisting of both indirect and direct elicitation techniques as well as quantitative and qualitative data for the analysis of current usage attitudes in England, it is important to keep in mind the purpose of each test that forms part of this approach as well as their respective methods since these could have implications for my interpretation of the results.

Before tackling the analysis, I will describe the data obtained through the questionnaire and how it was prepared for the analysis, before providing the details on the statistical tests used for the analysis. The first part of the online questionnaire contains the eleven investigated usage problems. Each usage problem will be discussed in detail as follows: first I will present corpus data for evidence of usage, before any sociolinguistic variation in the attitudes of the questionnaire respondents will be identified. For the corpus analysis, I will draw on the British National Corpus (BNC) and whenever a comparison with American English is made, on the Corpus of Contemporary American English (COCA) (see § 5.5). Secondly, the contextual preference of the usage problem investigated will be demonstrated on the basis of the questionnaire respondents’ contextual acceptability judgments, before I discuss any possible correlations between the respondents’ degree of judgment certainty as well as
on what grounds the acceptability judgment was made (see § 5.3.1). Did respondents base their judgments, i.e. whether a particular feature was considered acceptable or unacceptable, on self-reported knowledge of a rule or on their intuition? Whether the overall acceptability judgment, or the question whether the questionnaire respondents found a particular usage problem acceptable or not, correlates with a specific degree of certainty or judgment basis could provide a new perspective on the usage debate in the sense that the respondents’ linguistic security with regard to usage problems is tested. This perspective made it possible to explore whether there is a difference between respondents expressing a prescriptive or a descriptive attitude with regard to the certainty about their usage judgment and whether they based their judgment on a rule they had learnt, say, in school, or rather a gut feeling. The third step adds a qualitative dimension to this quantitative test in that it comprises an analysis of additional comments made by the survey participants on the usage problem; this will be achieved by identifying recurring and prominent themes in the comments. A qualitative dimension will contribute positively to the understanding of usage attitudes as such comments can provide further insights into a simple acceptable/unacceptable judgment.

To position the general public’s attitudes in the usage debate, it is also necessary to include the views held by usage guide authors. This will be achieved through a survey of the HUGE database. As mentioned in Section 6.1, I will use a slightly modified version of Yáñez-Bouza’s (2015) tripartite categorisation to determine a usage problem’s treatment as either “advocated”, “neutral” or “criticised”. The HUGE survey is, however, restricted to those usage problems which are included in the database.

In order to be able to carry out a comparison of usage attitudes with Mittins et al.’s (1970) *Attitudes to English Usage* in Chapter 9, the average
acceptability ratings of the eleven investigated usage problems will be performed, which will be calculated by averaging all contexts indicating an acceptable judgment. This type of calculation was also used by Mittins et al. (1970) and consequently will allow me to make a careful comparison between the Mittins study and my own investigation of usage attitudes in the area studied. In interpreting the results of the comparison, however, I will take into account the slightly different methodologies used and populations surveyed. While Mittins et al.’s study focussed on students and educationalists, only a small number of members of the general public were included in their sample. The focus of my study is, however, on the attitudes of a sample of the general public. Hence, the results of such a comparison should only be understood as an indicator of possible tendencies of changing usage attitudes.

7.2. Results of the Questionnaire

While the online questionnaire was completed by 230 respondents from all over England, it is known that online questionnaires are prone to a self-selection bias (Olsen, 2008, pp. 809–810; Bethlehem, 2010, p. 162; Toepoel, 2016, p. 200). This bias is also indicative of specific traits shared by the respondents, such as a general interest in language or eagerness to make one’s opinion public. Not only does the self-selection bias influence the questionnaire sample, the means of how the questionnaire was made available equally influences the composition of the survey sample. Posting the questionnaire on various social media sites, such as Facebook and Twitter, I tried to increase its visibility by targeting younger age groups. However, in order to reach older members of the general public who might not be members of these social media sites, a call for participants was included in a newsletter of the University of the Third Age (U3A), which attracted a lot of interest from older informants. Other means of making the questionnaire more widely available included
hanging out leaflets at the annual English Grammar Day held on 4 July 2014 at the British Library, as well as putting up posters at various locations in London during my research stay at Queen Mary University of London between January and March 2014. The self-selection bias in combination with the means of questionnaire distribution clearly had an effect on the resulting survey sample and its representativeness of the survey population, as can be seen from Table 7.1 below.

Table 7.1 Full sample breakdown

<table>
<thead>
<tr>
<th>Gender</th>
<th>18–25</th>
<th>26–30</th>
<th>31–40</th>
<th>41–60</th>
<th>over 60</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>10</td>
<td>8</td>
<td>13</td>
<td>10</td>
<td>27</td>
<td>68</td>
</tr>
<tr>
<td>Female</td>
<td>37</td>
<td>15</td>
<td>23</td>
<td>27</td>
<td>60</td>
<td>162</td>
</tr>
<tr>
<td></td>
<td>47</td>
<td>23</td>
<td>36</td>
<td>37</td>
<td>87</td>
<td>230</td>
</tr>
</tbody>
</table>

From the table above, the overrepresentation of female and older respondents becomes obvious. In order to minimize these biases in the data analysis, I proportionally and randomly stratified the survey sample according to the 2011 Census of England and Wales for the two social variables age and gender. This resulted in the survey breakdown illustrated in Table 7.2.

Table 7.2 Proportionally stratified sample breakdown

<table>
<thead>
<tr>
<th>Gender</th>
<th>18–25</th>
<th>26–30</th>
<th>31–40</th>
<th>41–60</th>
<th>over 60</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>8</td>
<td>8</td>
<td>13</td>
<td>9</td>
<td>17</td>
<td>55</td>
</tr>
<tr>
<td>Female</td>
<td>9</td>
<td>8</td>
<td>14</td>
<td>9</td>
<td>17</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>16</td>
<td>27</td>
<td>18</td>
<td>34</td>
<td>112</td>
</tr>
</tbody>
</table>
In the 2011 Census, 49 per cent of the English population were male, and 51 per cent were female. Hence, the proportional stratification resulted in 57 female and 55 male participants to be randomly selected from the set of completed questionnaires. Consequently, the sample size was reduced from 230 questionnaires to 112. Age constituted a more complex social variable as different age categories were applied in the 2011 England Census and my attitudes survey. While the census made use of nine age categories ranging from ‘Age 18 to 19’ to ‘Age 90 and over’, I decided to use only five age categories, as already illustrated in Table 7.2 above, which was also due to the regulations of the ethics committee. According to these regulations, the survey participants needed to be over the age of 18. The census data was combined into five strata which corresponded roughly with the categories applied in the questionnaire. I decided to merge the age categories 41–50 and 51–60 used in the online questionnaire into one category, not only to reflect the census age category more closely, but also to compensate for the lack of male respondents falling into these two age categories. Because, all in all, only nine men from these age categories completed the online questionnaire, I decided to include all nine male participants in the proportionally stratified sample for the age group in question. The underrepresentation of middle-aged respondents and overrepresentation of younger and older age groups, which is probably due to the self-selection bias and the means of survey distribution mentioned above, led to a few differences between the England census data, the Proper English Usage (PEU) survey sample and the proportionally stratified sample. Therefore, some age groups, especially the 41–60 one, are underrepresented. The differences between the 2011 Census data, the full PEU survey sample and the stratified sample are indicated in the comparison presented in Table 7.3 below.
Table 7.3 Comparative overview of samples and the 2011 Census data (%)

<table>
<thead>
<tr>
<th>Age strata</th>
<th>England Census 2011</th>
<th>PEU full survey sample</th>
<th>PEU stratified sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>stratum 1</td>
<td>18–25</td>
<td>11.9</td>
<td>20.4</td>
</tr>
<tr>
<td>stratum 2</td>
<td>26–30</td>
<td>8.8</td>
<td>10.0</td>
</tr>
<tr>
<td>stratum 3</td>
<td>31–40</td>
<td>26.3</td>
<td>15.7</td>
</tr>
<tr>
<td>stratum 4</td>
<td>41–60</td>
<td>24.7</td>
<td>16.1</td>
</tr>
<tr>
<td>stratum 5</td>
<td>over 60</td>
<td>28.4</td>
<td>37.8</td>
</tr>
<tr>
<td>Total</td>
<td>100 %</td>
<td>100 %</td>
<td>100 %</td>
</tr>
</tbody>
</table>

As Table 7.3 shows, stratifying the survey sample benefits representativeness as the stratified sample reflects England’s population more closely than the PEU full survey sample. The underrepresentation of the age group 41–60, which shows a 8.3 per cent difference between the UK Census and the PEU stratified survey, consequently led to an overrepresentation of other age groups, especially the two containing the youngest participants of the survey.

As part of the social and educational background information requested at the end of the questionnaire, I asked the participants to provide information on their ethnicity as well as on their first language(s). I chose to use similar if not the identical labels as used in the 2011 UK Census, as shown in Table 7.4 below. While the majority of the stratified sample stated their ethnicity as “White”, namely 104 participants, a few questionnaire respondents had a different ethnicity, which can be seen in Table 7.4. As discussed in Chapter 5, England’s population shows different degrees of ethnic diversity. A comparison with the ethnic background of the stratified sample respondents to the composition of the population of the Golden Triangle, as illustrated in Figure 5.2 above, shows that the stratified sample resembles the regions South England and East of England more closely than the highly ethnically diver-
sified capital London. Ethnic diversity, moreover, often also entails multilingualism. In Table 7.5, an overview of the linguistic background of the stratified sample respondents is given.

Table 7.4 PEU stratified sample respondents’ ethnicities in percentages (raw figures in brackets)

<table>
<thead>
<tr>
<th>Ethnicity</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>92.9 (104)</td>
</tr>
<tr>
<td>Mixed/Multiple Ethnic Groups</td>
<td>1.8 (2)</td>
</tr>
<tr>
<td>(e.g. White and Black African ...)</td>
<td></td>
</tr>
<tr>
<td>African/Caribbean/Black British</td>
<td>1.8 (2)</td>
</tr>
<tr>
<td>Other ethnic groups</td>
<td>3.6 (4)</td>
</tr>
<tr>
<td>Total</td>
<td>100 (112)</td>
</tr>
</tbody>
</table>

Table 7.5 PEU respondents’ linguistic background in percentages (raw figures in brackets)

<table>
<thead>
<tr>
<th>First language(s)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>English</td>
<td>86.6 (97)</td>
</tr>
<tr>
<td>German</td>
<td>6.3 (7)</td>
</tr>
<tr>
<td>bilingual</td>
<td></td>
</tr>
<tr>
<td>(English/Spanish, English/German, English/Jamaican Patois, English/Finnish/Norwegian)</td>
<td>3.6 (4)</td>
</tr>
<tr>
<td>Italian</td>
<td>1.8 (2)</td>
</tr>
<tr>
<td>Danish</td>
<td>0.9 (1)</td>
</tr>
<tr>
<td>Polish</td>
<td>0.9 (1)</td>
</tr>
<tr>
<td>Total</td>
<td>100 (112)</td>
</tr>
</tbody>
</table>

As the ethnicity breakdown of the stratified sample respondents reveals, the majority stated that English was their first language, followed by seven German native-speakers and four bi- or multilingual speakers. Speakers of other languages who live in England and participate in society are considered part of the speech community described in this study. Hence, restricting the sample to English native speakers would only cause a misrepresentation of the speech community.
Furthermore, the education level of the 112 questionnaire respondents constitutes an important factor as it is one of the four social variables investigated in this study. As can be seen from Table 7.6 below, the majority of the questionnaire respondents seem to be well-educated. A comparison with the 2011 England Census, which shows a greater diversity between the different education levels with roughly 28 per cent of the English population falling into the highest level described as “degree or above” (ONS, 2012b), indicates that this difference is caused by a sampling error most likely due to the self-selection bias of online questionnaires and the chosen sampling techniques.

Table 7.6 PEU respondents’ education level in percentages (raw figures in brackets)

<table>
<thead>
<tr>
<th>Education Level</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCSEs</td>
<td>0.9 (1)</td>
</tr>
<tr>
<td>GCE A levels</td>
<td>2.7 (3)</td>
</tr>
<tr>
<td>Other certificate (e.g. HND)</td>
<td>9.8 (11)</td>
</tr>
<tr>
<td>Degree or above</td>
<td>86.6 (97)</td>
</tr>
</tbody>
</table>

It is therefore important to bear in mind the rather homogenous composition of the stratified sample with regard to level of education when investigating and discussing usage attitudes in this study. In order to enable a more meaningful comparison between education levels possible, I divided the sample presented in Table 7.6 above into two groups: university-educated and non-university-educated respondents. An overview of this categorisation can be found in Table 7.7. The combination of the three education levels GCSE, GCE A levels and other certificate (e.g. HND) seems necessary given the low number of questionnaire respondents with lower levels of education.

Table 7.7 PEU categorisation: university-educated vs non-university-educated (raw figures in brackets)

<table>
<thead>
<tr>
<th></th>
<th>non-university-educated</th>
<th>university-educated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degree or above</td>
<td>13.4 (15)</td>
<td>86.6 (97)</td>
</tr>
</tbody>
</table>
Being a marker of social class membership as well, level of education will be used in this study as a social variable (Block, 2014, p. 3). Since social class is an important yet highly complex and disputed variable, a more socially stratified sample would be needed to be able to conduct a thorough analysis of social class. A list of all professions is provided in Appendix H. Although the sample contains a relatively large number of teachers and other language professionals, it cannot be assumed that their attitudes towards disputed usage features differ. As I have argued above (see § 3.3.1), language awareness is a crucial component of usage attitudes. This component does not necessarily have to be equated with the level of education, or a speaker’s profession. It cannot be assumed, for instance, that all teachers are aware of disputed usages, as indeed will be confirmed in the analysis. The reason for different degrees of awareness lies in the fact that teacher training as well as the teaching of English has changed considerably in England (see § 2.3.1). This is also evident from a government survey of teachers published in 1998 which concluded that “younger teachers had generally not been taught grammar explicitly as part of their own education” (The Grammar Papers, 1998, p. 26). Nonetheless, the professions of the questionnaire respondents are presented together with other background information in the qualitative analysis of comments.

The proportionally stratified sample will be conducted in the Statistics Programme SPSS 23. Since my data is not normally distributed, I will make use of a range of non-parametric tests to determine any correlations between the elicited data. Making use of the Mann-Whitney U-test, which compares the differences between ordinal or continuous dependent variables and independent variables, I will first identify any sociolinguistic significant correlation between the dependent variable acceptability rating and the independent social variable age (all $p$-values are two-tailed). To identify any significant correlations between acceptability ratings and the independent social variable
gender, I will make use of a chi-squared test (χ²) to refute the null hypothesis according to which there is no difference between the acceptability judgments of women and men. In case the null hypothesis, i.e. that there is no difference between any specific social groups and their usage judgments, is refuted, an Odds Ratio (OR) will be calculated to identify whether one of the two traditional gender categories “male” vs “female” is more likely to find a particular usage problem more acceptable than the other. Whether nativeness plays a role in the variation of usage attitudes will be investigated by making use of Spearman’s correlation coefficient, which will show any significant differences between the two groups “native” and “non-native” speakers and the acceptability ratings, as to whether a particular stimulus sentence was considered acceptable or unacceptable. Spearman’s correlation tests will also be used to identify any possible differences between acceptability ratings and level of education, i.e. university-educated or non-university-educated respondents. I will also state the effect size for each of the investigated variables, which describes further the strength of the identified phenomenon. In order to assess the influence of all four social variables on acceptability judgments and to identify a possible covariance between these independent variables, I conducted a binary logistic regression analysis for each usage problem (Field, 2013, p. 761). This analysis will enable me to see whether the social variables age, gender, nativeness and level of education co-vary in their influence on the dependent variable. The binary logistic regression analysis is used to create a model which predicts the presence or absence of a specific characteristic (Acton et al., 2009, p. 258). In my study, this specific characteristic constitutes an acceptable judgment made by the respondent as to whether a stimulus sentence was considered acceptable or not acceptable. Since I am interested in

3 Effect sizes were calculated by making use of the effect size calculator available at http://www.polyu.edu.hk/mm/effectsizefaqs/calculator/calculator.html.
determining the influence of the four social variables investigated in this study, i.e. age, gender, nativeness and level of education (university-educated or non-university-educated), on the outcome variable, I decided to use the forced entry method in which all predictors, i.e. the social variables, are included in the model (Field, 2013, p. 322). In order to compare the categorical predictors to a referent, I determined that age groups are compared to the highest age group, non-native speakers are compared to native speakers, male speakers to female speakers and non-university-educated speakers to university-educated speakers. The reason why I chose these reference groups is not only previous studies (cf. Mittins et al., 1970; Trudgill, 1974; Albanyan & Preston, 1998), but also the results of the non-parametric tests conducted prior to the binary logistic regression analysis. Having provided a detailed overview of how the analysis of the questionnaire will be conducted, I will now move on to present the analysis for the eleven usage problems.

7.2.1. The usage problems

7.2.1.1. Different from/than/to

Having described the prescriptivists’ issue with different than in Section 6.2 above, I will first discuss the corpus evidence of different from/than/to, before presenting the sociolinguistic analysis of the elicited attitude data and a qualitative analysis of comments. To complete the analysis, an overview of the treatment of this particular usage feature in the HUGE database is provided.

As described in the preceding chapter, different to and different than have been found to vary in frequency of usage in British and American English. A corpus study conducted by Busse and Schröder (2010b, pp. 97–98) showed that different to was the second most frequent variant of the different from/than/to issue in British English, while different than was in American English. While differently than shows a strikingly high frequency rate of 1,087
tokens in COCA, it produced merely 10 tokens in the BNC, as can be seen in Table 7.8 below. Looking into these differences in more detail by comparing the usage frequencies of *different than* in both corpora, it does not come as a surprise to find this particular variant more frequently in COCA, with 4,569 tokens, which correspond with the highest normalised frequency of 20.45 tokens per million words in spoken section of the corpus. In comparison to COCA, the BNC includes only a total of 50 tokens for *different than*. With a normalised frequency of 1.81 tokens per million words the spoken subsection of the BNC also shows the highest frequency rate.

Table 7.8 Overview of *different than* and *differently than* in BNC and COCA

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Different than (n=50)</td>
<td>Freq. per mil</td>
<td>18</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>10</td>
</tr>
<tr>
<td>Differently than (n=10)</td>
<td>Freq. per mil</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>COCA</th>
<th>Spoken</th>
<th>Fiction</th>
<th>Magazine</th>
<th>Newspaper</th>
<th>Acad.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Different than (n=4,569)</td>
<td>Freq. per mil</td>
<td>2,237</td>
<td>480</td>
<td>470</td>
<td>873</td>
</tr>
<tr>
<td>Differently than (n=1,087)</td>
<td>Freq. per mil</td>
<td>339</td>
<td>95</td>
<td>183</td>
<td>204</td>
</tr>
</tbody>
</table>

What needs to be borne in mind, however, are the differences between the two corpora in terms of size and collection dates (see § 5.5). Therefore, looking into normalised frequencies and reporting the subsection in which *different(ly)* *than* is most frequently found can aid a better understanding of actual usage.

Confirming Busse and Schröder’s findings of secondary preference differences between British and American English, a complementary corpus search of *different to* shows that this variant is indeed more frequently found...
in the BNC, as the spoken subsection of the corpus shows the highest frequency of 13.35 tokens per million words, which corresponds with a much lower frequency in COCA, which only shows 1.52 tokens per million words in the same subsection.

As the focus of this study, however, is on a sociolinguistic investigation of usage attitudes, I will attempt to show whether any significant correlations can be found between the social variables investigated and the questionnaire respondents’ acceptability judgments, i.e. whether respondents found the stimulus sentence (S1. *The Americans look at this differently than the British*) acceptable or not. Looking at the social variables age, gender, nativeness and level of education, I found the only variable which shows a significant correlation with the acceptability ratings is age. While gender ($\chi^2 (1) = 0.55, p = .814$), nativeness ($r_s = -.081, p = .395$) and education level ($r_s = .059, p = .534$) show no significant difference in acceptability ratings, applying the Mann-Whitney $U$-test showed that age has an effect on attitudes towards the use of *differently than* in that younger respondents are more likely to find the construction acceptable ($Mdn = 31–40$-year-olds), while older respondents tend to find it rather unacceptable ($Mdn = 41–60$-year-olds, $U = 1128$, $p = .024, r = -.21$). The small effect size, however, weakens the assumption that the differences between acceptability ratings vary significantly. Yet, smaller effect sizes are not unusual in small samples such as the one used in this study (Field, 2013, pp. 79–80). A binary logistic regression analysis was conducted to predict an acceptability judgment of *differently than* using the social variables age, gender, level of education (university-educated or non-university-educated) and nativeness as predictors. If none of the social variables are included in the model, the model would make predictions with an accuracy rate of 60.7 per cent. The overall prediction accuracy was increased
to 66.1 per cent in the model including all predictors, and the results of the analysis are presented in Table 7.9.

Table 7.9 Results of binary logistic regression: differently than

<table>
<thead>
<tr>
<th>Included</th>
<th>95 % CI for exp b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B(SE)</td>
</tr>
<tr>
<td>Nativeness</td>
<td>0.25 (0.77)</td>
</tr>
<tr>
<td>Gender</td>
<td>0.12 (0.41)</td>
</tr>
<tr>
<td>Age (18–25)</td>
<td>1.20 (0.64)</td>
</tr>
<tr>
<td>Age (26–30)</td>
<td>0.85 (0.63)</td>
</tr>
<tr>
<td>Age (31–40)</td>
<td>1.35 (0.59)*</td>
</tr>
<tr>
<td>Age (41–60)</td>
<td>1.03 (0.62)</td>
</tr>
<tr>
<td>Level of education</td>
<td>-0.16 (0.61)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.39 (0.44)</td>
</tr>
</tbody>
</table>

Note $R^2 = .06$ (Hosmer & Lemeshow), Model $\chi^2(7) = 8.683, p = .276, p < .05^*$

Testing the fit of the model showed that the proposed model produced a good fit model ($\chi^2(7) = 8.683, p = .276$). The Wald statistics for the overall effect of age on the outcome variable are, however, not significant ($p = .129$) despite being significant for the age group of 31–40-year-olds ($p = .021$). The exp $b$ value indicates the changes in the odds ratios as a result of a comparison of different units, i.e. age groups for instance (Field, 2013, p. 786). This value shows an odds ratio of 3.85 for the age group 31–40-year-olds in comparison to the reference group of over-60-year-olds, which means that the odds of finding this particular usage feature unacceptable rise with the age of the participants. These findings are in line with the results of the Mann-Whitney $U$-test, which showed a significant correlation of age with acceptability judgment. However, the overall effect of age is no longer significant when taking into account all social variables.

Having analysed the sociolinguistic variation of attitudes towards differently than, I will now discuss the contextual preference of S1. Furthermore, possible correlations between the respondents’ acceptability judgments
and their degree of certainty as well as judgment basis will be investigated. While 39.3 per cent of all judgments made by the questionnaire respondents fell into the context ‘unacceptable’, the average acceptability rating of *differently than* in stimulus sentence S1 amounts to 32.6 per cent. What becomes clear from the contextual acceptability rating of *differently than* in Figure 7.1 is the formality scale of the three media of language discussed above: writing, online/mobile and speech (see § 5.3.1). The use of *differently than* in language used in CMC contexts falls neatly in between the two traditional contexts of written and spoken language, thus confirming Crystal’s (2006a, pp. 51–52) description of CMC as the new third medium. The contextual preferences of *differently than* are presented in the order of increased frequency of acceptability.

![Figure 7.1 Contextual acceptability in percentages: differently than](image)

The stimulus sentence S1. *The Americans look at this differently than the British* is considered acceptable in formal writing by 17.9 per cent of the informants only. The acceptability rate increases along the formality and media
scales and reaches its highest judgment in the informal speaking context, in which 53.6 per cent of all questionnaire respondents find the use of differently than acceptable.

Making use of a Mann-Whitney U-test, I was able to determine a statistically significant difference between respondents expressing negative and positive judgments with regard to the level of certainty about their acceptability judgment ($U = 1105, p = .009, r = -.25$). According to the statistical analysis, acceptability judgments determining the usage’s unacceptability were made with a higher level of certainty ($Mdn = “absolutely certain”) than usage judgments of acceptability ($Mdn = “somewhat certain”). To test whether the different respondents’ acceptability ratings correlate significantly with their basis of judgment, a Fisher’s exact test was used as well as Spearman’s correlation coefficient to indicate the direction of the correlation. These tests showed that respondents’ who consider differently than acceptable tend to base their judgments on a feeling rather than their knowledge of a rule ($r_s = .230, p = .014$), which also resulted in a significant difference according to the performed Fisher’s exact test ($p = .023$).

A qualitative analysis of the 45 comments made by the questionnaire respondents produced three main topics, i.e. concerning offering corrections, distinguishing personal usage from that of others, and highlighting contextual usage. The above-mentioned findings of contextual preferences and the sociolinguistic variation of usage attitudes with respect to this particular usage problem will be complemented by metalinguistic comments of the questionnaire respondents which will enable a better understanding of and greater insights into their usage attitudes. The overlapping character of some topics identified in the respondents’ comments makes a classification of comments not always straightforward. An overview of the identified themes can be found
listed for each sample sentence separately in Appendix J. For practical purposes, I decided to italicise the usage features in the comments of the questionnaire respondents in the examples below and throughout this study.

The topic of offering corrections brings to light an awareness among the questionnaire respondents of the stigmatised and disputed status of differently than in the stimulus sentence they were presented with in the online questionnaire. As can be seen from examples (1) – (4), the respondents identified differently than as the culprit in the stimulus sentence.

1. Things differ from one another, not to or than. (Retired dental surgeon, over 60 years old, male)
2. One differentiates ‘from’ and compares ‘to’. ‘than’ is just incorrect. (Retired, over 60 years old, male)
3. A thing is either different ‘TO’ or different ‘FROM’ another thing not different ‘THAN’. (Security consultant, over 60 years old, male)
4. This is interesting – even though the ‘rule’ is about different than/from/to, it’s the use of ‘differently’ that jars here. I think ‘in a different way from (or than)’ would sound more natural. (Editor, 31–40 years old, female)

These comments also indicate the respondents’ extension from differently than to the distinction between different from/than/to. Comments (1) and (2) exhibit prescriptive views in that only different from is considered the correct variant. What is furthermore interesting in comment (1) is that the respondent argues that the stimulus sentence is unacceptable because the verb differ is traditionally followed by from. A similar reasoning is provided in comment (2), which further extends the explanation to the verb compare being followed by to, indicating that different to is not a suitable option either. These comments contrast with comment (3), in which the respondent claims that both
*different to* and *from* are acceptable. Comment (4) is intriguing in that it illustrates how *differently than* is linked to *different from/than/to* in general and in that the respondent, a female editor, expresses her issue with the stimulus sentence by concentrating on the adverbial use of *differently* and stating that it “jars”.

The second main topic constitutes a frequently recurring pattern in the metalinguistic comments studied in this thesis: distinguishing between personal usage and that of others. Examples (5) – (8) are examples of this theme and will be discussed in more detail.

(5) It’s a common usage. But one I don’t like.  
   (Education adviser, over 60 years old, male)

(6) I know this is not how I would say it.  
   (Specialist tutor for adult dyslexic students, over 60 years old, female)

(7) When I say unacceptable I mean I wouldn’t use it.  
   (Manager in a museum, 41–60 years old, female)

(8) Not how I would say it as I would say “from” rather than the American way used here.  
   (Retired, over 60 years old, female)

What these comments have in common is not only the distinction between the respondents’ personal usage and that of others, but also a distancing from the usage represented in the stimulus sentence which they perceive as unacceptable. Comments (5) – (7) state that the stimulus sentence does not represent the respondents’ own usage, while the respondent in (8) distinguishes between her own usage of favouring the construction *different from*, and that of American English speakers’, who she thinks would make use of the stimulus sentence presented. What became apparent in my analysis of such metalinguistic comments was a distinction that was made between American
and British English, as indeed in example (8). This is the third and last topic I would like to discuss here on the basis of examples (9) – (11).

(9) The use of “than” is American and grates.  
(PhD student, 26–30 years old, female)

(10) In British English it should be “differently to” - the example is acceptably [sic] to users of American English; I would never correct an American for saying it!  
(retired Primary and EFL teacher, over 60 years old, female)

(11) I understand it should be ‘from’ the British but in an informal context I would not correct this mistake in my own or other’s language.  
(English teacher, 18–25 years old, female)

As can be seen from these examples, different than is associated with American English by all three informants. In comment (9) a female PhD student states that this particular usage “grates”. While comments (10) and (11) both provide corrections as in those that were discussed in examples (1) – (4), albeit in relation to differently to and differently from respectively, the retired teacher in (10) emphasises the acceptability of different than in American English. Interestingly, the English teacher in (11) provides further contextual information in that the descriptive use of different than in informal context would not trigger a correction with her, be it in her own usage or that of others. This last comment also emphasises the overlapping character of some of the comments.

Adding a qualitative dimension in the form of the respondents’ comments to the analysis of usage attitudes towards different from/than/to emphasises the variability of attitudes towards usage problems in that the context in which particular usage problems appear seems to play a role. Furthermore, offering corrections and being able to extend the occurrence of differently than to the usage problem different from/than/to illustrates how a part of the general
public possesses a high degree of awareness towards this particular usage feature. That differently than in the stimulus sentence is associated with American English seems to confirm this notion, which is frequently disseminated in usage guides and other reference works, such as the OED, as discussed above. Therefore, it does not seem surprising to find respondents commenting on differences between Americans and British speakers with regard to usage. What is, however, important is the distancing of respondents to the allegedly incorrect use of differently than, as was demonstrated in comments (5) – (8). The comments analysed serve as an indication of the social salience of differently than and different from/than/to as a usage problem. Having analysed the respondents’ usage attitude data on differently than, I will now turn to the treatment of this particular usage feature in the HUGE database.

In order to provide an overview of how differently than has been treated by authors whose usage guides are included in the HUGE database, it is necessary to bear in mind that this particular usage feature has two further variants: differently to and differently from. The focus of this analysis is, however, on differently than. Differently than is discussed in 30 British usage guides included in the HUGE database. To classify these, I will make use of a slightly modified version of Yáñez-Bouza’s (2015) tripartite categorisation of usage precepts into “criticised”, “advocated” and “neutral”. For this particular usage problem, I have also included a category “not mentioned” to indicate that the variant differently than was not discussed in a particular entry. All in all, differently than turns out to be the variant most often criticised since it appears in thirteen entries, followed closely by differently to, which is criticised ten times. The prescribed variant differently from was never criticised, which was to be expected as it is widely accepted and acknowledged as being part of Standard English. What needs to be borne in mind
is that usage guide authors tend to discuss the usage problem *different from/than/to*, which forms the basis of this analysis. The adverbial form *differently than* is, however, occasionally mentioned, as for example by Howard (1993; p. 124), who states that “[t]han is particularly useful after differently: ‘they do things differently in New York *than* in London’”. Before providing an overview of the treatment of *differently than* in the usage guides included in HUGE in Table 7.10 below, I will exemplify the categories used as follows.

Criticised

XCIX. DIFFERENT THAN. I found your affairs had been managed in a different manner than what I had advised. Ibid. A different manner than is not English. We say different to and different from; to the last of which Expressions I have in another Place given the Preference, as seeming to make the best Sense. (Baker, 1770, p. 100)

Neutral

Different than is an established idiom in American English, but is not uncommon in British use… Both different to and different than are especially valuable as a means of avoiding the repetition and the relative construction required after different from in sentences like *I was a very different man in 1935 from what I was in 1916* (Joyce Cary). This could be recast as *I was a very different man in 1935 than I was in 1916* or than in 1916…This construction is especially common when different is part of an adverbial clause (e.g. in a different way) or when the adverb differently is used, and has been employed by good writers since the seventeenth century … (Burchfield et al., 1984, pp. 101–102)

Advocated

Since the 18th century, different than has been singled out by critics as incorrect, but it is difficult to sustain the view in modern standard English that one version is more correct than the others. There is little difference in sense between the three, and all of them are used by respected writers. (Butterfield, 2007, p. 40)

Not mentioned

Sticklers would rather have ‘different from’ than ‘different to’ but some good writers (including Charlotte Bronte) have given ‘different to’ respectability. We shall not distinguish here. (Blamires, 1994, p. 20)
The first example listed here, which presents the earliest discussion of different than in a usage guide included in HUGE, namely Baker’s (1770, p. 100), labels different than as “not English”. Interestingly, Baker states that the other two variants could be found in English and he refers to an entry in his usage guide which, however, also describes the use of different to as an “impropriety” he would like to see banished from the English language (Baker, 1770, pp. 7–8). Burchfield et al. (1984, p. 101) on the other hand describe the difference in usage between American and British English and refrain from passing an explicit judgment on the use of different than. The stigmatisation of different than is also discussed in Butterfield (2007, p. 40), who however argues that there is “little difference between the three [variants]”. Since there are three possible variants for this construction, Blamires’ The Queen’s English (1994) is interesting as it does not mention the often stigmatised and criticised different than, but rather distinguishes between different from and different to, which makes this particular instance an example for the last category “not mentioned”. An overview of all usage guides discussing different from/than/to is given in Table 7.10. What needs to be mentioned here is that the tables discuss the treatment of the usage problems investigated include both the date of the first publication and the edition used in the HUGE database (in brackets).

As can be seen from this table below, the majority of British usage guides discussing this particular usage problem criticise the use of different than, while only three advocate the variant. That different than is not mentioned in usage guides discussing the usage issue of which preposition is to follow different seems to be a temporal phenomenon. This neglect of different than could indicate a shift with respect to the prominence of the other two variants. The historical treatment of different than in HUGE is shown in
Figure 7.2 below, which illustrates the above-mentioned chronological development of this phenomenon.

Table 7.10 reatment of different than (“criticised”, “neutral”, “advocated” and “not mentioned”) in British usage guides

<table>
<thead>
<tr>
<th>Attitude</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>not mentioned</td>
<td>Alford1864, Fowler&amp;Fowler106(1922), Fowler1926, Gowers1948, Vallins1953(1960), Blamires1994</td>
</tr>
<tr>
<td>Total</td>
<td>30</td>
</tr>
</tbody>
</table>

Figure 7.2 illustrates that the beginning of the stigmatisation history of different(ly) than can be traced back to the late-eighteenth century. Interestingly, different(ly) than only re-emerged as a usage problem in the 1960s and was mainly criticised until the 1980s. Only in the 1980s did usage guide authors start to approve of different(ly) than, which, however, does not mean that they unanimously agreed on this feature’s acceptability. Until the 1940s, different(ly) than was not mentioned in the usage guides as problematical, as the focus at the time was rather on the distinction between different from and different to. A notable exception is Baker’s two editions of his usage guide published in 1770 and 1779 respectively. These findings contribute to Busse and Schröder’s (2010b, p. 97) argument that in the twentieth century different from became the preferred variant in British English. That the focus shifted to stigmatising different than from the 1950s onwards could be due to the increasing importance and spread of American English after the Second
World War (Bauer, 1994, pp. 65–66). Figure 7.2 below shows how a prescriptive treatment of *differently than* seems to become more prominent in the 2000s.

![Figure 7.2 Diachronic treatment of *differently than* in British publications](image)

To summarize, with an average acceptability rating of 32.6 per cent, *differently than* showed a clear contextual preference in terms of acceptability judgments. While the stimulus sentence S1 (*The Americans look at this differently than the British*) indicates an acceptability rating of only 17.9 per cent in the formal writing context, it achieved its highest acceptability rating of 53.6 per cent in the informal speaking context. As can be seen from Figure 7.1 above, the three media of language – written, online, and spoken usage – fall neatly into an order of increased acceptability across the formality scale. The Mann-Whitney $U$-test brought to light how the social variable age showed an initially significant correlation with acceptability ratings, which would indicate an increase in intolerance towards the use of *differently than* with age. The overall effect of age on the acceptability judgment was, however, diminished in the binary logistic regression test which accounts for all possible
effects of the social variables investigated. Furthermore, questionnaire respondents who made prescriptive judgments show a greater degree of certainty and tend to base their judgments on the self-reported knowledge of a rule rather than a gut feeling.

My analysis of qualitative metalinguistic comments does not only reveal that some respondents possess a high degree of awareness of the usage problem at hand, but that they are also able to extend the adverbial *differently than* to the overarching issue of distinguishing between *different from*, *than* and *to*. Additionally, the comments show how respondents tend to distance themselves from a particular usage feature which they have perceived as unacceptable by differentiating between their own usage and that of others. As a result, the affiliation of *different(ly) than* with American English is enforced. Corpus evidence supports this finding, as *different(ly) than* occurs more frequently in COCA than in the BNC.

Analysing the entries in HUGE, a clear shift in the discussion and treatment of *different from/than/to* can be detected in the sense that until the 1940s *different than* was barely discussed in the advice literature. Since usage guides are a reaction to usage rather than an attempt to pre-empt usage, the spread of American English in the UK after the Second World War could pose a possible source for the stigmatisation of *different than* in British usage guides (cf. Bauer, 1994, pp. 65–66). Although *different than* has been found to be acceptable by some usage guide authors, such as Peters (2004), the feature remains highly disputed and stigmatised.

7.2.1.2. **Data are**

Having provided an insight into why the use of *data is* is considered problematical in Section 6.3 above, I will now turn to the analysis of the perception data. However, the findings of my corpus analysis are presented first. It does not come as a surprise to find the highest frequency of *data are* in the academic
subsection of the BNC given the stimulus sentence’s (S2. *The data are often inaccurate*) formality and academic style. *Data are* scores a normalised frequency rate of 20.35 tokens per million words in this particular context. *Data is*, on the other hand, records a lower normalised frequency rate of 13.11 tokens per million words in the academic subsection of the corpus, which is also the subsection with the highest frequency rating. Since the spread and influence of American English has been associated with the increased use of the disputed variant *data is* in British English (cf. Peters, 2004, p. 140), I will also draw on COCA to identify possible differences in usage frequencies between American and British English; this needs to be considered with care given the corpora’s characteristics described in Section 5.5 above. Hence the corpus findings presented throughout the study should be viewed as illustrations of usage tendencies in British and American English respectively. The focus of the corpus analysis does not lie on a diachronic comparison of usage frequencies, but on identifying general usage tendencies and interpreting them in relation to the corpora’s subsections. The results of my corpus analysis are presented in Table 7.11.

**Table 7.11 Overview of *data is* and *data are* in BNC and COCA**

|-----|--------|---------|----------|-----------|----------|-------|-------|
| *Data is*  
*(n=452)* | Freq. per mil | 11 | 2 | 17 | 8 | 98 | 201 | 115 |
| *Data are*  
*(n=491)* | Freq. per mil | 0 | 0 | 5 | 0 | 99 | 312 | 75 |

<table>
<thead>
<tr>
<th>COCA</th>
<th>Spoken</th>
<th>Fiction</th>
<th>Magazine</th>
<th>Newspaper</th>
<th>Acad.</th>
</tr>
</thead>
</table>
| *Data is*  
*(n=1,859)* | Freq. per mil | 191 | 58 | 419 | 281 | 910 |
| *Data are*  
*(n=2,889)* | Freq. per mil | 42 | 18 | 351 | 158 | 2,320 |
That both variants obtain the highest frequencies in the academic subsection of COCA confirms the findings of the BNC. Additionally, the corpus search in COCA for *data are* resulted in a normalised frequency of 22.43 tokens per million words in the academic register, while *data is* only resulted in 8.80 tokens per million words, as can be seen in the table above.

In order to obtain a more detailed insight into the diachronic development of *data are* and *data is* in American English and so to determine whether American English could potentially have influenced British English, as claimed by usage guide authors, I will make use of COHA, a corpus of historical American English which was described in detail in Section 5.5 above. The results of this search are presented in Table 7.12 below.

**Table 7.12 Overview of data is and data are in COHA by decade since 1810**

<table>
<thead>
<tr>
<th>Decade</th>
<th><em>Data is</em> (n=177)</th>
<th><em>Data are</em> (n=438)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Freq. per mil</td>
<td>Freq. per mil</td>
</tr>
<tr>
<td>1810</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td>1820</td>
<td>1 0.14</td>
<td>0 0</td>
</tr>
<tr>
<td>1830</td>
<td>0 0</td>
<td>1 0.07</td>
</tr>
<tr>
<td>1840</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td>1850</td>
<td>0 0</td>
<td>2 0.12</td>
</tr>
<tr>
<td>1860</td>
<td>1 0.06</td>
<td>9 0.53</td>
</tr>
<tr>
<td>1870</td>
<td>0 0</td>
<td>9 0.48</td>
</tr>
<tr>
<td>1880</td>
<td>0 0</td>
<td>1 0.05</td>
</tr>
<tr>
<td>1890</td>
<td>0 0</td>
<td>4 0.19</td>
</tr>
<tr>
<td>1900</td>
<td>1 0.05</td>
<td>14 0.63</td>
</tr>
<tr>
<td>1910</td>
<td>3 0.13</td>
<td>13 0.57</td>
</tr>
<tr>
<td>1920</td>
<td>2 0.08</td>
<td>22 0.86</td>
</tr>
<tr>
<td>1930</td>
<td>10 0.41</td>
<td>33 1.34</td>
</tr>
<tr>
<td>1940</td>
<td>8 0.33</td>
<td>62 2.55</td>
</tr>
<tr>
<td>1950</td>
<td>10 0.41</td>
<td>16 0.65</td>
</tr>
<tr>
<td>1960</td>
<td>11 0.46</td>
<td>42 1.75</td>
</tr>
<tr>
<td>1970</td>
<td>11 0.46</td>
<td>37 1.55</td>
</tr>
<tr>
<td>1980</td>
<td>26 1.03</td>
<td>53 2.09</td>
</tr>
<tr>
<td>1990</td>
<td>50 1.79</td>
<td>68 2.43</td>
</tr>
<tr>
<td>2000</td>
<td>43 1.45</td>
<td>52 1.76</td>
</tr>
</tbody>
</table>
A basic search of both variants in COHA shows that the variant *data is* has gained ground and has become more frequent in comparison to the accepted variant from the early mid-twentieth century onwards. The disputed variant resulted in 0.41 tokens per million words in the 1930s, while at the same time *data are* only obtained a standardised frequency rate of 1.34 tokens per million words. As can be seen from Table 7.12 above, *data is* has become increasingly more frequent over time, yet it has not exceeded the frequency of the prescribed variant *data are*. To provide a historical overview of the development of these two variants in British English, a corpus search of the Hansard Corpus was conducted. My findings for this corpus search are presented in the overview in Table 7.13.

Table 7.13 Overview of *data is* and *data are* in Hansard Corpus

<table>
<thead>
<tr>
<th>decade</th>
<th><em>data is</em> (n=429)</th>
<th><em>data are</em> (n=954)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Freq. per mil</td>
<td>Freq. per mil</td>
</tr>
<tr>
<td>1800</td>
<td>1 0.20</td>
<td>0 0</td>
</tr>
<tr>
<td>1810</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td>1820</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td>1830</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td>1840</td>
<td>0 0</td>
<td>0 0</td>
</tr>
<tr>
<td>1850</td>
<td>0 0</td>
<td>1 0.03</td>
</tr>
<tr>
<td>1860</td>
<td>0 0</td>
<td>4 0.12</td>
</tr>
<tr>
<td>1870</td>
<td>0 0</td>
<td>1 0.03</td>
</tr>
<tr>
<td>1880</td>
<td>0 0</td>
<td>2 0.03</td>
</tr>
<tr>
<td>1890</td>
<td>0 0</td>
<td>3 0.06</td>
</tr>
<tr>
<td>1900</td>
<td>2 0.03</td>
<td>18 0.28</td>
</tr>
<tr>
<td>1910</td>
<td>5 0.06</td>
<td>19 0.24</td>
</tr>
<tr>
<td>1920</td>
<td>7 0.10</td>
<td>24 0.33</td>
</tr>
<tr>
<td>1930</td>
<td>8 0.08</td>
<td>35 0.37</td>
</tr>
<tr>
<td>1940</td>
<td>12 0.13</td>
<td>11 0.12</td>
</tr>
<tr>
<td>1950</td>
<td>10 0.08</td>
<td>12 0.10</td>
</tr>
<tr>
<td>1960</td>
<td>23 0.15</td>
<td>25 0.16</td>
</tr>
<tr>
<td>1970</td>
<td>30 0.18</td>
<td>46 0.28</td>
</tr>
<tr>
<td>1980</td>
<td>110 0.60</td>
<td>271 1.47</td>
</tr>
<tr>
<td>1990</td>
<td>110 0.62</td>
<td>219 1.24</td>
</tr>
<tr>
<td>2000</td>
<td>111 1.25</td>
<td>263 2.97</td>
</tr>
</tbody>
</table>
This corpus shows that the use of *data are* has experienced a slump with regard to its normalised frequency in 1940. *Data are* shows a frequency rate of 0.12 tokens per million words, while *data is* ranks 0.13 tokens per million words. Before 1940, *data are* had predominantly been used and showed persistently higher frequency rates than *data is*, as can be seen from Table 7.13 below. Interestingly, the Hansard corpus bears evidence of a resurgence of the prescribed variant. This trend becomes strongly evident from the 1980s onwards when *data are* shows an increased normalised frequency of 1.47 tokens per million words compared to 0.60 for *data is*. For the last decade, the 2000s, my analysis of the Hansard Corpus indicates how *data are*, with a normalised frequency rate of 2.97 tokens per million words, is considerably more frequent than *data is*, which has a normalised frequency rate of 1.25. It needs to be borne in mind, however, that the normalised frequencies as well as raw figures in COHA and the Hansard Corpus are fairly low. Nonetheless, the corpus evidence illustrates general usage tendencies and shows how *data is* has been used more frequently in American English from the mid-twentieth century onwards. The findings of the Hansard Corpus indicate a shift occurring in the mid-twentieth century affecting the frequencies of usage of both variants, but it also highlights how *data are* has gained ground in parliamentary speeches from the 1980s onwards. Therefore, the corpus evidence demonstrates how *data are* is not only being used more frequently than *data is* in both BNC and COCA, but also that the feature is associated with formality, as the academic subsections showed the highest frequency rates in the corpora and a high degree of formality may be assumed to be characteristic of parliamentary speeches. The association of *data are* with formality is supported by the complete lack of occurrences of *data are* in the spoken, fiction and newspaper subsections of the BNC, as shown in Table 7.11.
As for the sociolinguistic analysis, the social variables age ($U = 1109, p = .076, r = -.17$) and gender ($\chi^2 (1) = 0.1, p = .946$) in relation to acceptability ratings showed no statistically significant correlations. Similar results were obtained in the analysis of acceptability ratings and the social variable education (university-educated vs non-university-educated) ($r_s = .114, p = .232$). Interestingly, nativeness showed a weak positive correlation with acceptability ratings ($r_s = .278, p = .003$), indicating that native speakers tend to express a greater acceptability towards the use of *data are* in the stimulus sentence investigated. Using a binary logistic regression analysis to determine the influence of the social variables investigated in this study on the relationship identified between the social variable nativeness and acceptability judgment, I was able to identify a prediction accuracy of 67 per cent, if none of the social variables are included in the model as predictors. The application of a forced data entry method resulted in the proposed model with an increase of the overall prediction accuracy to 71.4 per cent. A summary of this model is provided in Table 7.14 below.

Table 7.14 Results of binary logistic regression: *data are*

<table>
<thead>
<tr>
<th>Included</th>
<th>95% CI for $\exp b$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$B(\text{SE})$</td>
</tr>
<tr>
<td>Nativeness</td>
<td>$-1.85 (0.78)^*$</td>
</tr>
<tr>
<td>Gender</td>
<td>$-0.28 (0.45)$</td>
</tr>
<tr>
<td>Age (18–25)</td>
<td>$-1.30 (0.66)$</td>
</tr>
<tr>
<td>Age (26–30)</td>
<td>$-0.08 (0.73)^*$</td>
</tr>
<tr>
<td>Age (31–40)</td>
<td>$-0.45 (0.62)$</td>
</tr>
<tr>
<td>Age (41–60)</td>
<td>$-0.49 (0.69)$</td>
</tr>
<tr>
<td>Level of education</td>
<td>$-0.85 (0.63)$</td>
</tr>
<tr>
<td>Constant</td>
<td>1.59 (0.52)*</td>
</tr>
</tbody>
</table>

Note $R^2 = .10$ (Hosmer & Lemeshow), Model $\chi^2 (7) = 13.99, p = .051, p < .05^*$
Testing the fit of the model showed that the inclusion of all predictors produced a good fit model ($\chi^2(7) = 13.99$, $p = .051$). The Wald statistics confirmed that nativeness is the social variable whose inclusion in the model is most significant ($p = .017$). When comparing the age group of 26–30-year-olds to the reference group of over-60-year-olds, a significant correlation can be identified between these two variables ($p = .0498$). However, the Wald statistics for the overall effect of age indicate that this effect is cancelled out ($p = .367$). The $exp$ $b$ value indicates that being a non-native speaker increases the odds for obtaining a negative acceptability judgment, as the odds ratio is .16. These findings confirm the above-mentioned effect of nativeness with regard to the acceptability judgment on *data are*.

As mentioned above, the stimulus sentence (S2. *The data are inaccurate*) seems to be associated with formal contexts due to its formality and academic style. Figure 7.3 below shows the contextual acceptability distribution of *data are*. What is intriguing, however, is that this particular usage problem was considered unacceptable by 33 per cent of all respondents. As can be seen in the figure, the stimulus sentence is considered widely acceptable in all contexts ranging from 41.1 per cent in informal CMC to 63.4 per cent in formal writing, which resulted in an average acceptability rating of 48.5 per cent. It is interesting to see that this average acceptability rating is considerably lower than the one obtained by Mittins et al., whose data for *data is* obtained an average rating of 69 per cent.
Figure 7.3 Contextual acceptability in percentages: *data are*

Figure 7.3 furthermore shows that both CMC-related contexts, informal and formal online/mobile usage, are no longer situated between the spoken and written language media, but precede the speaking and written media for this usage problem.

One way of assessing the respondents’ attitudes towards *data are* is to identify any possible correlation between the general acceptability rating of the stimulus sentence and the respondents’ certainty and basis of judgment (see § 5.3.1). While there was no statistically significant difference in terms of degree of certainty about their judgments between respondents who found this stimulus acceptable and those who did not ($U = 1254$, $p = .329$, $r = -.09$), a statistically significant difference between these two groups and their basis of judgment was identified using a Spearman’s correlation coefficient and Fisher’s exact test ($p = .004$). The group of respondents who found *data are* acceptable tend to base their judgments more frequently on a rule than those who rated the stimulus sentence containing the prescribed variant as unacceptable ($r_s = -.276$, $p = .003$). This finding adds an interesting perspective to the
discussion of usage attitudes as it indicates how those who report that they
base their judgment on a rule consider the stimulus sentence acceptable, while
those who do not base their judgment on a gut feeling. As discussed in Chapter
2, norma loquendi or customary usage seems to have influenced the latter
group’s judgment.

To turn to the analysis of comments, the respondents provided 53 com-
ments altogether. Analysing these comments brought to light two main topics:
offering a correction or explanation and commenting on personal usage. What
needs to be mentioned here, however, is that the corrections provided affected
the prescribed variant, which was changed into the variant data is. Explan-
ations for these corrections were provided by those who seem to follow de-
scriptive customs, while those who are aware of data being traditionally con-
sidered a plural noun offer an explanation for their judgments. Comments (12)
– (15) serve as examples of the first and most frequently mentioned topic of
offering a correction of data are or providing an explanation for the accept-
ability of data is.

(12)  data should always be plural.
      Retired arts consultant, over 60 years old, female)

(13)  Data is the plural, datum is the singular.
      (Security consultant, over 60 years old, male)

(14)  The data is or was inaccurate.
      (Civil servant, 31–40 years old, female)

(15)  “The data” is singular, so it should read “The data IS ......”
      (Retired, over 60 years old, female)

While (12) and (13) seem to be elaborations of respondents who express tradi-
tional views on this particular usage feature, comments (14) and (15) correct
the prescriptive variant data are into the descriptive variant data is. Awareness
of the growing acceptability of data is and its changing use as a plural are
illustrated in comments (16) and (17), which emphasise the difference between prescribed norms and customary usage.

(16) It sounds horribly unnatural, but I do know that ‘data’ is both technically plural and singular. I feel, however, that it is becoming naturalised in English and soon it would not be uncommon to see ‘datas’ as a plural.
(Student, 18–25 years old, female)

(17) The word ‘data’ although technically plural is generally acceptable as a singular noun.
(Stay-at-home mother, 31–40 years old, female)

The distinction between norm and custom is a vital one, which often seems to be forgotten in the usage debate. Yet, comments such as (16) and (17) illustrate how prescribed, traditional norms can be perceived by members of the general public. All in all, five of the comments make use of the word “technically” to refer to the traditional prescriptive norm of *data* as a plural noun, which is, however, followed by respondents stating that its use as a singular noun is commonly accepted as well, as it is for example illustrated in (17). Comments (18) – (20) provide further insights into how the use of the prescribed variant is perceived.

(18) I feel that Latinate plural agreements are still expected in formal contexts, especially if failure to use them might mark one out as ignorant. However, using this in an informal context could make one seem rather pompous or pedantic.
(English teacher, 31–40 years old, male)

(19) “data” is plural, however, using it in that context informally makes you sound a bit stuck up.
(Writer/Journalist, 26–30 years old, male)

(20) Feels wrong but I’m pretty sure it’s grammatically correct.
(Student, 18–25 years old, female)
Not only does the respondent in (18) comment on the Latinate origin of the rule governing *data* as the plural form of *datum* as well, both respondents in (18) and (19) elaborate on the contextual use of *data are*. What is intriguing about those two comments is that they show an insight into the part of what consequences non-compliance with the norm or custom can cause. The respondent in (18) argues that the use of *data are* is “still expected in formal contexts” and that not complying with these expectations would be perceived as “ignorant”. Yet, using the prescribed variant in informal contexts would be considered as “pompous or pedantic” (18) and “a bit stuck up” (19), as some respondents commented, which shows that usage has strong social connotations. Lastly, the comment in (20) serves as an example of how norm awareness and customary usage can affect speakers, such as this female student who, despite being certain about the grammatical correctness of the sentence, feels that something is “wrong” with the stimulus sentence.

The final component of the analysis constitutes the HUGE database analysis. 28 British publications discuss Latinate plurals, three of which do not mention *data are* explicitly, but rather discuss the issue at hand by using other Latinate plural nouns used in a singular manner such as *agenda* and *graffiti*. Before providing an overview of the categorisation of the usage entries into “criticised”, “neutral” and “advocated”, examples of each category are given below.

### Criticised

DATA is plural only (*The d. are, not is, insufficient.*/*What are the d.?*/*We have no d.*); the singular, comparatively rare, is *datum*; one of the *data* is commoner than a *datum*; but *datum-line*, line taken as a basis, is common.

(Fowler, 1926, p. 108)

### Neutral

This is a Latin plural and is generally used with a plural verb in English:

The data available are inadequate.
However, there is a growing tendency to consider *data* as a collective noun grouping together individual objects and to attach a singular verb to it:

The data he has accumulated is sufficient for our purposes. (Bailie & Kitchin, 1988, p. 95)

**Advocated**

Originally *data* was a plural noun: *These data are all wrong*. But it’s now widely used as a collective singular noun: *Let me know when all this data has been entered in the computer*. Both usages are acceptable in standard English. The singular is becoming more common than the plural, and it’s the standard usage in the field of computers. (Ayto, 1995, p. 85)

As can be seen from these examples, Fowler’s advice is bluntly straightforward, despite his acknowledgement of the prescriptive singular *datum* being rather rare. Bailie and Kitchin’s *The Essential Guide to English Usage* (1988), on the other hand, contains a description of the use of *data* as a singular. However, the authors do not give advice on whether to use *data are* or *data is*. The example for the advocated category provides an insight into how the use of *data is* has been promoted in the past few decades by connecting this development to the field of computer sciences. Ayto (2002, p. 85) states that both usages can be considered “acceptable in standard English”.

Table 7.15 contains an overview of all usage guides and their tripartite categorisation into “criticised”, “neutral” and “advocated”. For the sake of completeness, the three usage guides which do not discuss the issue of *data are* explicitly but rather focus on other plurals are included in the category “not mentioned”.

As can be seen from Table 7.15 the majority of the usage guides that deal with the usage problem *data is/are* take a neutral stance on the issue by not passing an explicit judgment or stating a contextual preference for either of the two usages. While nine usage guides criticise the use of *data is*, the usage is advocated in six of the 28 usage guides. What needs to be borne in
mind, however, is that advocating *data is* does not necessarily mean condemning *data are*. As shown in the example representing an advocated usage advice above, both variants are considered acceptable.

Table 7.15 Treatment of *data are* (“criticised”, “neutral”, “advocated” and “not mentioned”) in British publications

<table>
<thead>
<tr>
<th>Treatment</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>not mentioned (3)</td>
<td>Vallins1953(1960), Blamires1994, Lamb2010</td>
</tr>
<tr>
<td><strong>Total:</strong> 28</td>
<td></td>
</tr>
</tbody>
</table>

Criticism on *data is* in the usage guide tradition can be traced back to Fowler’s *A Dictionary of Modern English Usage* (1926). Figure 7.4 below illustrates how this usage problem is truly an issue of the twentieth century. *Data is* was widely condemned until the 1970s. The only usage guide published in this particular decade is Wood’s *Current English Usage* (1962, p. 68), which mentions the use of singular *data* as “a collective denoting a single body of facts”. Wood argues that *data is* is acceptable in certain contexts, though he leaves them unspecified. Technological advancements and the use of *data* as a singular in computing seems to have influenced usage guide authors’ precepts towards *data is* (Ayto, 1995, p. 85). Figure 7.4 shows how usage entries discussing *data is* in a neutral or advocating manner have become more frequent from the 1980s onwards.
An intriguing trend in the treatment of *data is* can be detected in the 2000s, when we see an increase in criticism of the descriptive use of *data* as a singular noun. Among the three usage guides criticising the use of *data is* are Sayce’s *What Not to Write* (2006) and Heffer’s *Strictly English* (2010). Both authors use *ipse-dixit* pronouncements in their advice, as exemplified by Sayce (2006, p. 44), who briefly states the following: “The word ‘data’ is plural. The singular form is ‘datum’. The data are reliable (*not* the data is reliable)”. This recent increase in criticism towards the use of *data is* is in line with Tieken-Boon van Ostade arguing that the Age of Prescriptivism is now (cf. Chapter 2). Both Sayce (2006) and Heffer (2010) neglect contextual differences of *data* as a singular and its overwhelming use in computer sciences.

Although the Latinate origin of the distinction between singular *datum* and plural *data* constitutes the original usage conundrum (see § 6.3) and is often wrongly resorted to by prescriptivists, it seems as if *data* has extended its original meaning, in being taken as the plural of *datum*, to reflect a single collection of various facts. The use of the term *data* in technology seems to have promoted its widespread use as singular. The corpus evidence shows

![Figure 7.4 Diachronic treatment of *data is* in British publications](image-url)
only minor differences in frequency between *data is* and *data are* in both British and American corpora; *data are* occurs more frequently than *data is* in the academic subsections of the corpora, which is in line with the formality of this subsection. However, interesting tendencies can be identified in the Hansard Corpus of parliamentary speeches. While *data is* was seemingly gaining ground before the 1980s, a shift can be identified in this decade, which saw the resurgence of *data are*.

My sociolinguistic analysis of the respondents’ acceptability judgments revealed a main effect for the variation between acceptability and nativeness, while no other social variable showed significant correlations. The importance of the social variable nativeness was confirmed in the binary logistic regression analysis. Despite the fact that they did not show a difference in the degree of certainty, the questionnaire respondents showed a clear difference in their judgment basis. Those respondents who found *data are* acceptable stated basing their judgment on the knowledge of a rule, while those who found it unacceptable reported basing their judgments on a gut feeling. These findings strengthen the distinction between norms and customary usage. With an average acceptability rating of 48.5 per cent, *data are* shows a fairly high acceptability.

The HUGE analysis of usage entries identified the origin of this usage problem’s stigmatisation in Fowler’s *A Dictionary of Modern English Usage* (1926). Before the advent of the internet in the 1990s, *data is* was first treated in a neutral manner in the 1970s. In subsequent years, the use of *data is* was increasingly advocated, as was shown in Figure 7.4. Yet again, the 2000s show a resurgence of proscriptions against *data is*.

### 7.2.1.3. Flat adverb: *go slow*

After having provided a description of flat adverbs in the preceding chapter (§ 6.4), I will now undertake a detailed analysis of my usage attitude data by
first providing a summary of a recent corpus study conducted by Lukač and 
Tieken-Boon van Ostade (forthc.), before adding the perspectives of laypeople 
and usage guide authors on the use of the flat adverb in the stimulus sentence 
(S3. *That’s a dangerous curve; you’d better go slow*). In a comparative corpus 
analysis conducted by Lukač and Tieken-Boon van Ostade (forthc.) it was 
shown that the flat adverb *go slow* occurred more frequently in COCA than in 
the BNC, in which both variants, *go slow* and *go slowly*, show almost equally 
high frequency rates. This study showed that *go slow* occurred most frequently 
in COCA in the subsections “fiction” and “magazines”, while in the BNC the 
spoken and fiction subsections show the highest frequencies (Lukač & 
Tieken-Boon van Ostade, forthc.). The findings of the corpus analysis con-
ducted by Lukač and Tieken-Boon van Ostade (forthc.) raise the question of 
whether *go slow* as a representative of flat adverbs or flat adverbs in general 
should be considered Americanisms. I will go into this question below.

My statistical analysis of a possible correlation between the social vari-
ables age, gender, nativeness and education level, and the obtained accept-
ability ratings produced the following results. While age ($U = 868$, $p = .357,$ 
r = –.09), education level ($r_s = .070$, $p = .466$) and nativeness ($r_s = –.012,$ 
p = .899) did not show any significant difference between those informants 
who rated the stimulus sentence as acceptable and those who did not, gender 
showed a significant difference ($\chi^2 (1) = 5.233$, $p = .022$). Women turned out 
to be three times more likely to deem the stimulus sentence unacceptable than 
men ($OR = 3.81$). This result confirms Lukač and Tieken-Boon van Ostade’s 
findings (forthc.) for the social variable gender. That women tend to reject the 
use of *go slow* could be a case of overt prestige, since women have been found 
to prefer standard variants (Trudgill, 1974, p. 94). Previous sociolinguistic 
studies proved that women tend to favour the standard variant or the variant
that carries more prestige (Trudgill, 1974, p. 94). The binary logistic regression analysis showed that a model which only includes the constant, i.e. excluding all social variables, predicts 80.4 per cent of all variation correctly. When all predictors are included in the model by means of a forced data entry method, the prediction accuracy of this model does not increase but remains the same. In Table 7.16 below, an overview of the proposed model is provided.

The model presented below proved to be a good fit ($\chi^2(7) = 8.82$, $p = .266$). The Wald statistics, however, confirmed that gender does indeed contribute meaningfully to the model ($p = .030$). The exp $b$ value indicates that if the gender category is increased (1 = “male”, 2 = “female”), the odds of obtaining a negative acceptability judgment increase with an odds ratio of 3.24. This confirms that women are three times more likely to reject the use of *go slow*, while a possible influence of other social variables included in the analysis can be ruled out.

<table>
<thead>
<tr>
<th>Included</th>
<th>$B(\text{SE})$</th>
<th>Lower</th>
<th>exp $b$</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nativeness</td>
<td>-0.03 (0.91)</td>
<td>0.16</td>
<td>0.97</td>
<td>5.78</td>
</tr>
<tr>
<td>Gender (18–25)</td>
<td>1.17 (0.54)*</td>
<td>1.12</td>
<td>3.24</td>
<td>9.36</td>
</tr>
<tr>
<td>Age (26–30)</td>
<td>0.95 (0.87)</td>
<td>0.47</td>
<td>2.59</td>
<td>14.35</td>
</tr>
<tr>
<td>Age (31–40)</td>
<td>1.13 (0.78)</td>
<td>0.67</td>
<td>3.08</td>
<td>14.29</td>
</tr>
<tr>
<td>Age (41–60)</td>
<td>0.20 (0.73)</td>
<td>0.29</td>
<td>1.22</td>
<td>5.13</td>
</tr>
<tr>
<td>Level of education</td>
<td>-0.17 (0.70)</td>
<td>0.21</td>
<td>0.85</td>
<td>3.34</td>
</tr>
<tr>
<td>Constant</td>
<td>0.56 (0.50)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note $R^2 = .08$ (Hosmer & Lemeshow), Model $\chi^2(7) = 8.82$, $p = .266$, $p < .05^*$

As for the analysis of contextual preferences, the flat adverb *go slow* shows a clear divide between formal and informal contexts in which the stimulus sentence (S3. *That’s a dangerous curve; you’d better go slow*) has been found acceptable by my own informants. While only 19.6 per cent of all
responses fell into the unacceptable category, Figure 7.5 below shows how the stimulus sentence was rated by the majority of respondents in all three informal contexts. Given the style and register of the stimulus sentence, it is no surprise to find the highest acceptability rating in the informal speaking context. With an average acceptability rating of 43.9 per cent, an unexpected decrease in acceptability can be noticed when comparing the stimulus sentence to Mittins et al.’s study which showed an acceptability rating of 54 per cent (Mittins et al., 1970, p. 109). Nevertheless, a comparison with the Mittins study should only be understood as presenting tendencies of possible changes having affected usage attitudes, as attitudes are notoriously difficult to compare given the methodological differences. Whether the lower average acceptability rating is due to such methodological differences, i.e. highlighting the usage problem and drawing on a different sample of informants, or to an increase in prescriptive attitudes over the years remains to be answered.

![Figure 7.5 Contextual acceptability in percentages: go slow](image)
Performing a $U$-test showed that there was a significant difference with regard to the different judgment groups and their respective reported degree of certainty ($U = 649, p = .003, r = -.19$). Yet, a more detailed analysis shows that the median in both groups constitutes “absolutely certain”. According to the analysis of the judgment basis, a positive correlation between acceptability and basing the judgment on intuition, i.e. a gut feeling, could be determined ($r_s = .328, p = .000$), with Fisher’s exact test showing a significance level of $p = .001$.

Investigating the metalinguistic comments obtained ($n = 42$) could provide an insight into why the questionnaire respondents expressed a lower acceptability overall when compared to Mittins et al.’s study. Many of the comments dealt with correcting the stimulus sentence, which indicates an awareness of the stigmatised feature. Examples of corrections can be found in comments (21) – (23) below.

(21) ‘Slow’ is an adverb & must take the -$ly$ ending.
(Retired dental surgeon, over 60 years old, male)

(22) ‘slow’ is an adjective, so the adverb ‘slowly’ should be used.
(Retired school teacher, over 60 years old, male)

(23) go slowly. Adverb, not adjective please!
(Retired arts consultant, over 60 years old, female)

All three respondents whose comments are represented here make use of appropriate terminology to refer to the issue, and furthermore provide a correction stating that instead of $slow$, the adverb $slowly$ should be used. Comment (21) seems to take a particularly firm stance on the issue by stating that the adverb “must take the -$ly$ ending”.

A number of respondents also mentioned other issues with this particular stimulus sentence, such as the use of a semi-colon or contractions. Comments (24) – (26) illustrate such issues and will be discussed in more detail.
While the respondent in (24) argues that the semi-colon used in the stimulus sentence would be perceived as arrogant in informal contexts, the respondent in (25), who described his occupation as that of an old nuisance, criticises the conceptual representation of slowing down which, according to him, would be better phrased using the comparative *slower*. The last comment discussed here, (26), highlights yet another issue which needs to be borne in mind. Contractions as in *that’s* and *you’d* are often perceived as informal features. As mentioned above, Mittins and his colleagues restricted this particular stimulus sentence in the choice of context with only informal contexts being available to their questionnaire respondents. This example illustrates the importance of selecting stimuli carefully. In order to make a tentative comparison with Mittins et al.’s study possible, I chose to make use of the same stimulus sentence.

A frequently established pattern identified in metalinguistic comments is the distinction between personal usage and the usage of others, which I would also like to illustrate briefly here. Comments (27) and (28) hint at the type of people who use flat adverbs such as *go slow*.

(27) *This is very American, I would say ‘slowly’ but this phrasing wouldn’t make me cringe.*

(Account manager for a charity, 26–30 years old, female)
(28) It should have a colon. In addition, the last word should be “slow-ly”; I despise the use of adjectives as adverbs.

(PhD student, 26–30 years old, female)

Comment (27) includes an association of flat adverbs with American English, which has already been mentioned above. Stating that this stimulus sentence “is very American”, the respondent, who is British, distances herself from this usage further by arguing that she would use the adverb *slowly*. In contrast to her rather lenient attitude towards flat adverbs, the PhD student in (28) corrects not only the semi-colon, but goes on to state that she despises flat adverbs.

Metalinguistic comments such as the ones discussed above provide a more detailed insight into usage attitudes and enable a better understanding of them. By analysing these comments, it was possible to see that the stimulus sentence used could potentially have influenced the respondents in their judgments, as the sentence contained features characteristic of informal language, such as contractions. Furthermore, the use of written stimuli to elicit attitudes towards spoken contexts needs to be mentioned here as problematical in a survey like this. These issues were, unfortunately, not raised in the pilot phase of the questionnaire, or I could have adapted the sentence accordingly (see § 5.3.1).

Having analysed the questionnaire respondents’ attitudes towards the flat adverb *go slow*, I will now present the findings of the analysis of usage guide entries discussing this particular usage feature. Out of the 39 British usage guides included in HUGE, 25 discuss flat adverbs. Since the database uses the term “slow/slowly” as a label for flat adverbs in general, the entries investigated also contain other flat adverbs. Categorising the entries of these usage guides on the basis of their treatment, i.e. whether the use of flat adverbs is explicitly criticised or advocated, or whether no explicit judgment is made but rather a contextual preference is stated, resulted in ten usage guides falling
into the “criticised” category, thirteen into the “neutral” and two into the “advocated” category. Examples of each category are presented below, followed by an overview of all 25 usage guides and their categorisation in Table 7.17.

**Criticised**

As explained on page 29, adverbs are words used to describe verbs, adjectives or other adverbs and are often formed by adding -ly to the adjective:

- The teacher was cross
- She spoke to me crossly (i.e. in a cross way)
- He was a heavy man
- He moved heavily (i.e. in a heavy way)

In recent years, however, many sports commentators have chosen to ignore this distinction and say such things as Federer is serving beautiful or Woods drove his tee shot perfect. In fact, this usage has become so common that it may almost be considered the norm. But only if one is a sports commentator. For anyone else, it is ungrammatical and unacceptable. (Taggart, 2010, p. 52)

**Neutral**

Slow or slowly As with quick, slow often replaces the correct grammatical form slowly. Markings on roads read SLOW, rather than SLOWLY, and workers decide to go-slow as a form of industrial action. In commands or very short sentences, especially following the verb ‘go’, slow is often the usual form: ‘Be careful and go slow’. In most other cases, especially in writing or in longer sentences, slowly is the correct form to use: ‘she drove slowly through the village’; ‘let’s go slowly until we see how things work out’. (Howard, 1993, p. 371)

**Advocated**

Go slow is an accepted idiom. The normal adverb is slowly, and the comparative more slowly, but we say ‘The car went slower and slower until it came to a standstill’. Perhaps we feel the word to be semi-adjectival, descriptive of the speed; and in any case more and more slowly would sound awkward. (Wood, 1962, p. 217)

The “criticised” example provided above describes how flat adverbs have increasingly spread in society, yet the author of the usage guide in question, Taggart (2010, p. 52), identifies sports commentators as being particular prone to the use of flat adverbs, which she calls “ungrammatical and unacceptable”. Howard’s (1993, p. 371) advice, cited as an example of the neutral treatment
of flat adverbs, distinguishes between different contexts and favours the use of \textit{slowly} in written contexts. Wood’s description of the flat adverb \textit{go slow} as “accepted idiom” (1962, p. 217), on the other hand, constitutes the most lenient treatment of flat adverbs in the usage guides investigated. Table 7.17 includes the detailed categorisation of usage guide entries.

Table 7.17 shows that flat adverbs are rarely advocated in the British usage guides included in HUGE. In order to analyse the diachronic development of the treatment of flat adverbs in these usage guides, I will provide an overview in Figure 7.6 below.

Table 7.17 Treatment of flat adverbs (“criticised”, “neutral” and “advocated”) in British publications

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>advocated</strong></td>
<td>Wood1962(1970)</td>
</tr>
</tbody>
</table>

This overview shows that flat adverbs were already criticised from the earliest days of the usage guide tradition onwards, as the feature already occurs in both editions of Robert Baker’s usage guide, published in 1770 and 1779 respectively. The difference between these two editions is the addition of a prescriptive comment on \textit{tolerable well/good}, which Baker advocated instead of \textit{tolerably} in his second edition (1779, p. 34). What followed was a period of criticism and neutral usage advice which, however, ended in a short period of acceptance in the 1960s and 1970s. This diachronic overview illustrates how flat
adverbs have consistently featured in usage guides from the 1960s onwards. Figure 7.6 below, however, clearly highlights how from the 1980s onwards prescriptive and proscriptive usage advice has been gaining ground.

To conclude, flat adverbs, such as the iconic *go slow*, have not only been considered old chestnuts in the usage debate, but they have also been frequently associated with American English. Based on an earlier study conducted by Lukač and Tieken-Boon van Ostade (forthc.), the corpus evidence showed that *go slow* does indeed occur more frequently in COCA than in the BNC. Although Lukač and Tieken-Boon van Ostade identified both gender and age as showing significant correlations in their attitude study with acceptability judgments, my own analysis, which exclusively concerned British users, revealed only an influence of gender on usage attitudes. The women in my sample are three times more likely to reject flat adverbs than men, which could hint at overt prestige. This finding was also confirmed in the binary logistic regression analysis. The reason for the difference between the two studies very likely lies in the different survey samples and methods applied.

![Figure 7.6 Diachronic treatment of flat adverbs in British English](image-url)
The analysis of contextual preferences of the questionnaire respondents for the stimulus sentence (S3. *That’s a dangerous curve; you’d better go slow*) showed the highest acceptability rating of nearly 75 per cent in the informal speaking context, while the lowest acceptability rating of 18.8 per cent was obtained for the formal written context. Hence, the average acceptability rating of the stimulus sentence containing *go slow* was 43.9 per cent, which was surprisingly lower than the average acceptability rating of 54 per cent identified by Mittins and his colleagues in the late 1960s (Mittins et al., 1970, p. 109). While the qualitative comments further provided an insight into how the use of flat adverbs is perceived, contextual differences were also commented on by the respondents. What needs to be emphasised here is the possible bias of other features of the stimulus sentence used in my investigation, such as the appearance of the semi-colon and the two contractions. These issues could have influenced the respondents’ judgments as a result of which the judgment could not only have affected the occurrence of the flat adverb but also that of the semi-colon or the contractions. The HUGE database analysis showed that flat adverbs, such as the iconic *go slow*, has a long history in the advice literature. Furthermore, usage guide authors tend to predominantly consider flat adverbs critically, as usage of the feature is rarely advocated. Again, the last three decades covered by the database show an increased occurrence of critical treatments of flat adverbs such as *go slow* in HUGE.

### 7.2.1.4. Like as an approximative adverb

Of the two nonstandard functions of *like* investigated in this study (see § 6.5), attitudes towards the use of *like* as an approximative adverb were elicited by means of the stimulus sentence (S4 *The new restaurant is like 2 minutes up the road*). Firstly, corpus evidence of the use of *like* as an approximative adverb will be provided in this section before a sociolinguistic analysis of the
questionnaire respondents’ attitudes is tackled. Conducting a corpus search in the BNC, I limited my search to patterns reflecting the stimulus sentence. Hence, making use of the POS-tagger in BNC, I searched for constructions involving *is/was/are/were like* being followed by a numerical expression. The same search was conducted for *about*. The corpus search showed that *about* \((n = 1,764)\) is used considerably more frequently than *like* \((n = 29)\) as an approximative adverb. The findings of the corpus search were limited to both adverbs fulfilling the approximative function, while *like’s* comparative functions, its use as a discourse particle, as well as *about* in the phrase *to be about* were carefully excluded. Examples of these types are given below in 5.a) and 5.b).

5. a) As a matter of fact we were like two atom bombs -- we’d go off together and there would be this tremendous explosion but we’d come down together, too. (CH8, written (biography))

    b) The superb women’s road movie Thelma And Louise is about two pals who cock a snoot at men in general during an impromptu crime odyssey. (CBC, written (newspaper other social))

Looking at *like* as an approximative adverb from a qualitative perspective, I was able to identify a similar notion as expressed by D’Arcy (2006, p. 342), who states that *like* co-occurs with words referring to age only to express vagueness. In such contexts, *like* behaves syntactically in a similar manner as the traditional approximative adverbs *about* and *roughly* (see § 6.5). Below, two examples of *like* as an approximative adverb are given, one expressing vagueness in connection with age (6.a) while the other one expresses vagueness concerning quantity (6.b).

6. a) No, I was just getting charged full fares when I was like thirteen.(KPF, spoken (conversation))

    b) Oh I think that was one of two. The other one was like four hundred and ninety nine pounds. (KPU, spoken (conversation))
In order to add a contrastive perspective, I looked at the corpus findings of approximative adverb *like* in more detail, which made it possible to identify the majority of occurrences in spoken contexts, namely 27 out of 29 approximative adverbs. By contrast, the approximative adverb *about* shows a greater occurrence in all subsections of the BNC. Since *like* as an approximative adverb seems only recently to have been gaining ground and replacing *about*, a corpus search of the BNC can only be taken as an indication of trends due to the composition of the corpus (see § 5.5). In order to provide a clearer overview of the development of *like* as an approximative adverb, a study making use of more up-to-date corpora such as D’Arcy’s study (2006) would be needed.

As was discussed in Section 6.5, D’Arcy (2006) showed the importance of age and gender in the variability in the use of approximative adverb *like* in Canadian English. Whether these social variables also play a role in the perceptions of this particular usage feature in British English will be examined here. My statistical analysis reflected the tendencies of lexical replacement of *about* by *like* in that the age of my informants proved to play a crucial role in the acceptability of the stimulus sentence ($U = 1046$, $p = .000$, $r = –.19$). This means that younger respondents tend to find the stimulus sentence acceptable ($Mdn = 31–40$-year-olds), while those who find it unacceptable tend to be older ($Mdn = 41–60$-years old). The other social variables, such as education level ($r_s = .091$, $p = .342$), gender ($\chi^2(1) = .001$, $p = .975$) and nativeness ($r_s = –.037$, $p = .695$), did not show any statistically significant differences. The results of the binary logistic regression analysis showed that the prediction accuracy of a model excluding all social variables is 58 per cent, which increases to 67 per cent if all predictors are included in the model illustrated in Table 7.18.
Testing the fit of the model showed that the model presented in Table 7.18 below is a poor fit ($\chi^2(7) = 20.32, p = .005$), which could be due to the quality of the data collected. Nonetheless, the Wald statistics confirmed the meaningful contribution of the social variable age to the model ($p = .002$), with the groups of 18–25-year-olds ($p = .016$), 26–30-year-olds ($p = .003$), and 41–60-year-olds ($p = .002$) displaying levels of statistical significance when compared to the reference group of over-60-year-olds. The $\exp b$ values of these three groups indicate that older respondents exhibit higher tendencies to reject the use of *like* in S4. This is also illustrated by their odds ratios of 4.08, 8.87 and 10.29 respectively. This confirms the significance of the age factor which has already been identified in the Mann-Whitney $U$-tests. These findings are in line with Chambers’ (2000) conclusion of lexical change to be found with specific age groups.

Table 7.18 Results of binary logistic regression: *like*

<table>
<thead>
<tr>
<th>Included</th>
<th>$B(SE)$</th>
<th>Lower</th>
<th>$\exp b$</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nativeness</td>
<td>0.44 (0.73)</td>
<td>0.37</td>
<td>1.55</td>
<td>6.52</td>
</tr>
<tr>
<td>Gender</td>
<td>0.04 (0.43)</td>
<td>0.44</td>
<td>1.04</td>
<td>2.44</td>
</tr>
<tr>
<td>Age (18–25)</td>
<td>1.57 (0.65)*</td>
<td>1.34</td>
<td>4.08</td>
<td>17.21</td>
</tr>
<tr>
<td>Age (26–30)</td>
<td>2.18 (0.74)*</td>
<td>2.08</td>
<td>8.87</td>
<td>37.83</td>
</tr>
<tr>
<td>Age (31–40)</td>
<td>0.71 (0.56)</td>
<td>0.68</td>
<td>2.03</td>
<td>6.05</td>
</tr>
<tr>
<td>Level of education</td>
<td>2.33 (0.74)*</td>
<td>2.40</td>
<td>10.29</td>
<td>44.06</td>
</tr>
<tr>
<td>Constant</td>
<td>0.56 (0.50)</td>
<td>0.26</td>
<td>0.86</td>
<td>2.92</td>
</tr>
</tbody>
</table>

Note $R^2 = .13$ (Hosmer & Lemeshow), Model $\chi^2(7) = 20.32, p = .005^*, p < .05^*$

The analysis of contextual preference of the stimulus sentence produced the following results, presented in Figure 7.7 below. The stimulus sentence showed its highest acceptability rate in the informal speech context with 55.4 per cent. Interestingly, the percentage of unacceptability ratings, namely 42 per cent, seems to suggest an almost equal divide between respondents with
respect to acceptability judgments. The stimulus sentence obtained an average acceptability rate of only 17.9 per cent. Figure 7.7 also shows how the stimulus sentence is considered to be informal rather than formal, since the formal contexts show the lowest acceptability ratings.

The analysis of the degree of certainty showed that there was a statistically significant difference between those informants who found the stimulus sentence acceptable and those who did not ($U = 1192, p = .006, r = -.19$), yet the median for both groups is the same ($Mdn$ = “absolutely certain”). Additionally, a significant positive correlation was identified when I analysed the respondents’ stated judgment basis, which showed that those who found the stimulus sentence acceptable tend to base their judgments on a gut feeling, while those who find it unacceptable report basing their judgments on a rule ($r_s = .244, p = .009$, Fisher’s exact test $p = .012$). All this hints at a pattern describing a significant correlation between judgments indicating unacceptability and basing such judgments on a self-reported knowledge of a rule.

Figure 7.7 Contextual acceptability in percentages: *like*
As for the qualitative analysis of comments made by the questionnaire respondents, 49 respondents commented on the stimulus sentence that contained like. As the analysis of contextual preference showed, a relatively large proportion of usage judgments fell into the unacceptable category, namely 42 per cent. That is why it does not come as a surprise that the majority of the respondents distinguish between their own personal use and that of others. In particular respondents foreground the use of like by young speakers who are accused of frequently using and even over-using this particular feature. Comments (29) – (33) exemplify this theme.

(29) I’m old, so would never use it.
   (Education adviser, over 60 years old, male)

(30) There is a generational issue here. I would never use ‘like’ in this way. But young people do.
   (Retired educational publisher, over 60 years old, male)

(31) If you use like in any other sense than to show appreciation or to compare things being alike its showing that you’re like an idiot - only half literate popstarlets [sic] talk like that!
   (Petrophysicist in oil company, 31–40 years old, female)

(32) Completely breaks the rules but I do it in informal contexts, so it must be acceptable, right?
   (Digital marketing consultant, 31–40 years old, female)

(33) I am guilty of this!
   (Manager in a museum, 41–60 years old, female)

The association of this usage feature with young speakers was frequently mentioned in the comments obtained from the questionnaire respondents, and it transpires in the examples provided above. Both respondents in (29) and (30), who believe this feature to be found in the speech of young speakers, argue that their generation would not use like in this manner. Comments (31) and (32) were made by 31–40-year-old female professionals whose views are
intriguingly divergent. The respondent in (31) states that using this particular feature would mark the speaker as “an idiot” and goes on to describe the stimulus sentence as being representative of “half literate popstarlets”. The respondent in (32) on the other hand admits breaking “the rules” in informal contexts, yet she also expresses a degree of uncertainty by questioning the acceptability of this feature. In stark contrast to (31), the respondent in (33) simply claims to be “guilty” of this type of usage herself. Analysing such comments enables a better understanding of the generational difference in usage which the comments seem to express.

What becomes obvious from the comments obtained is the different functions of *like* as perceived by the questionnaire respondents. While some consider *like* to be a filler, such as the respondents in comments (34) and (35), others find the use of *like* in the place of standard approximative adverbs such as *about* or *approximately* problematical. Comments (36) – (38) below mention this particular issue.

(34)  Like ‘like’ is another way of saying ‘errr’. Also don’t like ‘like’ the use of numerals with text.
(Security consultant, over 60 years old, male)

(35)  “Like” is completely unnecessary - allows speaker time to think
(Retired primary and EFL teacher, over 60 years old, female)

(36)  This is sloppy yob speech. What use does the word *like* have here? *About* is the correct word.
(Retired, over 60 years old, male)

(37)  I haven’t yet adopted ‘like’ to mean ‘approximately.
(Retired, over 60 years old, female)

(38)  Did it ever harm anyone to use ‘around’ and ‘approximately’?
(Student, 18–25 years old, male)
The difference in perception of the various functions of *like* is intriguing. While the use of *like* as a filler has frequently been studied, as discussed above (see § 6.5), the function of *like* as an approximative adverb still has to be investigated further, especially in British English. Comments such as (36), (37) and (38) show how *like* is used instead of other adverbs and is corrected to *about, approximately* and *around*. Replacing such adverbs with *like* is often considered incorrect, irritating and redundant, as indeed the comments in (39) and (40) illustrate.

(39) ‘Like’ is superfluous.  
(Retired accounts clerk, over 60 years old, female)

(40) please can we scrap ‘like’ from the English dictionary  
(Publishing, 31–40 years old, male)

While (39) serves as a representative example of many comments on the unnecessary use of *like* in the stimulus sentence, the respondent in comment (40) expresses negative sentiments towards the use of *like* in general, in that he thinks it should be removed from dictionaries. Given the stimulus sentence’s overall negative perception, it does not come as a surprise to find no positive comments on the use of *like* with the informants.

Since studies on the approximative adverb *like* are not only relatively rare, with its distinction from other functions of *like* being relatively recent (cf. D’Arcy, 2006), this feature’s function as an approximative adverb is not yet discussed in usage guides included in HUGE. *Like* is mainly discussed in terms of its alleged misuse and occurrence in place of the conjunction *as* in comparatives such as in the example quoted in Howard (1993, p. 32): “She did it perfectly like she always does”. Taggart’s *Her Ladyship’s Guide to the Queen’s English* (2010), however, also contains a condemnation of the function of *like* as a discourse particle. Commenting on the use of *like*, Taggart states that it is “one of the most overused and misused words in English and
was, even before the distressing colloquialism *And I’m, like, yeah, whatever came into being*” (2010, p. 76). In a similarly strict fashion, Heffer (2010, p. 125) describes discursive *like* as “abominable usage” characteristically used by young speakers. It seems as if *like*’s function as a discourse particle overshadows its other functions and is yet to be picked up by usage guide authors. Nevertheless, the questionnaire data as well as the findings of recent studies, in particular D’Arcy’s works, hint at the increasing importance of *like* as an approximative adverb as well.

To summarize my analysis, I have shown that while many functions of *like* have been extensively investigated, such as its use as a discourse particle and quotative, the feature’s approximative function has only recently attracted scholars’ attention. D’Arcy (2006) shows, however, how in Toronto English *like* is replacing other standard approximative adverbs, in particular *about*. As Chambers mentions in respect to lexical changes, perceptions of the change are perceived differently in society and associated with a particular group. My findings show that usage attitudes towards the use of *like* as an approximative adverb revealed that age shows indeed a statistically significant correlation with acceptability ratings, which means that older informants are less likely to accept the use of *like* in the stimulus sentence used in this investigation. The importance of the main effect of age was confirmed in the binary logistic regression analysis. That this particular feature obtained a low average acceptability rating of only 17.9 per cent illustrates how almost half the respondents find this usage feature unacceptable or barely acceptable in formal contexts, as illustrated in Figure 7.7 above. If *like* is found to be acceptable, it is clearly associated with informal contexts.

The corpus evidence showed that *about* is the most frequent approximative adverb in the BNC. Nonetheless, occurrences of the use of *like* as an approximative adverb have also been recorded. The time period covered by
the BNC needs to be taken into account here. *Like* as an approximative adverb is not yet included in the advice literature. This phenomenon has already been described in that usage guides are a reaction to language use rather than a preemptive measure against future usage. Yet, my qualitative analysis of comments strengthens my assumption that *like* as an approximative adverb can be considered an emerging usage problem.

### 7.2.1.5. Americanisms: *burglarize*

As described in Section 6.6, the verb *burglarize* is the product of a word formation process which is realised by the addition of the derivational suffix *-ize*. However, *burglarize* is not a very common word in British English, as can be seen from the corpus search. While the *-ize* spelling did not produce any results for the word in the BNC, the *-ise* variant showed only one occurrence in the Miscellaneous subsection of the corpus, which under closer inspection turned out to be a religious text from the period 1975 to 1984. Conducting the same searches in COCA, it does not come as a surprise to find no occurrences of the *-ise* spelling variant, as this spelling is considered to be British (cf. Biber et al., 1999, p. 402). Yet the *-ize* variant produced an overall frequency rate of 39 counts. The spoken subsection as well as the newspaper subsection produced the highest standardised frequency ratings of 0.19 and 0.07 tokens per million words respectively. Hence, the corpus search does indeed indicate that *burglarize* is more frequent in American English than in British English.

By including the stimulus sentence (S5. *The bank was burglarized twice last week*) in the online questionnaire, attitudes towards the use of *burglarize* were elicited. The statistical analysis of the elicited usage attitudes showed no correlation between usage ratings and the social variables investigated. That neither age (*U = 1068, p = .353, r = −.09*), education level of the questionnaire respondents, i.e. whether they were university-educated or not (*r_s = .113, ...
Current Usage Attitudes in England: the Online Questionnaire (Part One) 233

$p = .237$, gender ($\chi^2 (1) = .107, p = .743$), nor nativeness ($r_s = -.147, p = .121$) showed a significant difference which could be indicative of a possible consensus within the speech community. A binary logistic regression analysis revealed that a model excluding the social variables investigated has a prediction accuracy of 74.1 per cent which, however, decreases to 72.3 per cent in a model containing all predictors. This model is presented in Table 7.19 below.

Table 7.19 Results of binary logistic regression: burglarize

<table>
<thead>
<tr>
<th>Included</th>
<th>$B(\text{SE})$</th>
<th>Lower exp $b$</th>
<th>Upper exp $b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nativeness</td>
<td>0.86 (0.72)</td>
<td>0.58</td>
<td>2.37</td>
</tr>
<tr>
<td>Gender</td>
<td>0.17 (0.46)</td>
<td>0.48</td>
<td>1.19</td>
</tr>
<tr>
<td>Age (18–25)</td>
<td>0.43 (0.77)</td>
<td>0.34</td>
<td>1.54</td>
</tr>
<tr>
<td>Age (26–30)</td>
<td>0.58 (0.76)</td>
<td>0.40</td>
<td>1.78</td>
</tr>
<tr>
<td>Age (31–40)</td>
<td>1.00 (0.66)</td>
<td>0.74</td>
<td>2.07</td>
</tr>
<tr>
<td>Age (41–60)</td>
<td>0.96 (0.71)</td>
<td>0.66</td>
<td>2.62</td>
</tr>
<tr>
<td>Level of education</td>
<td>-0.80 (0.83)</td>
<td>0.09</td>
<td>0.45</td>
</tr>
<tr>
<td>Constant</td>
<td>-1.74 (0.57)*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: $R^2 = .06$ (Hosmer & Lemeshow), Model $\chi^2 (7) = 7.24, p = .405, p < .05^*$

The model’s fit was determined as good ($\chi^2 (7) = 7.24, p = .405$). As can be seen from Table 7.19 above, none of the social variables is able to predict the acceptability judgment, which is in line with the findings of the non-parametric tests. As for the contextual analysis of usage judgments, the overwhelming unacceptable-rating of 74.1 per cent of the stimulus sentence supports the findings of the sociolinguistic analysis. Figure 7.8 below shows how the stimulus sentence produced a fairly evenly distributed acceptability judgment in the other contexts, with the informal speaking context showing a slightly greater acceptability rating of 19.1 per cent than the other contexts. The stimulus sentence obtained an average acceptability rating of only 13.4 per cent.
While there was no significant correlation between the usage judgment of the respondents and the basis of their judgments ($r_s = .143$, $p = .133$, Fisher’s exact $p = .117$), a significant difference was found with regard to the respondents’ certainty about their judgment ($U = 807$, $p = .001$, $r = -.28$). Despite finding this difference, the median states that both groups were absolutely certain about their judgment ($Mdn = “absolutely certain”)$. The contextual preference of this particular usage feature thus seems to be in line with my sociolinguistic analysis of usage attitudes and reinforces the argument for a consensual agreement about the unacceptability of *burglarize*.

The above-mentioned finding indicating a consensus about the stimulus sentence’s unacceptability is further strengthened by the analysis of the 59 comments obtained from the questionnaire respondents, who frequently mention the foreignness of *burglarize* and even claim doubting the existence and legitimacy of this particular word. Comments (41) – (43) below illustrate the foreignness of *burglarize* and identify it as an American usage feature.
(41) ‘Burglarized’ is a vulgar Yankeeism.
(Retired dental surgeon, over 60 years old, male)

(42) burglarized? burgled, I suppose you mean! Horrible Americanism.
(Retired arts consultant, over 60 years old, female)

(43) ‘To burglarise’ is an over-the-top Americanism. It’s ‘to burgle’ - or simply ‘rob’.
(Student, 18-25 years old, male)

Calling burglarized a “vulgar Yankeeism”, the respondent in (41) condemns this particular usage in no uncertain terms. In a similarly strong manner, the retired arts consultant in (42) corrects burglarize to burgle before describing burglarized as an “[h]orrible Americanism”. The description of this usage feature as “an over-the-top Americanism” in (43) does not seem as harsh as the comments (41) and (42). About half of all comments on burglarize hinted at its foreignness and frequently associated it with American English.

An equally frequently recurring label is the condemnation of burglarize as a “non-word”, a notion which is also often connected to ain’t in American English (Curzan, 2014, p. 60). By stating that burglarize is not a word, respondents enforced the legitimacy of the British variant to burgle. Comments (44) – (46) illustrate this theme.

(44) burglarized is not a word in any circumstance and so I would correct this.
(English teacher, 18-25 years old, female)

(45) Burglarized simply isn’t a word in my dialect.
(Software engineer, 31–40 years old, male)

(46) Burglarized is not a word.
(Accountant, 26–30 years old, female)

While the respondent in (44) claims she would correct burglarized as it was “not a word in any circumstance”, not even considering its possible American
status, the respondent in (45) restricts his decision to his own dialect, in which he claims *burglarized* cannot be found. This notion is also reflected in (46). While the qualitative analysis has so far shown the respondents’ associations of *burglarize* as an example of American English and as a non-word, other respondents condemn the stimulus sentence containing *burglarize* for other reasons. The comments provide an insight into how the use of *burglarize* will be perceived by some respondents and also into who the respondents believe makes use of this feature. Comments (47) – (49) deal with these issues.

(47) I hate this kind of daft error. Makes the speaker sound idiotic. (Manager in a museum, 41–60 years old, female)

(48) This would mark someone as fairly uneducated (& perhaps trying to appear otherwise) - why not use the shorter *burgled*? (University lecturer, 41–60 years old, male)

(49) Only chavs make up stupid words (Digital marketing consultant, 31–40 years old, female)

All three comments quoted here attribute a lack of intelligence to a speaker making use of this particular usage feature. While the respondent in (47) argues that *burglarize* is a “daft error”, which results in the speaker being perceived as “idiotic”, the respondent in (48) attributes the stimulus sentence to a “fairly uneducated” speaker whose aim may be to sound educated. The last comment (49) marks *burglarized* as a “stupid” word which would only be used by a speaker of lower social standing. The use of the derogatory term ‘chavs’ as alleged users of the Americanism to *burglarize* emphasises this respondent’s negative attitude towards this particular usage feature. The term ‘chavs’ has become widely discussed in the last decade in Great Britain as this derogative term has been used in a discussion on social changes affecting British society through which the white working class has become a demon-
ised, marginal group, according to Jones (2016, pp. 8–9). What can be gathered from the analysis of my respondents’ comments is that the overall negative attitude towards *burglarize* as represented in the stimulus sentence is based on the feature’s associations with American English and its being labelled non-word. In addition, the comments also indicate that the use of *burglarize* as well as its users will be perceived in a strongly negative light.

Since the HUGE database does not include Americanisms as a category of usage problems, I searched for entries which included the term ‘Americanism’. The results of this analysis should therefore not be considered as a mere description of the concept ‘Americanism’. Which features are considered to be part of this category? How many usage guide authors make use of this term? First of all, it is surprising to see only a small number of usage guides mentioning the term ‘Americanism’ in general. Only nine out of the 39 British publications, published between 1770 and 2010, discuss this type of usage problem by referring to them as Americanisms. Yet, it has to be noted that these nine usage guides discuss different usage problems that contain this term, which could hint at the author’s subjective judgment of what should be included in usage guides (Weiner, 1988, p. 175). Some of these usage problems are the already discussed *different than* and *real* as a flat adverb used instead of *really*, lexical items such as *railroad* and *to run for office*, as well as morphosyntactic issues such as the use of the sentence adverbial *hopefully* or *to have got for to have*.

Table 7.20 includes an overview of Americanisms discussed in the nine usage guides. The table includes, however, only Americanisms which were considered controversial, while features which have been adopted into British English, such as the nonstandard *kinda* discussed in Peters (2004, p. 307), have not been included in the overview below. The majority of the usage features included in Table 7.20 below constitute lexical items. The issue of verb
conversions is, however, mentioned in *Pocket Fowler’s Modern English Usage* edited by Allen (1999, p. 311), who criticises this practice using the example of *to hospitalize* which, according to him, is “regarded with some suspicion in BrE, but is standard in AmE”.

Table 7.20 Overview of Americanisms discussed in British publications

<table>
<thead>
<tr>
<th>Author</th>
<th>Americanisms Discussed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vallins1953</td>
<td>to meet up with, to lose out on, I don’t have (for I haven’t got), to aim to do (for aim at doing), the first time in years (in instead of for), due to, like for as, elevator, automobile, fall, sidewalk, way above, way below, way up, way back, baby (to refer to one’s girlfriend), the use of considerable with material things, <em>Do you have a match?</em>, <em>in the event that</em> (for <em>in the event of</em>), <em>fixings</em> (for <em>trimmings</em>), <em>to protest the decision</em> (for <em>to protest the decision</em>), to raise, to reckon, southpaw, to stem from, through (for <em>Monday through Friday</em>)</td>
</tr>
<tr>
<td>Gowers1965</td>
<td>overly, overview, have (got), American pronunciation</td>
</tr>
<tr>
<td>Burchfield/Weiner/Hawkins1984</td>
<td>Oklahoma, I guess, to check up on, to win out, to lose out, the sentence adverb hopefully, color, theater, gotten (for got), dove (for dived), snuck (for sneaked), pronoun he to refer back to impersonal one, real, sidewalk, candies, gas, blank, faucet, comforter, first floor (for ground floor), name for (for name after), public school, pacifier, wash up (for washing one’s face and hands), around (for about), American pronunciation, climax, to consult with, importantly, through (e.g., Monday through Friday), transportation</td>
</tr>
<tr>
<td>Greenbaum&amp;Whitcut1988</td>
<td>overly, overview, have (got), American pronunciation</td>
</tr>
</tbody>
</table>

**Table 7.20** Overview of Americanisms discussed in British publications.
The lack of discussion of this linguistic practice in general in usage guides becomes even more evident when a comparison is made with news style guides such as the *BBC News Styleguide*. Style guides such as the *BBC News Styleguide* (Allen, 2003) or *Guardian Style* (Marsh & Hodson, 2010) are interesting publications, since media institutions make their house style available to the public and so foster their image as language guardians. Allen included an extensive section on Americanisms in the BBC’s style guide, in which verb conversions are discussed as well. Verb conversions are also discussed in *The Guardian and Observer’s Style Guide* (2015), *The Times Style and Usage Guide* (2003), *The Telegraph Style Guide* (2010) and *Guardian Style* (2010), though only the verbs *to impact* and *to hospitalise* are discussed there. It thus seems that the issue is more typically considered a feature that should be discussed in newspaper style guides than in usage guides, which suggests that
they may be considered more typical of journalistic jargon than of usage generally.

Turning nouns into verbs is a linguistic practice which is frequently associated with American English. Yet, *euthanize*, an often-mentioned example of alleged American practice, was first recorded in the English *The Times* in the mid-twentieth century, so it appears to have been common practice in British English, too. Nevertheless, words such as *euthanize* and the investigated *burglarize* have become prototypical examples of American English usage, which is confirmed through the findings of my corpus search. *Burglarize* does indeed occur more frequently in COCA than in the BNC.

Verb conversions as a linguistic practice of turning nouns into verbs, either by maintaining the noun or by adding derivational suffixes, is viewed critically by the questionnaire respondents, as becomes apparent from the fact that it obtained a high unacceptability rating of 74.1 percent and a low average acceptability rating of merely 13.4 per cent. Analysing the comments obtained from the respondents confirms the word’s association with American English – in the eyes of my British informants, that is. This association is explained through the process of what is referred to as “othering” (Dervin, 2012, p. 187) and associating *burglarize* as a non-word. The comments also revealed how users of the word would be perceived by the respondents, who state that its use is associated with a lack of education as well as intelligence. Given the feature’s high unacceptability rate, it does not come as a surprise to find that none of the social variables investigated in this study correlate with the acceptability judgment.

Since the British usage guides in HUGE do not include Americanisms as a category as such, I conducted a full-text search of the database using the term ‘Americanism’, which revealed that only nine of the 39 British publications actually discuss Americanisms. What this search also brought to light
is the subjective judgments of the usage guide authors as to what constitutes an Americanism. That these nine usage guides rarely discuss verb conversions such as *burglarize* is not unexpected in light of the fact that this feature is often considered to be an Americanism. A comparison with style guides, such as the *BBC’s News Styleguide* (Allen, 2003) but also various others, makes this even more intriguing, as verb conversions such as *to hospitalize* and *to impact* are frequently discussed in the style guides consulted. This strongly suggests that this feature is associated with journalistic jargon.

### 7.2.1.6. Less than

Having described the issue with using *less* with countable nouns in Section 6.7 above, I will now turn to the analysis of the perception data as well as corpus data. My corpus search for *less than* followed by a cardinal number plus noun or noun phrase showed that *less than* frequently collocates in this pattern with *per*, *cent*, *years*, *hours*, *months* and *weeks*, with *less than two years* being the most frequent pattern occurring in the BNC (73 times). A total of 1,248 instances including *less than* being followed by a countable noun are included in the BNC. By contrast, *fewer than* followed by a cardinal number plus noun was much rarer, showing a frequency rate of 127 occurrences. The corpus search brought to light collocations of *fewer than* with *per*, *people*, *times*, *employees*, *countries* and *years*. The most frequent pattern, *fewer than 20 people*, occurred only four times in the BNC. The collocates of *less* and *fewer* reveal a difference in that *less than* seems to be used more frequently for temporal indications, as in example (7.a) below, while *fewer than* occurs more frequently with countable nouns such as *employees* and *countries* as shown in (7.b). However, the corpus analysis shows that *less than* is also indeed used with countable nouns.
7. a) The money has been raised in less than two years. (K97, written (newspaper))

b) Almost 90 per cent of companies have fewer than fifty employees. (B1U, written (academic))

Having presented the corpus evidence of actual usage, I will turn now to the analysis of the usage attitudes of the questionnaire respondents towards the use of less than in the stimulus sentence (S6. Pay here if you have less than 10 items). Similar to the usage problem burglarize discussed in Section 7.2.1.5, no significant correlation between the social variables age ($U = 1219$, $p = .686$, $r = -.04$), education level ($r_s = .041$, $p = .664$), gender ($\chi^2 (1) = .514$, $p = .473$) and nativeness ($r_s = -.142$, $p = .134$), and acceptability ratings could be identified. The binary logistic regression analysis revealed that a model excluding predictors, i.e. the social variables investigated, has a prediction accuracy of 71.4 per cent. The application of the forced data entry method in which all social variables are included in the model resulted in the proposed model in Table 7.21 below. Its prediction accuracy does not increase and remains at 71.4 per cent.

Table 7.21 Results of binary logistic regression: less than

<table>
<thead>
<tr>
<th>Included</th>
<th>95 % CI for exp b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Included</td>
</tr>
<tr>
<td>Nativeness</td>
<td>1.51 (1.11)</td>
</tr>
<tr>
<td>Gender</td>
<td>0.39 (0.44)</td>
</tr>
<tr>
<td>Age (18–25)</td>
<td>0.36 (0.69)</td>
</tr>
<tr>
<td>Age (26–30)</td>
<td>–0.29 (0.64)</td>
</tr>
<tr>
<td>Age (31–40)</td>
<td>0.26 (0.61)</td>
</tr>
<tr>
<td>Age (41–60)</td>
<td>0.22 (0.66)</td>
</tr>
<tr>
<td>Level of education</td>
<td>–0.19 (0.63)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.55 (0.46)</td>
</tr>
</tbody>
</table>

Note $R^2 = .04$ (Hosmer & Lemeshow), Model $\chi^2 (7) = 4.75$, $p = .691$, $p < .05^*$
The model presented in the table proves to be a good fit ($\chi^2(7) = 4.75$, $p = .691$). The outcome of the binary logistic regression analysis confirms the findings of the non-parametric tests while taking into account any possible covariance between the four social variables. This means that attitudes towards the use of less than do not seem to vary according to any social group.

Analysing the contextual preference of the stimulus sentence showed that the questionnaire respondents rated this particular feature as fairly acceptable, in particular in the informal contexts in which less than was considered acceptable by the majority of the respondents. While less than obtained its highest acceptability rating of 60.7 per cent in the informal spoken context, the stimulus sentence seems to be less acceptable in formal contexts, which showed the lowest acceptability rating of only 22.3 per cent in the formal writing context. Less than obtained an average acceptability rating of 40.2 per cent. An overview of the contextual preference of less than is presented in Figure 7.9 below.

![Figure 7.9 Contextual acceptability in percentages: less than](image-url)
Despite identifying a significant difference between acceptability judgments and the degree of certainty ($U = 759, p = .000, r = -.39$), both groups indicated that they were absolutely certain about their judgment ($Mdn = \text{“absolutely certain”}$). However, a clear statistical difference was found between those who find less than acceptable and those who do not, with their judgment basis showing a positive moderate correlation ($r_s = .418$, $p = .000$, Fisher’s exact test $p = .000$). This means that those who rated the stimulus sentence as acceptable said that they based their judgment on a gut feeling, while those who found it unacceptable based their decision on self-reported rule knowledge.

Analysing the metalinguistic comments ($n = 51$) made by the questionnaire respondents on this particular usage feature showed a high degree of awareness about the usage conundrum revolving around the use of less with countable nouns. The two most prominent themes I would like to discuss here are providing a correction and the widespread acceptability of less instead of fewer (see further Appendix I). Comments (50) – (53) represent the first theme and deal with the arguments used by the respondents as to why fewer should be used in the stimulus sentence instead of less.

(50) People seem to have given up using the word “fewer”, which is correct when talking of a number of items.  
(Retired, over 60 years old, female)

(51) Fewer. FEWER than 10 items. You can’t chop an item in half.  
(Youth worker, 26–30 years old, female)

(52) FEWER!! It’s countable.  
(Literacy Consultant, 41–60 years old, female)

(53) ‘fewer’ is correct here because you can count ‘items’ - old rule about countable and un-countable [sic] things applies.  
(Retired school teacher, over 60 years old, male)
All four comments selected here argue for the countability of *items*. A youth worker elaborates on her correction of *less* to *fewer* in (51) by stating that one cannot “chop an item in half”. A retired school teacher in (53) not only emphasises the correctness of *fewer*, but also explains the necessity of applying the “old rule about countable and un-countable [sic] things”. What becomes apparent from the analysis of the comments is how the traditional rule advocating *fewer* still seems to hold true for some respondents. The above-mentioned argument of using *less* with countable nouns when quantity is regarded as a single inclusive unity was not mentioned by any of the respondents. What was, however, mentioned was the widespread acceptability of *less* than with countable nouns. This theme, which will be exemplified on the basis of comments (54) – (57), often co-occurred with providing a correction.

(54) Grammatically, it should be ‘fewer than’ for countable nouns. However, in reality, this ‘rule’ is widely broken.
(Teacher, 31–40 years old, female)

(55) Should be ‘fewer,’ but nearly everyone says ‘less.’
(Retired solicitor, over 60 years old, male)

(56) Sadly the difference between *less* and *fewer* is just about dead now.
(Retired educational publisher, over 60 years old, male)

(57) The ‘less’/ ‘fewer’ distinction is a lost cause.
(Education adviser, over 60 years old, male)

*Less* being used for countable nouns is found to be so widespread that it is considered acceptable by respondents such as those in (54) and (55), who both provide a correction of *less* by changing it to *fewer*. However, the loss of the distinction is bemoaned as well, as by the retired educational publisher in (56) and the education adviser in (57). The comments show that there seems to be a clear divide between those who have accepted the change, albeit somewhat
grudgingly, and those who insist on the traditional use of *fewer*. This divide is aptly illustrated by the two respondents in comments (58) and (59) below.

(58) Simple rule here: anyone who get upset about *fewer*/*less* distinctions needs to get out more (or to stay in less!).
(Teacher, 41–60 years old, male)

(59) Should be *fewer than*, but it is used so often in shops catering for the ignorant that one has come to accept it.
(Retired, over 60 years old, male)

While the teacher in (58) argues in favour of using *less* with countable nouns by referring to common usage, a notion which can also be identified in (55) above, the retired male respondent in (59) provides a correction and argues that shop signs using *less than* with countable nouns cater “for the ignorant”.

Out of the 39 British publications in HUGE, 28 usage guides discuss the distinction between *less than* and *fewer than*. This particular usage feature was first dealt with critically in Robert Baker’s *Reflections on the English Language* (1770), while the three most recent usage guides included in HUGE continue this feature’s stigmatisation. Baker’s advice is provided together with two other usage guide entries below to illustrate the categorisation applied.

**Criticised**

This Word is most commonly used in speaking of a Number; where I should think *Fewer* would do better. *No fewer than a Hundred* appears to me not only more elegant than *No less than a Hundred*, but more strictly proper. (Bakker, 1770, p. 55)

**Neutral**

Strictly speaking, the rule is that fewer, the comparative form of few, is used with words referring to countable things, including people: *fewer books*; *fewer than ten contestants*. Less, on the other hand, is used with things which cannot be counted: *less money*; *less music*. In addition, less is normally used with numbers when they are on their own, e.g. *less than 10,000*, and with expressions of measurement or time: *less than two weeks*; *less than four miles away*. To use less with
countable things, as in less words or less people, is widely regarded as incorrect in standard English. It is a well-known usage point in English – so much so that an upmarket British store chain was forced by public demand to change the check-out signs in its food supermarkets from ‘Less than five items’ to ‘Fewer than five items’. (Oxford A–Z, 2007, p. 59)

**Advocated** When you have a plural noun describing items that can be counted, use ‘fewer’. If it’s a plural noun describing items that can’t be counted, use ‘less’. Fewer books, fewer data, fewer people. Less work, less food, less coverage. The exception is when ‘less’ is followed by ‘than’ and an amount of something (such as distance, time or money). It’s less than 3 metres. It takes less than an hour to get there. It costs less than £5. (Sayce, 2006, p. 53, bold in original)

Categorising the usage entries into advocated, neutral and criticised usage advice highlights a shift in how the distinction between fewer and less is discussed. The traditional distinction between fewer and less is mostly upheld, yet their comparative usage in the form of fewer than and less than has been subject to change since a number of usage guide authors such as Sayce (2006, p. 53) quoted above advocate the use of less than with countable nouns. The reasoning of usage guide authors is based on the above-mentioned notion that the following noun should be considered a single inclusive unit. Frequently mentioned exceptions of less than with countable nouns are measurements of time and distance (cf. Howard, 1993; Burt, 2002; Sayce, 2006). Table 7.22 provides a detailed overview of the categorisation of the 28 usage guides dealing with this usage problem.

Table 7.22 below indicates how a shift in perception most likely has occurred through which less than has become an acceptable exception to the traditional prescriptive rule, which is still overwhelmingly promoted in seventeen usage guides of the 28 usage guides. By providing a diachronic illustration of the categorisation in Figure 7.10, it is possible to pinpoint the decade in which less than started to be considered acceptable. The first neutral
advice with respect to the use of *less than* can be found in Swan’s *Practical English Usage* published in 1980. From the 1990s onwards, a tendency towards more lenient usage advice on the *less than* issue can be identified as more usage guide authors deemed *less than* acceptable.

Table 7.22 Treatment of *less than* (“criticised”, “neutral”, and “advocated”) in British publications

<table>
<thead>
<tr>
<th></th>
<th>Critics (17)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Neutral (4)</td>
</tr>
<tr>
<td></td>
<td>Advocated (7)</td>
</tr>
<tr>
<td></td>
<td>Total: 28</td>
</tr>
</tbody>
</table>

Figure 7.10 Diachronic treatment of *less than* in British publications
Despite this fairly recent increase in acceptability in the usage guides, a considerable number of usage guide authors continue to dismiss the use of *less than* with countable nouns. Most notably, the three most recent usage guides included in HUGE promote the traditional distinction between *less* and *fewer* (cf. Taggart, 2010; Heffer, 2010; Lamb, 2010). Mentioning the supermarket signage debate, Lamb (2010, p. 170), for example, shows his support for the prescriptive corrections by stating the following: “A supermarket was persuaded to change the signs *Five items or less* to *Five items or fewer*, which is correct”.

The traditional distinction between *fewer than* and *less than* based on the countability of nouns seems to constitute a recurring usage problem in the usage debate. Whether constructions such as *10 items* should be treated as countable or not depends on whether they are perceived as an inclusive unity or as single entities. My corpus search showed an interesting difference between collocation patterns in that *less than* seems to collocate more frequently with temporal quantities, such as *years, months, hours* and *weeks*, than *fewer than* does. The majority of the questionnaire respondents indicated that the stimulus sentence was acceptable in informal contexts. The average acceptability rating of 40.2 per cent suggests that the stimulus sentence has a fairly high acceptability rate in comparison with other usage problems discussed in this study (see §. 6.2.5). Only 28.6 per cent of all usage judgments fell into the unacceptable category. Therefore, it also does not come as a surprise to find no sociolinguistic stratification of the questionnaire respondents’ usage attitudes according to any of the social variables investigated. The usage feature’s general acceptability is also mentioned by the respondents in the obtained comments.

As for the HUGE analysis, the distinction between *less than* and *fewer than* is indeed a recurring problem in the usage debate and it may therefore be
considered an old chestnut, all the more so since this particular item has been
present in the usage guides since the tradition’s earliest days. While its first
critical treatment can already be found in Baker’s *Reflections on the English
Language* (1770), the usage feature has only recently become viewed as ac-
ceptable. Nevertheless, it has to be mentioned that this feature is still widely
condemned by usage guide authors who take a prescriptive stance on this
issue, such as the three latest additions to the HUGE database.

7.2.1.7. The double negative
As described in Section 6.8 above, the stimulus sentence (S7. *He wasn’t seen
nowhere after the incident*) was included in the questionnaire to elicit attitudes
towards the use of double negatives. Before these attitudes are investigated in
regard to their sociolinguistic stratification, I will provide an overview of the
occurrence of double negatives in the BNC. Using the POS-tagger, a BNC
corpus search for the double negative used in the stimulus sentence and its
standard variant *He wasn’t seen anywhere* was conducted to identify patterns
including the contraction of *not* being followed by a verb and the negative
adverb *nowhere*. Sixteen patterns were identified, with the most frequent one
being *n’t going nowhere* (*n* = 5), *n’t go nowhere* (*n* = 5), *n’t got nowhere*
(*n* = 3). While eleven of these sixteen occurrences can be found in the spoken
subsection of the BNC with a normalised frequency rate of 1.10 tokens per
million words, four appeared in the fiction subsection, which reflects a nor-
malised frequency rate of 0.25 tokens per million words, and one single occur-
rence was recorded in the non-academic subsection reflecting a normalised
frequency rate of 0.06 tokens per million words. That only sixteen occurrences
of this particular double negative pattern are recorded in the BNC does not
come as a surprise, as written texts are usually edited and double negatives are
usually not considered acceptable in writing. The four instances of double
negatives in the fiction subsection are all found in reported speech, which indicates that the double negative is a feature frequently found in spoken contexts. By contrast, a corpus search substituting the Standard English variant *anywhere* for *nowhere* in the sentence showed a considerably higher overall frequency rate of 237 identified patterns of which *n’t go anywhere* (*n* = 51), *n’t get anywhere* (*n* = 49) and *n’t going anywhere* (*n* = 18) are the most frequent. Besides the generally higher number of patterns including the standard variant, all subsections of the corpus contain these occurrences, with the spoken subsection showing the highest normalised frequency rate of 9.23 tokens per million words followed by the subsections fiction (4.48 tokens per million words), magazine (2.20 tokens per million words) and newspapers (2.01 tokens per million words).

Turning to the sociolinguistic analysis of usage attitudes, I was able to show that none of the social variables investigated in this study showed a significant correlation with acceptability ratings for this feature. While the social variables education level (*r* = .078, *p* = .416), gender (*χ²*(1) = .313, *p* = .576) and nativeness (*r* = –.047, *p* = .623) resulted in clear non-significant correlations, age (*U* = 815, *p* = .079, *r* = –.17) showed the lowest *p*-value and was closest to statistical significance. In order to identify any possible influence of the independent variables on the acceptability judgment, I conducted a binary logistic regression analysis which accounts for any possible covariance between these independent variables. The binary logistic regression analysis showed that a model including only the constant, which means that none of the social variables were included in the model first, has an overall prediction accuracy of 78.6 per cent. The inclusion of all social variables in the proposed model, which is illustrated in Table 7.23 below, does not, however, change the prediction accuracy.
Table 7.23 Results of binary logistic regression: double negative

<table>
<thead>
<tr>
<th></th>
<th>95 % CI for exp b</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Included</td>
</tr>
<tr>
<td>Nativeness</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Age (18–25)</td>
<td></td>
</tr>
<tr>
<td>Age (26–30)</td>
<td></td>
</tr>
<tr>
<td>Age (31–40)</td>
<td></td>
</tr>
<tr>
<td>Age (41–60)</td>
<td></td>
</tr>
<tr>
<td>Level of education</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td></td>
</tr>
</tbody>
</table>

Note: $R^2 = .07$ (Hosmer & Lemeshow, Model $\chi^2 (7) = 8.54, p = .287, p < .05^*$

While the model presented in the table above is considered a good fit ($\chi^2 (7) = 8.54, p = .287$), the Wald statistics indicate that the overall effect of age is not significant ($p = .177$) in spite of the seemingly meaningful contribution of the group of 18–25-year-olds to the model when compared to the reference group of over-60-year-olds ($p = .032$). The exp $b$ value indicating the changes in the odds ratios in the comparison of these two age groups furthermore highlights the odds of making a negative acceptability judgment increase with age ($OR = 5.55$). Yet, the overall effect of age is not significant, as the covariance of all social variables is taken into account in the binary logistic regression.

The contextual preference analysis revealed an overwhelming rejection of the stimulus sentence: 78.6 per cent deemed *He wasn’t seen nowhere after the incident* as unacceptable. Few respondents found the stimulus sentence with the double negative acceptable in formal contexts. The formal writing context shows the lowest acceptability rate of 0.9 per cent. Surprisingly, this trend is found in the informal contexts as well. Only the informal spoken context registers 18.8 per cent of acceptable judgments. Thus, the low average acceptability rating of 7.6 per cent does not come as a surprise. Figure 7.11 below illustrates the contextual preference ratings in detail.
With an average acceptability rating of merely 7.6 per cent, the double negative seems to be widely considered unacceptable. This overall rejection of the double negative by the majority of respondents is reflected in the respondents’ high degree of certainty of their judgments. A statistical difference between acceptability and unacceptability ratings was identified ($U = 608$, $p = .000$, $r = -.52$), which showed, however, that the median of both groups was “absolutely certain”. Nonetheless, an analysis of the judgment basis resulted in no significant differences between those who found the stimulus sentence acceptable and those who did not ($r_s = .010$, $p = .919$, Fisher’s exact test $p = 1.000$).

51 respondents provided a comment on the stimulus sentence including the double negative. My analysis of these comments revealed that there was a high degree of awareness of double negatives among the respondents, which not only showed in the fact that respondents were able to identify the issue
using the correct terminology, but also in their corrections to the standard variant *anywhere*. This awareness is also accompanied with comments indicating a possible confusion caused by the double negative. The basis for this confusion is the above-mentioned rule of logic (cf. § 6.8) according to which two negatives make a positive, which has also been commented on by respondents. Comments (60) – (63) are examples of these findings.

(60) *Nowhere* should be *anywhere.*
   (Civil Servant, 31–40 years old, female)

(61) Double negative = positive
   (Stay-at-home mother, 31–40 years old, female)

   (Proof-reader, 41–60 years old, male)

(63) Double negative. Contradictory.
   (Old Nuisance, over 60 years old, male)

The comment provided by a civil servant who corrected *nowhere* into *anywhere* (60) serves as an example of the corrections made by questionnaire respondents. Comments (61) – (63) show the respondents’ awareness of the usage feature by the fact that they name the issue and provide an insight into their perception of double negatives causing ambiguity by applying the rule of logic. The rule is quoted in (61), while the informants in (62) and (63) argue that the double negative causes ambiguity or confusion.

How users of double negatives are perceived by the questionnaire respondents has also surfaced in the comments obtained and is illustrated by comments (64) – (67) below.

(64) Double negatives are the domain of stupids.
    (Digital marketing consultant, 31–40 years old, female)
In ‘standard English’ is should be ‘anywhere’; the use of the double negative would be fine in some contexts but could make the user sound uneducated in others.
(Teacher, 31–40 years old, female)

You can’t expect those with limited education to speak in correct sentences.
(Retired, private research, over 60 years old, female)

‘wasn’t seen nowhere’ sounds childish, but you would not expect a child to use the word incident.
(Engineering, 31–40 years old, male)

What becomes evident through the comments quoted here is the association of double negatives with a lack of education on the part of speakers who use them. While the respondent in (64) considers double negatives part of “the domain of stupids”, comments (65) and (66) make a similar, yet less harsh connection between the feature and lack of education. Lastly, the respondent in (67) argues that the double negative comes across as “childish”, which from a broader point of view could be connected to a lack of education.

Having discussed how users of double negatives are perceived by the questionnaire respondents, I would briefly like to discuss a number of comments which describe not only the notion of personal usage versus that of others, but also enable an insight into where double negatives are normally found.

This sounds awful. I come from a part of the UK that has a strong regional accent - I would never say this, even though I’ve heard it many times.
(Accountant, 26–30 years old, female)

This is standard usage in other dialects, but not mine.
(Software engineer, 31–40 years old, male)

Not for me unless I’m putting on an accent.
(Retired, over 60 years old, female)
(71) The sentence has a double negative – we are not in Spain.
(Retired dental surgeon, over 60 years old, male)

While all four comments describe double negatives as alien to their own personal usage, it becomes obvious that they are identified as dialectal features or even, as in (71), foreign to English. The respondent in (68) recounts how the area she is from “has a strong regional accent”, yet despite having heard double negatives many times, she claims never to use them herself. A similar notion is expressed by the respondent in (69). Comment (70) is interesting as the respondent states that she would use this kind of usage feature only when “putting on an accent”. This is indicative of a sociolinguistic practice called “stylising the other” which is defined by Rampton (1999, p. 421) as “a range of ways in which people use language and dialect in discursive practice to appropriate, explore, reproduce or challenge influential images and stereotypes of groups that they don’t themselves (straightforwardly) belong to”. As (70) shows, the respondent states that double negatives are not part of her language variety, yet they invoke to her an image or even stereotype she has concerning speakers who do use them. While these three comments acknowledge double negatives as English features, the retired dental surgeon in (71) distances himself from double negatives by placing them outside Great Britain.

Double negatives are discussed in 27 of the 39 British usage guides included in HUGE. Robert Baker’s *Reflections on the English Language* (1770) contains the first discussion of them. The categorisation of usage entries, however, showed that none of the usage guides advocate their use, which is most likely due to their aim of enforcing and promoting standard language features. Example entries of the “criticised” and “neutral” categories are found below, followed by a complete overview of the categorisation of usage entries in Table 7.24.
Avoid double negatives. They are offences against logic and, if they are an attempt at being funny, they fail. Sometimes they occur by accident: not obvious stinkers like “he said he would not never go there”, which can only be the product of illiteracy and stupidity, but a phrase … such as “of all the casualties, she was the least unscathed”. Unscathed is entirely the wrong word; it needs to be one of its antonyms, injured or hurt. …The obvious ones are not the problem, for all but the illiterate will spot them; it is those that come in longer sentences, usually including verbs that themselves have a negative import, that cause unexpected difficulties. Be especially alert to sentences such as “I cannot doubt that there may not be times when you feel like that”, which will confound almost everyone who reads or hears it. (Heffer, 2010, pp. 57–58, 132)

… So, if you don’t like double negative constructions, because you’re a standard English user, then that’s your privilege. You’ve been brought up that way. But don’t fall into the trap of thinking that there’s something intrinsically more logical about speaking or writing in that way. Or go searching for ambiguity where there are none to be found. If people say I ain’t got no money, they’ll never be misunderstood. Would you really wish to argue that these speakers are saying that they have got some? If people want to speak or write standard English, then, they’ve got to learn to avoid this kind of double negative construction, otherwise they’ll be severely criticized. But the critics must learn to criticize for the right reasons: it’s socially unacceptable, but it isn’t illogical. … (Crystal, 1984/2000, p. 46)

Table 7.24 below shows that twice as many British usage guides discuss double negatives critically than in a neutral manner. The two examples stated above show that while Heffer (2010, p. 132) associates double negatives with illiteracy, Crystal (1984, p. 46) emphasises the universal and widespread character of this linguistic feature which accidentally is absent from Standard English. He continues by advising his readers to avoid double negatives when using this particular variety, but warns them against being judgmental and treating this feature as inferior. The diachronic development of the treatment of double negatives is illustrated in Figure 7.12 below.
Table 7.24 Treatment of double negative (“criticised” and “neutral” only) in British publications

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total: 27</td>
<td></td>
</tr>
</tbody>
</table>

Despite being discussed in the oldest usage guide in the HUGEC collection, double negatives are not dealt with until the 1920s, with the notable exception of The Vulgarities of Speech Corrected (1829). From the 1980s onwards, double negatives seem to have come under increased criticism, as can be seen in Figure 7.12.

Figure 7.12 Diachronic treatment of double negatives in British publications

That the majority of usage guides investigated criticise the use of double negatives hints at the currency of this particular usage feature, which is most
likely due to their aim of maintaining the standard variety in this respect. The widespread use of double negatives in British dialectal varieties makes the double negative an interesting feature for analysis, particularly so since it is not considered to be part of Standard English, which is after all the domain within which usage guides operate. The assumption that two negatives make a positive, which stems from logic and has been dismissed by Pullum and Huddleston (2002, p. 847), still seems to hold true for some of the questionnaire respondents who argue that double negatives could cause ambiguity.

This feature has produced an overwhelmingly negative result from the questionnaire respondents, with almost 79 per cent calling the stimulus sentence unacceptable, which confirms the findings of Tieken-Boon van Ostade (2005) and De France (2010) discussed above (see § 6.8). This consequently resulted in a very low average acceptability score of only 7.6 per cent. When the stimulus sentence was found acceptable, this was the case in the informal contexts with informal speech showing the highest acceptability rate of 18.8 per cent (cf. Figure 7.12). Such an overall negative evaluation of the double negative indicates a relatively high probability that usage judgments will not vary according to any of the social variables investigated, i.e. age, gender, nativeness and level of education. This finding was also confirmed by the results of the binary logistic regression analysis.

My analysis of the comments provided by the questionnaire respondents brought to light a number of interesting themes which help us understand the respondents’ judgments better. Based on this analysis the negative perception of users of double negatives became clear in that an association was made between double negatives and a lack of education. Drawing on stereotypical images of speakers, some of the respondents distance themselves from the nonstandard use of double negatives, despite making remarks on their
widespread occurrence even in their own direct vicinity. Interestingly, this stereotypical image of uneducated double negative users is reproduced by respondents who comment on making use of this linguistic feature when “putting on an accent” (74). This form of language styling further distances the respondents’ language use from the usage of others.

The HUGE analysis showed not only that double negatives were first discussed in Baker’s *Reflection on the English Language* (1770), but it further highlights how, apart from one usage guide published in 1829, the double negative seems very much a usage problem of the twentieth century. Its overwhelmingly negative treatment in the advice literature can be traced back to the aim of such usage guides to promote and fix the standard language variety.

7.2.1.8. The dangling participle

Dangling participles as a usage problem have been described in detail in Section 6.9 above and the analysis of attitudes towards the stimulus sentence containing this particular feature (S8. *Pulling the trigger, the gun went off*) is presented in the present section. Not only are dangling participles widespread (cf. Bartlett, 1953; Hayase, 2011), they also come in many different shapes, which makes a corpus analysis complex. To make up for this, I will draw on the findings of Hayase’s study, who investigated a selection of 96 types of present participle constructions in the BNC. Among those 96 types are participles such as *approaching, entering, leaving* and *walking*. Restricting the selection to dangling participles solely occurring at the beginning of a phrase, Hayase (2011, pp. 92–93) obtained 956 examples which were divided into five categories based on their description of a situation: “Cognition, Motion, Perception, State, and Action”. Of these five categories, dangling participles falling into the category “Cognition” were the most frequent. An example of such a dangling participle is quoted in Hayase (2011, p. 94): “Comparing them
to the English Baroque woodwinds, it is clear that they became the prototype and standard for English makers well into the 18th century”. In the absence of a separate corpus study undertaken by myself, Hayase’s study at any rate proves the actual occurrence of dangling participles.

As for the analysis of usage attitudes and contextual preferences of the stimulus sentence, the following results were obtained. From my sociolinguistic analysis it can be gathered that no individual age group \((U = 1451, p = .484, r = -.07)\) deviates in their usage judgment. The other social variables investigated, education level \((r_z = -.026, p = .784, r = .036, p = .850)\) and nativeness \((r_z = -.030, p = .754)\), did not show a significant correlation with acceptability ratings either, which was a finding also consolidated in the binary logistic regression analysis. This analysis showed that a model only including the constant has an overall prediction accuracy of 50 per cent. The application of a forced data entry method, however, revealed that the prediction accuracy increased to 56 per cent when all social variables are included in the model presented in Table 7.25.

Table 7.25 Results of binary logistic regression: dangling participle

<table>
<thead>
<tr>
<th>Included</th>
<th>B(SE)</th>
<th>Lower</th>
<th>exp b</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nativeness</td>
<td>-0.15 (0.69)</td>
<td>0.22</td>
<td>0.86</td>
<td>3.36</td>
</tr>
<tr>
<td>Gender</td>
<td>0.10 (0.40)</td>
<td>0.51</td>
<td>1.10</td>
<td>2.38</td>
</tr>
<tr>
<td>Age (18–25)</td>
<td>0.39 (0.60)</td>
<td>0.45</td>
<td>1.47</td>
<td>4.79</td>
</tr>
<tr>
<td>Age (26–30)</td>
<td>0.00 (0.61)</td>
<td>0.30</td>
<td>1.00</td>
<td>3.34</td>
</tr>
<tr>
<td>Age (31–40)</td>
<td>0.82 (0.55)</td>
<td>0.76</td>
<td>2.26</td>
<td>6.69</td>
</tr>
<tr>
<td>Age (41–60)</td>
<td>0.05 (0.60)</td>
<td>0.32</td>
<td>1.05</td>
<td>3.39</td>
</tr>
<tr>
<td>Level of education</td>
<td>0.18 (0.59)</td>
<td>0.38</td>
<td>1.20</td>
<td>3.78</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.32 (0.43)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note \(R^2 = .02\) (Hosmer & Lemeshow), Model \(\chi^2 (7) = 3.03\; p = .882, p < .05^*\)
Despite being a good fit model ($\chi^2(7) = 3.03, p = .882$), none of the predictors seem to make a significant contribution to it, which confirms the findings of the non-parametric tests applied above.

Figure 7.13 below shows the distribution of the contextual preference of the stimulus sentence including the dangling participle. This distribution is of interest since there seems to be an even divide between those who find the stimulus sentence acceptable and those who do not. A clear preference for informal contexts can be distinguished, with the informal speaking context receiving the highest acceptability rating of 42 per cent. The lowest acceptability rating can be identified in the formal speaking context which only obtained an acceptability rating of 10.7 per cent. Both CMC contexts fall neatly in-between the traditional speaking and writing language modes. With an average acceptability rating of 25.2 per cent, the dangling participle seems to have gained in acceptability when compared to Mittins et al.’s findings.
Regarding the level of certainty, a significant difference between acceptability groups and degree of certainty could be determined ($U = 1044$, $p = .001$, $r = -.28$). Those who deemed the stimulus sentence acceptable claimed to be somewhat certain of their assessment ($Mdn = “somewhat certain”)”, while those who found the stimulus sentence unacceptable expressed a higher degree of certainty ($Mdn = “absolutely certain”)”. This difference in degree of certainty is also reflected in the respondents’ judgment basis. While those who found the dangling participle acceptable tend to base their decision on a feeling, the respondents who indicated that the stimulus sentence is unacceptable seem to base their decisions on the self-reported knowledge of a rule ($r_s = .327$, $p = .000$, Fisher’s exact test $p = .001$). These findings hint at deviating degrees of awareness towards the dangling participle as a usage problem.

49 respondents provided a comment on the stimulus sentence with which I tested attitudes towards the dangling participle. One striking theme emerging from the analysis of comments is that of respondents offering a correction or identifying the problem at hand. Comments (72) – (74) below illustrate this particular theme.

(72)  Gun is the subject and cannot pull a trigger...
      (Retired Primary Headteacher, over 60 years old, female)

(73)  So, how does a gun pull its own trigger?! Confused subject of sentence.
      (Stay-at-home mother, 31–40 years old, female)

(74)  Clever gun.
      (Local authority, over 60 years old, female)

The three comments highlight the respondents’ awareness of the mismatch of subjects in the stimulus sentence. The gun’s ability to pull its own trigger is called into question by all three respondents. That such dangling participles
can cause confusion is a further theme which emerged from the respondents’ comments and is exemplified in (75) – (77).

(75) Sounds a bit dodgy.
(University lecturer, 41–60 years old, male)

(76) Unclear but possible.
(University lecturer, 41–60 years old, male)

(77) it is not clear who pulled the trigger - it would need to be in context and even then it would not be a proper sentence.
(Retired teacher, over 60 years old, female)

While the respondent in (75) describes the stimulus sentence as “a bit dodgy”, the respondent in (76) argues that the stimulus sentence could be “possible”, despite being “unclear”. These two comments illustrate that the stimulus sentence was perceived as flawed by some respondents. The respondent in (77) explains the ambiguity caused by the dangling participle with regard to who is responsible for the action of pulling the trigger, and goes on to notice how contextual information may clarify the confusion. Yet she states that context would “even then” not make the stimulus sentence “a proper sentence”. The role of context clearly seems to be important in that it can help clarify who pulled the trigger. Various respondents commented on this phenomenon, such as in comments (78) – (80).

(78) The gun didn’t pull the trigger. However in informal situations it depends on the utterance before/context, and if it is clear who pulled the trigger, then this would be communicatively effective.
(Freelance English Language, over 60 years old, female)

(79) I clicked unacceptable because of the lack of context - maybe there’s a situation this would make sense?
(Student, 18–25 years old, female)
(80) In informal speech native speakers could say this, but I’ve marked it ‘unacceptable’ because you’d need a very specific good context to know what they mean.

(University lecturer in German, 41–60 years old, female)

These comments prove the importance of the stimulus sentence’s context. While the respondent in (78) explains how context could make the stimulus sentence “communicatively effective” in an informal situation, both respondents in (79) and (80) explain how the lack of context influenced their decision to find the stimulus sentence unacceptable.

The HUGE analysis showed that the dangling participle was discussed in 27 of the 39 British publications, which form the basis of the ensuing discussion. The examples below illustrate the tripartite categorisation made to distinguish the treatment of the dangling participle in the HUGE database. It has to be mentioned that none of the 27 usage guides advocated the use of dangling participle.

**Criticised:**

The participle should normally have a proper ‘subject of reference’. C. T. Onions said that ‘a sentence like the following is incorrect because the word to which the participle refers grammatically is not that with which it is meant to be connected in sense’:

*Born in 1850, a part of his education was received at Eton.*

*(Correctly: Born in 1850, he received part of his education at Eton.)* (Burchfield, 1981, pp. 28–29)

**Neutral:**

Good English suggests that the laxness of older writers in their treatment of the participle (adjective) phrase was by no means absent from modern writers as long as the ‘loose’ phrase did not render the meaning ludicrous or nonsensical. To judge from the following examples, the modern journalist cares no more than Addison for the logical relationship, based upon position in the sentence, of noun and qualifier (that is, participle phrase). It is, indeed, possible that English today is swinging back to a freer syntax. (Vallins, 1953, p. 55)
A detailed overview of the categorisation is provided in Table 7.26 below from which the stark difference in number between the two categories becomes apparent. Only three usage guides provide a neutral treatment of the dangling participles, while 24 deem dangling participles unacceptable. This overall critical attitude towards the use of dangling participles is also illustrated in the diachronic overview of this feature’s treatment in usage guides provided in Figure 7.14 below.

Table 7.26 Treatment of the dangling participle (“criticised” and “neutral” only) in British publications

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>neutral</td>
<td>Vallins 1953(1960), Peters 2004</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
</tr>
</tbody>
</table>

Figure 7.14 Diachronic treatment of dangling participles in British publications
This syntactic issue was first critically discussed in Fowler and Fowler’s *The King’s English*, published in 1906. Therefore, the dangling participle is undoubtedly a usage problem of the twentieth century (cf. Tieken-Boon van Ostade & Ebner, 2017). The most lenient usage advice on this usage feature can be found in Peters’ *The Cambridge Guide to English Usage* (2004) from which an excerpt is quoted below.

Castigation of “dangling” constructions almost always focuses on sentences taken out of context. In their proper context of discourse, there may be no problem. … The third example [Now damaged in the stern, the captain ordered the ship back to port.] would sound natural enough in the context of narrative:

The bows of the vessel had been scarred by pack ice.

Now damaged in the stern, the captain ordered the ship back to port...

The narrative keeps the ship in the spotlight - in the topic position in both sentences ... If we rewrite the sentences to eliminate the dangling participles we lose the topicalizing effect they have. Any sentence in which they create a bizarre distraction should of course be recast. But if the phrase works in the context of discourse and draws no attention to itself, there’s no reason to treat it like a cancer in need of excision. (Peters, 2004, p. 138)

Peters emphasises how context can be decisive in whether a dangling participle causes ambiguity or not, by providing an example from literature and drawing attention to the issue of topicalization as a literary device. Correcting dangling participles would entail a loss of focus, as Peters (2004, p. 138) aptly points out.

With a traditional focus on written language, studies on dangling participles have often neglected the role played by context and have often merely focused on their occurrence in corpora (cf. Hayase, 2011). As is often argued, the syntactic mismatch of subjects can cause ambiguity and confusion, yet this ambiguity and confusion is undoubtedly consolidated in cases where no context is provided. Having made use of a stimulus sentence without any accompanying context, I aimed at the respondents’ reaction towards this particular use. As the results show, a clear divide between respondents can be identified in that half the respondents found the stimulus sentence acceptable, while the
other half found it acceptable only in various contexts. The dangling participle in the stimulus sentence obtained the highest acceptability rate of 42 per cent in the informal spoken context and the lowest rating of 10.7 per cent in the formal speaking context. A fairly similar stimulus sentence used by Mittins and his colleagues produced an average acceptability rating of 17 per cent in the late 1960s. Taking this finding into account, a tendency towards an increased acceptability of the dangling participle can be identified, as the average acceptability rating of 25.2 per cent shows, but only a very low one compared to other usage problems (Tieken-Boon van Ostade & Ebner, 2017).

My statistical analysis of possible correlations between usage judgments and the social variables investigated showed no significant results. However, respondents who discarded the stimulus sentence stated their judgment with a significantly higher level of certainty than those who found it acceptable. Furthermore, respondents identifying the stimulus sentence as unacceptable tend to base their decision on their knowledge of the rule. This is a pattern which has already been identified with other usage features discussed above (cf. Section 7.2.1.6). By adding a qualitative analysis of the obtained comments, I was able to identify not only an existing degree of awareness of the problematical status of dangling participles, but also to highlight the importance of context.

The analysis of usage entries included in HUGE emphasises an important finding, namely that the dangling participle has become a recurring usage problem in the advice literature only from the early twentieth century onwards. Furthermore, it has to be noted that usage guide authors overwhelmingly criticise this feature, even, as can be seen in Table 7.26 above, down to the three most recent usage guides included in the collection.
7.2.1.9.  

The use of *I* instead of *me*, as shown in the stimulus sentence (*S9. Between you and *I*, he will not be considered for this job*), has been described in detail in Section 6.10 above. The attitudes elicited through the online questionnaire will be analysed in this section; however, corpus data to show the feature’s actual use will be presented first. A corpus search including the exact phrase *between you and I* produced only two occurrences in the BNC. Both occurred in subsections related to spoken contexts, the spoken subsection and direct speech in the fiction subsection. The phrase *between you and me*, on the other hand, showed a higher frequency rate with 43 occurrences. The highest normalised frequency rate for *between you and me* was recorded in the fiction subsection with 2.07 tokens per million words. This subsection was followed by the magazine (0.41 tokens per million words) and spoken subsections (0.40 tokens per million words). Using the POS-tagger, I also searched for pronouns following the phrase *between you and*, which produced *between you and they* as the only other instance of nominative pronoun use in the BNC. While the frequencies for the use of nominative pronouns such as *I* or *they* is relatively low, one has to bear in mind that some of the material included in the BNC, such as edited writing and scripted speech, has most likely been proofread, as a result of which such perceived pronoun errors would have been corrected.

The sociolinguistic analysis of usage attitudes revealed that only age showed a significant difference with regard to usage judgment (*U* = 921, *p* = .000, *r* = −.36). In other words, the younger the respondents, the more likely it is that they find the stimulus sentence acceptable (*Mdn* = 31–40-years old). The median of those who found the stimulus sentence unacceptable lies at the age group comprising the ages of 41 to 60. The other social variables, i.e. education level (*r* = .107, *p* = .261), gender (*χ^2* (1) = .308, *p* = .579) and
nativeness ($r = -0.127$, $p = 0.183$), showed no statistically significant difference. Conducting a binary logistic regression analysis, I was able to show that the prediction accuracy of a model including only the constant increased from 53.6 to 67 per cent when all social variables were included. This model, which is presented in Table 7.27 below, was the result of a forced data entry method.

Table 7.27 Results of binary logistic regression: between you and I

<table>
<thead>
<tr>
<th>Included</th>
<th>95 % CI for exp $b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nativeness</td>
<td>0.59 (0.79)</td>
</tr>
<tr>
<td>Gender</td>
<td>–0.24 (0.42)</td>
</tr>
<tr>
<td>Age (18–25)</td>
<td>2.13 (0.70)*</td>
</tr>
<tr>
<td>Age (26–30)</td>
<td>1.77 (0.67)*</td>
</tr>
<tr>
<td>Age (31–40)</td>
<td>1.41 (0.58)*</td>
</tr>
<tr>
<td>Age (41–60)</td>
<td>1.16 (0.63)</td>
</tr>
<tr>
<td>Level of education</td>
<td>–0.58 (0.63)</td>
</tr>
<tr>
<td>Constant</td>
<td>–0.81 (0.47)</td>
</tr>
</tbody>
</table>

Note $R^2 = .12$ (Hosmer & Lemeshow), Model $\chi^2 (7) = 18.32$, $p = .011^*$, $p < .05^*$

While the model is of poor fit ($\chi^2 (7) = 18.32$, $p = .011$), the Wald statistics show that the predictor age does make a meaningful contribution to the model as the overall effect of age is significant at $p = .011$. In particular, the age groups 18–25 ($p = .002$), 26–30 ($p = .008$) and 31–40-year-olds ($p = .016$) display low levels of statistical significance when compared to the reference group of over 60-year-olds. As can be seen in the table above, the odds ratios for these three groups indicate that older age groups are more likely to find the use of between you and I unacceptable. These findings confirm the results of the Mann-Whitney $U$-test discussed above.

Similar to the dangling participle discussed in section 7.2.1.8, almost half the respondents considered this usage feature unacceptable. A difference between formal and informal contexts can also be identified in Figure 7.15.
The highest acceptability rating of 50 per cent was found in the informal speaking context, while the context with the lowest acceptability rate of 15.2 per cent was the formal writing context. With an average acceptability rate of 27.1 per cent, *between you and I* as a usage problem ranks relatively low in comparison to other usage problems discussed above, such as the double negative (cf. §. 7.2.1.7).

Regarding the degree of certainty, a statistically significant difference could be identified ($U = 1105$, $p = .002$, $r = -.30$) between the two acceptability groups, which showed that those who found the stimulus sentence unacceptable exhibited a higher degree of certainty ($Mdn = “absolutely certain”$) than those who did not ($Mdn = “somewhat certain”$). This is also reflected in the respondents’ judgment basis. The moderate positive correlation ($r_s = .309$, $p = .001$, Fisher’s exact test $p = .001$) identified between acceptability and judgment basis is indicative of the respondents’ tendency to base their decision on the knowledge of a rule, if the stimulus sentence including *between you and I* was considered as unacceptable.
Regarding the qualitative analysis, the stimulus sentence produced 43 comments, most of which contained a correction of the perceived error in the sentence or showed some awareness of the usage problem included in it. While some of these corrections are straightforward and brief, such as comment (81), others allow for a more detailed insight into why these corrections were made. Examples of the latter type of comment can be found in (81) – (84) below.

(81) Should be ‘me’ not I.
(Retired accounts clerk, over 60 years old, female)

(82) Should be ‘me,’ but enough people say ‘I’ to make it acceptable.
(Retired solicitor, over 60 years old, male)

(83) I know that the I should be me, but cannot remember why.
(Retired social worker, over 60 years old, female)

(84) The I should be a me but I do not know what the rule is - just know.
(Admin manager, over 60, female)

The respondent in (82) elaborates on his correction by stating that the common use of I instead of me has made it acceptable. In contrast to such straightforward corrections, those made by the respondents in (83) and (84) are accompanied by an admission on the part of the respondent that they either do not remember or are not familiar with the rule. Besides offering corrections, some respondents distinguished between their own usage and that of others. Within this theme, the topic of hypercorrection emerged, which will be discussed in detail below. Comments (85) – (87) illustrate this theme.

(85) This seems wrong, because you and I isn’t the subject of a clause here, so I would avoid this in formal and written contexts. However, ‘you and I’ is a common hypercorrection, and I probably do use it if not carefully considering my word choice.
(Software engineer, 31–40 years old, male)
The nominative and oblique cases are there for a reason. They’re not hard to tell apart, and I judge those who misuse them or hyper-correct themselves (or, even more irritatingly, others!).

(Student, 18–25 years old, male)

I would have written “you and me”- using “I” sounds like the speaker is trying too hard to be correct!

(Student, 18–25 years old, female)

In comment (85), a respondent explains how he would avoid using *between you and I* included in the stimulus sentence in “formal and written contexts”, yet he also admits using this phrase in careless situations. What is interesting in his comment is his identification of the phrase as a commonly used hyper-correction. The respondent in (86) also picks up on this phrase being prone to hypercorrections in that he claims to “judge those who misuse them [nominative and oblique cases] or hypercorrect themselves”. He goes on to note that he finds hypercorrecting the use of other speakers even more irritating. In comment (87), another student describes the phenomenon of hypercorrection as speakers who are “trying too hard to be correct”.

To conclude the analysis of the comments, I would like to discuss two comments which reflect a sentiment mentioned in Gowers (1954, p. 147), namely that the use of *I* in *It is I* is perceived as “pedantic”, while *It is me* is deemed unsuitable in written contexts (see § 6.10). The respondent in comment (88) not only distinguishes between written and spoken contexts, but also between different degrees of acceptability, stating that in writing she would be “unforgiving” when encountering this particular feature, while the same phrase in spoken contexts would be acceptable. A similar judgment is made by the respondent in (89), who finds that *between me and I* “commonly misused” although the stimulus sentence does not “sound immediately unusual or incorrect”.

(88) In writing, I don’t see why you would make this mistake as it’s a simple rule, so I would be unforgiving. In speech I think off hand you could easily make the mistake so it’s fine.
(English teacher, 18–25 years old, female)

(89) In formal writing, you should use the grammatically correct alternative. *Between you and me*. However, this is commonly misused and the above doesn’t sound immediately unusual or incorrect.
(Accountant, 26–30 years old, female)

These comments not only suggest a contextual difference between the two variants, which has already been identified in the contextual preference analysis above, but also show how customary usage influences speakers’ perceptions.

Out of the 39 British usage guides in the HUGE database, 32 discuss “I for me” as a usage problem. This particular usage feature was first critically discussed in the early nineteenth century in *The Vulgarities of Speech Corrected* (1829). The three most recent usage guides published in 2010 also contain criticism regarding the “I for me” issue. Since this category comprises different variants of the same type of usage problem, my analysis of usage entries focussed on *between you and I* in particular, i.e. pronouns following a preposition and being connected with *and*. Four usage guides, i.e. *Swan* (1980), *The Oxford Dictionary for Writers and Editors* (1981), *Blamires* (1994) and *Sayce* (2006), did not discuss this construction, but instead focussed on other “I for me” issues such as the *It is me/I* example discussed in Section 6.10. A detailed overview of the usage entry categorisation is provided in Table 7.28, while examples are provided of each of the categories used in my classification (only “criticised” and “neutral”). The classification brought to light that none of the usage guide authors advocated the use of this particular usage feature.
Yet there are educated people who lapse happily into the false grammar of "between you and I, "for you and I, ... So upmarket is this error, that *between you and I* has been called the "nob’s phrase".

This has been going on for centuries: Shakespeare wrote ‘All debts are cleared between you and I’; and for Pepys (1633-1703), kind words were exchanged ‘between my poor wife and I’. In our own time many distinguished interviewers and newscasters have given us ‘from you and I’, ‘from she and I’. No wonder that Robert Burchfield, a former chief editor of Oxford English Dictionaries, regrets that the ‘nob’s’ ungrammatical *between you and I* is ‘racing away into general, even educated use’. Nevertheless, ‘between you and I, ‘from you and I’, ‘for you and I’, etc. should be avoided, even in speech, because for many of us they are ungrammatical and slipshod. ....

These recommendations are subject to review at the end of the century, if not sooner, by which time between you and I, etc. may have become accepted usage, whatever we think, We cannot dispute this has long been the *Queen’s English*, since in 1954 Her Majesty pronounced on her return from a Commonwealth tour, ‘This is a wonderful moment for my husband and I’. (Howard, 1993, p. 209)

Neutral

(d) after *between* and *but* ( = except): ‘Between you and I’ has become a stock phrase—again in the belief that ‘you and I’ is more grammatical (and more genteel) than ‘you and me’. *Between* governs both pronouns: therefore both are accusative. In the same way *but* is often followed by the nominative. The following examples, all from Shakespeare, will illustrate the faults referred to:

‘All debts are cleared between you and I.’
   ‘There is none but he
   Whose being I do fear.’

‘I never saw a woman
But only Sycorax my dam and she.’

MEU, however, from common usage both past and present, justifies and even prefers the conjunctive use of *but* in this construction: in other words, ‘Whence all but he had fled’ is correct Mod.E. idiom. (Treble & Vallins, 1936, p. 41)
The advice offered by Treble and Vallins (1936, p. 41) was among the most lenient, as it contains an insight into the frequency of *between you and I* which they label as a stock phrase. Its idiomatic nature was also mentioned by Partridge (1942, p. 54), who went on to describe the increased use of this construction by educated speakers as well. By contrast, Howard (1993, p. 209) prescribes the use of the objective pronoun *me* in the phrase investigated. By referring to *between you and I* as a “nob’s phrase”, Howard (1993, p. 209) likewise pointed to its spreading use among educated speakers, news spokespeople and even Queen Elizabeth II.

From the overview in Table 7.28 below the mainly negative treatment of *between you and I* becomes evident. While eighteen usage guides proscribe the use of nominative pronouns in phrases such as *between you and I*, ten usage guides offer slightly more lenient usage advice by providing contextual restrictions and recommendations. An example of such treatment is the above-mentioned advice by Howard (1993, p. 209). In Figure 7.16 below, an illustration of the diachronic treatment of *between you and I* is presented.

### Table 7.28 Treatment of *between you and I* (“criticised” and “neutral” only) in British publications

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Examples</th>
</tr>
</thead>
</table>

Total: 32
The graph in Figure 7.16 below shows how *between you and I* came under heavy criticism only from the 1980s onwards, which is a trend that has crystallised in the discussion of usage problems so far.

![Figure 7.16 Diachronic treatment of between you and I in British publications](image)

The iconic phrase *between you and I* has been shown to be an interesting usage problem due to it apparently being the product of hypercorrection. This phenomenon has also been pointed out by some of the questionnaire respondents, who mistakenly describe the use of *between you and I* as an effort to sound more educated. However, the comments also reveal that such hypercorrections may have the opposite effect as users of such constructions are described as uneducated.

The statistical analysis which I undertook showed that age does indeed play a role when it comes to usage judgments on the stimulus sentence containing *between you and I* in that older respondents tend to dismiss the phrase as unacceptable, while younger ones tend to find it acceptable. Interestingly, an almost similar divide between respondents and their judgments can be identified here, as has also been identified for the dangling participle. Nearly
47 per cent of all usage judgments fall into the category unacceptable, while a clear tendency was found towards accepting the stimulus sentence in informal contexts. Comparing its average acceptability rate of 27.1 per cent to Mittins et al.’s study, in which the same construction obtained an acceptability rating of 23 per cent, shows the constructions continued disputed status more than four decades later.

Despite being mentioned first in the early nineteenth century, *between you and I* can clearly be called as a usage problem of the twentieth century. Although its treatment is predominately critical, an increase of criticism can be identified from the 1980s onwards. This pattern has been identified with other investigated usage problems as well, such as the double negative discussed above.

### 7.2.1.10. Split infinitive

In my description of the split infinitive in Section 6.11 above, I attempted to illustrate this particular feature’s status as a prototypical usage problem. Attitudes towards this phenomenon have been elicited by making use of the following stimulus sentence in the online questionnaire: S10. *He refused to even think about it*. Before presenting the sociolinguistic analysis, I will provide an overview of the occurrence of split infinitives in the BNC. Searching for the construction *to even think about* in the BNC showed nine instances, while the prescribed variant *to think even about* was not recorded in the BNC and *even to think about*, another more acceptable alternative, occurred with a raw frequency rate of eight occurrences. While this construction showed the highest normalised frequency rate of 0.25 tokens per million words in the fiction subsection of the BNC, the variant *to even think about* occurred most frequently in the magazine subsection, showing a comparable normalised frequency rate of 0.28 tokens per million words. In order to keep the corpus
search feasible, I restricted the search to adverbs directly preceding or following the infinitive.

Using the POS-tagger, I conducted a corpus search for split infinitive constructions with one adverb being placed between *to* and an infinitive. This search showed that the most frequently recorded split infinitive in the BNC is *to actually get* (36 instances), followed by *to really get* (27), *to actually do* (23) and *to even think* (21). A detailed overview of these four constructions with the highest frequencies and their prescribed alternatives is shown in Table 7.29. I chose to report only on the four most frequent constructions found in the BNC.

Table 7.29 Top 4 split infinitives and their prescribed variants (raw figures in parenthesis, BNC subsection with highest normalised frequencies in bold)

<table>
<thead>
<tr>
<th>disputed variant</th>
<th>prescribed variant post-infinitive position</th>
<th>prescribed variant pre-infinitive position</th>
</tr>
</thead>
<tbody>
<tr>
<td>to actually get (36)</td>
<td>to get actually (2)</td>
<td>actually to get (4)</td>
</tr>
<tr>
<td>spoken (3.31)</td>
<td>spoken (1.20)</td>
<td>spoken (0.30)</td>
</tr>
<tr>
<td>to really get (27)</td>
<td>to get really (35)</td>
<td>really to get (15)</td>
</tr>
<tr>
<td>spoken (1.20)</td>
<td>spoken (1.41)</td>
<td>spoken (1.20)</td>
</tr>
<tr>
<td>to actually do (23)</td>
<td>to do actually (2)</td>
<td>actually to do (6)</td>
</tr>
<tr>
<td>spoken (1.41)</td>
<td>spoken (0.20)</td>
<td>spoken (0.40)</td>
</tr>
<tr>
<td>to even think (21)</td>
<td>to think even (1)</td>
<td>even to think (31)</td>
</tr>
<tr>
<td>fiction (0.82)</td>
<td>miscellaneous (0.05)</td>
<td>fiction (1.19)</td>
</tr>
</tbody>
</table>

Table 7.29 above shows that the four most frequent split infinitive constructions and their prescribed variants either with the adverb in immediate post- or pre-infinitive position occur most frequently in spoken contexts. While the variants *to actually get* and *to actually do* show higher frequencies than their prescribed variants, *to get really* and *even to think* have higher frequency rates in the corpus than their disputed alternatives. Both *to get really*
and *even to think* could be more frequent due to speakers adding emphasis to the verb. That the placement of the adverb could cause a change in meaning, as in *Even he refused to think about it*, also needs to be borne in mind here. All three variants of the split infinitive construction *to even think* occurred most frequently in contexts other than the spoken subsection of the BNC.

Analysing possible correlations between usage judgment and the social factors investigated produced the following results. None of the social factors analysed showed a significantly different correlation with usage judgments (age ($U = 325$, $p = .052$, $r = -.18$), education level ($r_s = -.123$, $p = .196$), gender ($\chi^2(1) = .521$, $p = .470$), nativeness ($r_s = -.103$, $p = .278$)). The results of the binary logistic regression analysis were inconclusive, which is most likely due to the low number of judgments falling into the unacceptability category, as can be seen below. The use of the split infinitive was rejected by only ten informants who were all native speakers and university-educated. Furthermore, seven of these ten informants were over 60 years old. Hence, the group of informants rejecting the split infinitive is most likely too homogeneous for this kind of analysis and a larger sample would be needed to test the influence of all social variables on the outcome variable in a reliable manner.

The contextual preference analysis indicates the stimulus sentence’s widely acceptable status. With an average acceptability rating of 63.5 per cent, the split infinitive in the stimulus sentence shows the highest acceptability rating of 84.8 per cent in the informal speaking context and the lowest acceptability rating of 41.1 per cent in the formal writing context. Only 8.9 per cent of judgments fell into the unacceptable category. The contextual preference of the stimulus sentence investigating attitudes towards split infinitives is illustrated in Figure 7.17 below.
While no statistically significant difference between informants who find the stimulus sentence acceptable and those who do not could be identified ($U = 477, p = .697, r = -.04$), the analysis of the informants’ judgment basis showed that respondents who rated the stimulus sentence unacceptable tend to base their judgment on their familiarity with the rule. This also means that an acceptable judgment correlated positively with gut feeling ($r_s = .287, p = .003$, Fisher’s exact test $p = .007$).

42 respondents commented on the stimulus sentence (*He refused to even think about it*). The most persistently recurring themes identified in these comments included the notorious status of split infinitives as a language myth. The respondents not only named the issue at hand, but also provided an elaboration on what they think about this particular usage feature. Comments (90) – (94) represent this theme.

(90) I think the *even* should go before the *to*? But I’m not sure.
(English teacher, 31 – 40, Male)
(91) Don’t really mind a split infinitive.
   (Teacher, 41–60 years old, female)

(92) Good enough for Star Trek, good enough for anyone - to proudly split any infinitives.
   (Teacher, 41–60 years old, male)

(93) split infinitive!!!!!! however, I do know that this has changed over recent years - demonstrated by BBC.
   (Retired primary and EFL teacher, over 60 years old, female)

(94) The infamous split infinitive............ another battle that has been lost.
   (Retired language teacher, over 60 years old, female)

The comments quoted above not only reflect changes in the acceptability of split infinitives, they also provide a specialised insight into the attitudes of teachers. While the English teacher in (90) does indeed correct the split infinitive by moving even before to, he at the same time questions his decision and states his uncertainty about his correction. This comment needs to be viewed in contrast to the ones made by the respondents in (91) and (92), who are both teachers as well. In (91) a simple straightforward judgment of the acceptability of split infinitives is made. A similar notion is expressed by the teacher in (92) who, however, elaborates her judgment by referring to the famous opening of Star Trek and states “to proudly split any infinitives” herself. The last comments quoted above are made by two female retired teachers who express a slightly more negative perception of the changes affecting the acceptability of split infinitives. Following an outcry at identifying the split infinitive, the respondent in (93) goes on to weaken her initially negative reaction by stating that she has noticed a change in usage. In particular, she refers to the usage of the BBC, whose role of a language guardian has already been discussed in Chapter 2. This confirms not only how the language use of the BBC is subject to criticism from its audience, but also how the media reflects the language
use of society (see Bell, 1995, p. 23). Lastly, the respondent in (94) describes the “infamous split infinitive” as a lost battle.

Comments (95) – (97) deal with the theme of obtaining possibly negative judgments by other speakers triggered by the use of a split infinitive.

(95) Split infinitive. But I think these are fine! Others would disagree with me. I think it is absurd to base a rule on something that was formerly based on Latin, a completely different language family. But many would see this as an error...

(Lecturer, 31–40 years old, female)

(96) I wouldn't, personally, use a split infinitive in formal writing in case I was judged by someone even more pedantic than me. I think it sounds fine, though.

(English teacher, 41–60 years old, female)

(97) This is like the ‘fewer’ example: I would avoid this in writing, but I find the rule pedantic and pointless. However, sometimes one has to respect the prejudices of others, and I would do so in this case.

(University lecturer in German, 41–60 years old, female)

All three comments quoted here express the respondents’ tendencies to avoid splitting infinitives, since they realise that other speakers still perceive them as incorrect. The respondent in (95) emphasises the absurdity of the Latin origin of the split infinitive in English, yet concludes that split infinitives are still perceived as errors. In (96), a respondent claims to avoid split infinitives not to be “judged by someone even more pedantic than” her but continues by stating that the stimulus sentence “sounds fine”. The influence of those who regard split infinitives as errors is illustrated in (97), in which a respondent argues that he respects “the prejudices of others” and consequently avoids split infinitives, despite stating how “pedantic and pointless” this rule is. These comments provide an insight into how the influence of prescriptivists can
affect actual usage in that even those who accept split infinitives avoid such linguistic constructions in order not to commit a mistake in the eyes of others.

The last two comments I would like to discuss here deal with the theme of finding a historical reason for the split infinitive rule.

(98) The issue of split infinitives was raised to distinguish classes, not to identify ‘proper’ English.
(University lecturer, 41–60 years old, male)

(99) There isn’t really a sound or historically valid grammatical reason to ban split infinitives. In this sentence to move ‘even’ would detract from the force and meaning of the sentence. I think the ‘never split an infinitive’ argument is a bit feeble.
(Stay-at-home mother, 31–40 years old, female)

While the university lecturer in (98) argues that the rule against split infinitives was established “to distinguish classes, [and] not to identify ‘proper’ English”, the stay-at-home mother in (99) claims that the “feeble” rule against split infinitives has no “sound or historically valid grammatical reason”. These two comments complement the comments quoted above in that respondents acknowledge the outdated character of the rule; for all that, the influence of prescriptivists is retained and the myth about the split infinitive continues to be alive within the speech community.

In order to illustrate the history of this particular usage problem, I conducted a survey of all British usage guides included in HUGE. Being an old chestnut, the split infinitive is discussed in 34 British publications. The tripartite categorisation of the usage entries into “criticised”, “neutral” and “advocated” is illustrated in the examples below. The overall results of this categorisation are illustrated in Table 7.30 below. As can be seen from this overview, the majority of usage guide authors assume a neutral position. Figure 7.18 contains an overview of the diachronic development of the split infinitive discussion in British usage guides.
Criticised: A correspondent states as his own usage, and defends, the insertion of an adverb between the sign of the infinitive mood and the verb. He gives as an instance, “to scientifically illustrate.” But surely this is a practice entirely unknown to English speakers and writers. It seems to me, that we ever regard the to of the infinitive as inseparable from its verb. And when we have a choice between two forms of expression, “scientifically to illustrate,” and “to illustrate scientifically,” there seems no good reason for flying in the face of common usage. (Alford, 1864, p. 171)

Neutral: Avoid the split infinitive wherever possible; but if it is the clearest and/or most natural construction, use it boldly. The angels are on our side. (Partridge, 1942, p. 298)

Advocated: If you think a sentence will be more emphatic, clear or rhythmical, split your infinitive – there is no reason in logic or grammar for avoiding it. … If you can’t bring yourself to split an infinitive, at least allow others the freedom to do so. (Cutts, 1995, pp. 96–97).

Table 7.30 Treatment of the split infinitive (“criticised”, “neutral” and “advocated”) in British publications

<table>
<thead>
<tr>
<th>Type</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>criticised</td>
<td>Alford1864, Fowler&amp;Fowler1906(1922), Treble&amp;Vallins1936, Heffer2010</td>
</tr>
<tr>
<td>neutral</td>
<td>(23)</td>
</tr>
<tr>
<td>Total:</td>
<td>34</td>
</tr>
</tbody>
</table>

This usage feature was first discussed and criticised by Alford in 1864 and was subject to a rather negative treatment until the 1940s. Figure 7.18 shows, however, how the traditional stricture on splitting infinitives made way for a more moderate attitude towards this usage feature. From the 1980s onwards,
usage advice on the split infinitive issue shifted yet again and the practice of splitting infinitives even came to be advocated.

An exception to this gradual move toward acceptability is Simon Heffer’s *Strictly English* published in 2010. Not only does Heffer condemn split infinitives, he also wrongly attributes the origin of the rule against splitting infinitives to Lowth’s grammar.

This began with Latinists, notably Lowth, arguing that since the infinitive was intact in that language, it had better be as intact as possible in our own too. There is no reason in that sense why this should apply in English. However, the division of *to* from its verb was seized on by the Fowlers, correctly in my view, as inelegant. … for the sake of logic and clarity to and the verb whose infinitive it forms are always best placed next to each other rather than interrupted by an adverb. In nearly 30 years as a professional writer I have yet to find a context in which the splitting of an infinitive is necessary in order to avoid ambiguity or some other obstruction to proper sense. (Heffer, 2010, p. 62)

What also becomes apparent from this overview is that the split infinitive is a usage problem of the twentieth century, as it was discussed in only one earlier usage guide (cf. Tieken-Boon van Ostade & Ebner, 2017). While some usage
guide authors started advocating the splitting of infinitives in the 1980s, the myth of the split infinitive seems to have persisted in the usage guide tradition.

The myth surrounding split infinitives has made this particular usage feature not only a recurring old chestnut in the usage debate, but also turned it into a prototypical usage problem which enjoys widespread notoriety among speakers. The corpus evidence does not only show that split infinitives are in actual use, but also highlights their frequent occurrence in spoken contexts. This finding was illustrated on the basis of the four most frequent split infinitive constructions found in the BNC. As Crystal (2006a, p. 126) argues, splitting an infinitive is said to follow “the heartbeat of English”, by which it is said to make the construction sound more natural. That placing an adverb between the infinitive marker to and the infinitive can add emphasis or can cause a change in meaning further needs to be borne in mind (see § 6.11).

My sociolinguistic analysis showed that age again plays a crucial role in that older respondents are more likely to reject the split infinitive than younger ones. This generational difference in reflecting linguistic preferences in this respect could possibly result in split infinitives becoming irrelevant. With an average acceptability rating of 63.5 per cent, the stimulus sentence including the split infinitive has obtained the highest average acceptability rating of the usage problems investigated so far. In comparison with Mittins et al.’s study, which produced a 40 per cent average acceptability rating for a very similar stimulus sentence, a clear tendency of increased acceptability can therefore be identified. This growing acceptability of split infinitives is also reflected in the analysis of usage entries included in HUGE, which showed a clear move towards a more lenient treatment of split infinitives, with some writers even advocating its use. It should be mentioned that resistance towards the construction’s acceptability is still provided by one very late usage guide author, i.e. Simon Heffer, which stresses the fact that the selection of usage
problems to be included in a usage guide is subject to the usage guide author’s preferences. Although the split infinitive seems to enjoy an overwhelmingly high acceptability rate, its mythological status is still a fact. Despite mentioning the outdated character of the rule against splitting infinitives, respondents claim to follow the rule in order not to be judged negatively by others.

7.2.1.11. **Literally as an intensifier**

The changing meaning of *literally* as an intensifier has been described in Section 6.12 above, in which I have attempted to provide a more detailed insight into this particular usage feature’s stigmatisation. By including the stimulus sentence (S11. *His eyes were literally popping out of his head*) in the online questionnaire, attitudes towards the non-literal, hyperbolic use of *literally* were elicited. Before presenting the results of the sociolinguistic analysis which I conducted, I will report on the findings of two corpus-based studies of *literally* which made use of qualitative methods. Nerlich and Chamizo Domínguez (2003) conducted an analysis of occurrences of *literally* found in the *Bank of English*, a corpus of modern English varieties comprising 450 million words (Lee, 2010, p. 109). Grouping instances of *literally* into three categories depending on their meaning, Nerlich and Chamizo Domínguez (2003, pp. 202–204) show the different uses of *literally*. While the first group comprises instances in which *literally* is used in its original literal sense, the second group contains occurrences of *literally* which show the gradually shifting meaning of *literally* which is often used for rhetorical purposes (Nerlich & Chamizo Domínguez, 2003). The third and last group includes occurrences of *literally* which show a completed shift of meaning, i.e. *literally* meaning “the opposite of what it meant before” (Nerlich & Chamizo Domínguez, 2003, p. 203). While the original meaning of *literally* is illustrated in an example taken from the *OED* (s.v., *literally*) in 8.a), Nerlich
and Chamizo Domínguez’s study (2003, p. 203) provides examples of the intermediate and completely changed meanings of literally in 8.b) and 8.c) respectively.

8. a) She often became very angry with me for taking her literally.

b) Pamper yourself quite literally from head to toe!

c) I literally died laughing.

A similar categorisation as well as a corpus analysis was conducted by Claridge (2011), who made use of the BNC and the Santa Barbara Corpus of Spoken American English for this, adding a comparative dimension between British and American English to her study. Claridge (2011, pp. 100–101) distinguished between conventional uses of literally, semi-creative/conventional occurrences, and creative hyperbolic uses of literally. This tripartite categorisation corresponds with Nerlich and Chamizo Domínguez’s categorisation described above. Having compiled a sub-corpus containing occurrences of literally based on the BNC data, Claridge reports on the distribution of the 376 instances of literally. While the majority of the instances from this BNC sub-corpus, namely 85 per cent, fall into the first category comprising conventional uses of literally, only 8 and 7 per cent fall into the semi-creative/conventional and creative expression groups respectively. Thus, her findings serve as an indication of how frequent the different uses of literally actually are.

My sociolinguistic analysis of the usage attitudes obtained through the online questionnaire produced the following results. While education level ($r = .099, p = .297$), gender ($\chi^2(1) = .014, p = .905$), and nativeness ($r_s = −.009, p = .921$) did not show any statistically significant differences according to the acceptability ratings, usage attitudes, once again, seemed to vary according to age ($U = 982, p = .048, r = −.19$). These findings show that younger respondents exhibit the tendency to consider the stimulus sentence including
literally acceptable ($Mdn = 31–40\text{-years-old}$), while older respondents tend to reject the stimulus sentence ($Mdn = 41–60\text{-years-old}$).

The binary logistic regression analysis showed that a model containing only the constant has an overall prediction accuracy of 71.4 per cent which increases to 72.3 per cent in the proposed model presented in Table 7.31 below containing all social variables as predictors. Despite being a good fit ($\chi^2(7) = 6.20, p = .516$), the model does not show any statistical significance with regard to the Wald statistics. While age showed a significant correlation with acceptability ratings in the Mann-Whitney $U$-test, the binary logistic regression analysis indicates that the significance of age is diminished when all social variables are taken into account ($p = .368$).

Table 7.31 Results of binary logistic regression: *literally*

<table>
<thead>
<tr>
<th>Included</th>
<th>$B(SE)$</th>
<th>Lower</th>
<th>exp $b$</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nativeness</td>
<td>0.93 (0.79)</td>
<td>0.23</td>
<td>1.10</td>
<td>5.17</td>
</tr>
<tr>
<td>Gender</td>
<td>-.12 (0.45)</td>
<td>0.37</td>
<td>0.89</td>
<td>2.13</td>
</tr>
<tr>
<td>Age (18–25)</td>
<td>1.49 (0.84)</td>
<td>0.86</td>
<td>4.42</td>
<td>22.77</td>
</tr>
<tr>
<td>Age (26–30)</td>
<td>0.95 (0.74)</td>
<td>0.61</td>
<td>2.58</td>
<td>10.89</td>
</tr>
<tr>
<td>Age (31–40)</td>
<td>0.16 (0.57)</td>
<td>0.38</td>
<td>1.17</td>
<td>3.59</td>
</tr>
<tr>
<td>Age (41–60)</td>
<td>0.36 (0.65)</td>
<td>0.40</td>
<td>1.43</td>
<td>5.11</td>
</tr>
<tr>
<td>Level of education</td>
<td>-0.56 (0.62)</td>
<td>0.17</td>
<td>0.57</td>
<td>1.91</td>
</tr>
<tr>
<td>Constant</td>
<td>0.66 (0.46)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note $R^2 = .05$ (Hosmer & Lemeshow), Model $\chi^2(7) = 6.20, p = .516, p < .05^*$

The stimulus sentence was identified as acceptable in informal contexts by the majority of respondents, as all show an acceptability rate of above 50 per cent. The informal speaking context shows the highest acceptability rate of 68.8 per cent. In contrast, the stimulus sentence is considered less acceptable in the three formal contexts. It obtained the lowest acceptability rate of 10.7 per cent in the most formal context, formal writing. Overall, the stimulus
sentence which I used to investigate attitudes towards the use of literally obtained an average acceptability rating of 37.1 per cent. Nearly 30 per cent of the questionnaire respondents found the same stimulus sentence unacceptable, which is illustrated in Figure 7.19 below.

A statistically significant difference in degree of certainty could be identified ($U = 624, p = .000, r = −.48$) between respondents and their usage judgments. Those respondents who found the stimulus sentence unacceptable seemed to be motivated by a higher degree of certainty ($Mdn = “absolutely certain”) than those who found the stimulus sentence acceptable ($Mdn = “somewhat certain”). Furthermore, respondents calling the sentence unacceptable tended to base their judgment on the knowledge of a rule, while those who found it acceptable reported basing their judgments on a gut feeling ($r_s = .436, p = .000$, Fisher’s exact test $p = .000$).

![Figure 7.19 Contextual acceptability in percentages: literally](image)

Analysing the 49 comments obtained, I identified the most prominent and recurring themes for the stimulus sentence which I used to test speaker
attitudes towards *literally*. While numerous respondents commented on the widespread use and acceptability of *literally*, which is illustrated in comments (100) – (102), the unacceptability of *literally* was also frequently mentioned by respondents and is illustrated in comments (103) – (105).

(100) people get annoyed by use of ‘literally’, I am guilty of saying *literally*, literally all the time, literally.
(Student, 18–25 years old, male)

(101) ‘Literally’ is used a lot in this way (i.e. not literally!) and I think this meaning is so common that one could argue it is now acceptable. Speakers who are more prescriptive (or purist) would disagree I am sure. I do not use *literally* in this way but it use is extremely widespread.
(Lecturer, 31–40 years old, female)

(102) Of course they weren’t! However, the expression is used so often that it has become acceptable in informal verbal usage.
(Retired solicitor, over 60 years old, male)

While the respondent in (100) admits to using *literally* and produces a good example of his linguistic practice, the respondent in (101) emphasises how the use of *literally* as exemplified in the stimulus sentence is not only “so common that one could argue it is now acceptable”, but it is also “extremely widespread”. A similar notion is expressed in comment (102), made by a retired solicitor who argues that the frequent use of non-literal *literally* contributed to its acceptability. Yet he goes on to restrict the word’s acceptability to “informal verbal usage”. Both comments in (101) and (102) highlight how common and widespread usage has influenced the respondents’ perceptions of the acceptability of *literally* in a non-literal sense. In contrast to such comments, a large number of respondents also mention the unacceptability of non-literal *literally* (n = 28).

(103) Wrong use of ‘literally’.
(Retired accounts clerk, over 60 years old, female)
‘Literally’ is such a misused word!
(Stay-at-home mother, 31–40 years old, female)

“literally” is another dreadful word.
(Proof-reader, 31–40 years old, male)

The respondents in comments (103) to (105) briefly yet assertively state the unacceptability of the stimulus sentence containing literally. Being described as the “wrong use of ‘literally’”, “a misused” or “dreadful word”, literally in a non-literal sense clearly proves to be a current usage problem, as was already apparent from the comments (100) – (102) quoted above.

Studies on literally such as Nerlich and Chamizo Domínguez (2003) and Claridge (2011) differentiated three different stages of literally. Besides distinguishing between the word’s original meaning and its use as a hyperbolic intensifier, literally is also found to be used in an intermediate function whose meaning depends on the interpretation of the interlocutor. That the changing status of the word has been perceived by questionnaire respondents is evident from comments (106) and (107). The comments also showed evidence of the intermediate status of literally in that a number of respondents were uncertain of the scene described in the stimulus sentence. Comments (106) – (110) illustrate this.

The Oxford dictionary now says that “literally” means “not literally”, so who am I to argue?
(Writer/journalist, 26–30 years old, male)

Literally now can mean figuratively, I think I read this in a newspaper.
(English Teacher, 31–40 years old, male)

Comments (106) and (107), moreover, show awareness of the fact that the changing status of literally has been discussed in the media. While the re-
spondent in (106) refers to the adoption of *literally* into the “Oxford Dictionary”, the respondent in (107) mentions remembering the discussion of this event in a newspaper. Both comments, however, express the notion of authority, through which the status of *literally* seems to be legitimatised.

(108) Well were they or weren’t they?
(Student, 18–25 years old, male)

(109) literally? actually? physically?
(Retired arts consultant, over 60 years old, female)

(110) Hopefully they weren’t literally popping out as I wouldn’t know how to put them back in. Very dramatic & language is informal even slangy.
(Administrator, 26–30 years old, female)

While the respondents in (108) and (109) enquire the outcome of the scenario described in the stimulus sentence, another informant in (110) questions the process of eyes popping out, describing the language as “very dramatic” and “informal even slangy”. The intermediate use of *literally* is highly dependent on the interpretation of the respondents and whether they take *literally* to mean “really” or “figuratively”. All three respondents in the comments quoted above accepted the use of *literally* in the stimulus sentence, which further indicates how *literally* is associated with either the intermediate position or the non-literal sense of *literally*.

As mentioned in Section 6.12 above, the use of *literally* in a non-literal sense was first discussed in an American usage guide in 1918, namely Strunk’s *The Elements of Style*. The first reference to this particular usage problem in a British publication followed a few years later, in Fowler’s *A Dictionary of Modern English Usage* (1926). A total of 27 British usage guides discuss this usage problem. Applying the modified version of Yáñez-Bouza’s categorisation (2015) resulted in the overview presented in Table
7.32 below. Examples of each category are provided to illustrate the usage problem’s treatment in more detail.

**Criticised**

One of the most commonly misused words. A friend recently told Her Ladyship, rather alarmingly, that her two siblings were *literally* chalk and cheese. One is a slim, dark man; the other a solidly built woman with fair hair. They also have very different personalities. Metaphorically or figuratively, therefore, they are chalk and cheese. Literally, Her Ladyship can assure her readers, they are nothing of the sort. (Taggart, 2010, p. 77)

**Neutral**

2. Comment on the use of:
   (i) *virtually* in the following sentence: “Ruskin was *virtually* burned out when he was sixty”;
   (ii) *literally* in this sentence: “He *literally* glued his ears to the ground”. …

2. (a) Correct; *virtually* is here an antonym to *actually*.

(b) If a man took a glue-pot and a brush and, by some acrobatic contortions, with the glue fastened himself to the ground by the ears, *literally* would be correctly used in this sentence. I *literally* fly to a man’s help only if I go by aeroplane from where I am to where he is. To use *literally* with a metaphor is, obviously, to confuse the literal and the metaphorical. (See page 169). But the usage is very common; *literally*, in fact, loses its own literal meaning, and becomes an intensive or emphasising adverb in a kind of hyperbole. Vallins, 1951, pp. 199, 245)

**Advocated**

This word has a split personality: plain-speaking and tantalizing. In its primary sense, *literally* urges you to take a fact “according to the letter,” i.e. word for word or exactly as the utterance has it. Yet for most of the last two centuries it has also been used to underscore figures of speech or turns of phrase which could never be taken at face value: They were literally green with envy.

In cases like that, *literally* defies its literal sense and seems to press for factual interpretation of the idiom, however far-fetched. Readers are tantalized - caught between the urge to believe and disbelief. (Peters, 2004, p. 326)
While Taggart (2010) holds on to the traditional use of *literally*, Vallins (1951) describes the widespread use of *literally* in a non-literal sense, despite providing a correction of its use. Peters (2004), once again, turns out to be the most lenient in providing usage advice in that she acknowledges and advocates the different meanings of *literally*.

Table 7.32 Treatment of *literally* (“criticised”, “neutral” and “advocated”) in British publications

<table>
<thead>
<tr>
<th>Categorisation</th>
<th>Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>advocated (1)</td>
<td>Peters2004</td>
</tr>
</tbody>
</table>

Table 7.32 demonstrates how the traditional proscription against the use of *literally* in a non-literal sense prevails in the British usage guides included in HUGE. While 20 publications criticise this alleged new use, six take a neutral stance by accepting the usage problem in some contexts, and only one usage guide makes a clear distinction between the two uses and emphasises the role of the speaker in deciding the acceptability of this feature.

Placing this categorisation in a historical light, Figure 7.20 emphasises not only how *literally* is overwhelmingly criticised throughout the history of this usage problem, but also clearly illustrates how as an intensifier it is a typical usage problem of the twentieth century. Vallins’s (1951) acknowledgement of the widespread use of *literally* and its changing meaning as illus-
treated in the examples above seems to be an exception in the early stigmatisation history of \textit{literally} as an intensifier. Providing a neutral view by allowing contextual preferences became more frequent from the 1980s onwards. Yet, Figure 7.20 below also shows how prescriptions and proscriptions against the use of \textit{literally} in a non-literal sense have become more frequent.

![Figure 7.20 Diachronic treatment of the \textit{literally} as an intensifier in British publications](image)

The use of \textit{literally} as an intensifier has been the subject of a heated public debate in Great Britain which was fuelled by the \textit{OED}'s acceptance of \textit{literally}. Although this usage feature has come to be viewed as problematical only fairly recently, its earliest recorded use dates back to the eighteenth century (see § 6.12). Interestingly, this “new” use of \textit{literally} only became a regular feature in the usage debate after it was included in Strunk’s \textit{The Elements of Style} (1918), an American usage guide, followed by its treatment in Fowler’s \textit{A Dictionary of Modern English Usage} (1926) only a few years later. The semantic shift which \textit{literally} has been undergoing was captured and discussed in two studies which identify three different functions of \textit{literally}.

- **Advocated**
- **Neutral**
- **Criticised**
(Nerlich & Chamizo Domínguez, 2003; Claridge, 2011). Nevertheless, studies on the use of literally and of attitudes towards its hyperbolic use are relatively rare. The Mittins study included literally among the 55 items studied, and it obtained an average acceptability rate of 35 per cent. Interestingly, my own questionnaire results show only a slightly increased average acceptability rating of 37.1 per cent, which serves as a clear indicator of the disputed status of literally. The analysis of the comments demonstrates not only how respondents are aware of the changing meaning of literally, but also brings to light the possible confusion of respondents who are aware of the traditional and hyperbolic meaning of literally.

As far as the sociolinguistic analysis is concerned, age showed a statistically significant main effect, which was, however, diminished in the binary logistic regression analysis which takes any potential influences of the other social variables investigated into account. According to this main effect, younger respondents were more likely to accept the use of literally in the stimulus sentence than older ones. Respondents who made a negative usage judgment did so with a higher degree of certainty and stated basing their judgment on rule knowledge, while respondents who found the stimulus sentence acceptable showed a lower degree of certainty and tended to base their judgment on a feeling. Similar tendencies have been identified with already discussed usage problems and indicate that awareness of a particular usage feature may be translated into a more assertive judgment.

### 7.3. Concluding Remarks

The aim of the first part was to elicit attitudes towards usage problems by making use of a slightly less direct elicitation technique and of stimuli sentences which were partly taken from Mittins et al.’s Attitudes to English Usage (1970) and partly created for the purpose of this study. Therefore, a tentative comparison of changing usage tendencies could be achieved by comparing
essentially two snapshots of usage attitudes, which I will return to in Chapter 10. Although the usage features investigated were not highlighted, unlike in the case of the Mittins study, the respondents’ awareness of the stigmatised features often did come to light, showing that respondents were aware of the usage problems concerned; they commented on them accordingly. While most questionnaire respondents expressed their attitudes towards the usage problems studied, some stimulus sentences contained other features which seemed to be more salient to them. This was, for example, the case with the stimulus sentence used to elicit attitudes towards the dangling participle (see § 6.2.3), which contained a semi-colon whose alleged inappropriateness unfortunately attracted the attention of several respondents, thus deflating their attention from the issue at hand. The analysis of the second part of the questionnaire will be discussed in detail in the following chapter making the online questionnaire analysis complete.