**A corpus-based analysis of adjective amplification among native speakers and learners of English**

Martin Schweinberger

The University of Queensland, Australia

**Abstract**

This paper analyses adjective amplification by native speakers and learners of English with diverse language-backgrounds based on the *International Corpus of Learner English* and the *Louvain Corpus of Native English Essays* (ICLE). The study uses *Multifactorial Prediction and Deviation Analysis Using Regression/Random Forests* (MuPDARF) to evaluate which factors contribute to NNS making native-like use of *very* as the most frequent amplifier in the corpus data. The MuPDARF provides a detailed picture of divergences between NS and NNS of English and reports that NNS are significantly more likely to use *very* in a native-like manner in primed contexts and when amplifying emotional adjectives. In addition, NNS use *very* like English NS with low frequency adjectives in predicative contexts.

1. **Introduction**

The rise of learner corpus research (LCR) in the past three decades or so has advanced reliable empirical research of language background specific difficulties among learners from diverse language backgrounds. In addition, ever more sophisticated statistical modeling allows us to analyze both systematic effects of the language background on the acquisition process and interactions between factors that condition learners’ choices (see Gries 2018; Wulff and Gries 2019a). Thereby, advanced statistical modelling enables us to arrive at a more detailed understanding of differences between NS and NNS as these methods enable researchers to explain these differences with respect to cognitive mechanisms that underly native-like speech production (see, e.g., Gries and Adelman 2014; Gries and Deshors 2014; Wulff and Gries 2019b). The present paper combines LCR with recently introduced multivariate statistical methods to unearth systematic divergencies between NS and NNS in the domain of adjective amplification and to show how modern statistical methods may help us understand which factors contribute to native-like performance.

The phenomenon the present study focuses on is adjective amplification in English (see (1) and (2)). Adjective amplification is particularly relevant from the perspective of learners of English because amplification plays a crucial role in the social and emotional expression of speakers (see Labov 1985; Partington 1993; Peters 1994; Ito and Tagliamonte 2003: 258). In addition, adjective amplifiers represent key elements in expressing stance and they allow both NS and NNS to indicate social affiliations within a given speech community. However, acquiring the intricate social and linguistic multifunctionality of amplification appropriately represents a serious challenge even for advanced learners, because amplifier systems are particularly are a site of “fevered invention” and a domain that is particularly prone to change (Bolinger 1972: 18; Quirk et al. 1985: 590).

(1) Adjective amplification in predicative contexts

a. […] because they were *very poor*. (ICLE-BG-SUN-0003.1)[[1]](#endnote-1)

b. At the same time they made drugs *really popular*. (ICLE-PO-POZ-0046.3)

c. […] it is *absolutely irresponsible* to fight against another denomination […]. (ICLE-GE-SAL-0009.3)

d. […] they will *completely disappear* in fifty years at latest. (ICLE-RU-MOS-0020.6)

e. The countries are *so different* (culture, language, tradition) […]. (ICLE-FR-UCL-0067.1)

(2) Adjective amplification in attributive contexts

a. In others words, Europe has to become a *very strong* power. (ICLE-FR-UCL-0066.1)

b. Also if they are lucky and she gets a *really suitable* husband […]. (ICLE-SW-LND-0023.8)

c. […] you will have an *absolutely excellent* view on the match. (ICLE-DN-NIJ-0007.2)

d. These two *completely antagonistic* characters have nothing to envy one another. (ICLE-CZ-PRAG-0006.1)

Adjective amplification is a subtype of intensification and is related to the semantic category of degree which is why intensifying adverbs are also been referred to as degree adverbs or adverbs of degree (Quirk et al. 1985). Intensification ranges between very low intensity (downtoning) and very high intensity (amplification) (Quirk et al. 1985, 589-590). According to Quirk et al. (1985, 589-590), amplifiers “scale upwards from an assumed norm [while] downtoners have a lowering effect, usually scaling downwards from an assumed norm” (1985, 590). The current paper focuses exclusively on adjective amplification (see (1) and (2)) while leaving aside downtoning and therefore disregrads approximators such as *almost*, compromisers such as *more or less*, diminishers such as *partly*, and minimizers such as *hardly*. Within the category of adjective amplifiers, Quirk et al. (1985, 589-590) differentiate between maximizers such as *completely*, which denote the upper extreme of the scale (Quirk et al. 1985, 590) and boosters such as *very* that denote a high degree or a high point on the scale. Boosters, in particular, form an open class, which adopts new members to replace forms which have lost their expressiveness due to frequent use (see Quirk et al. 1985, 590).

Previous research has produced a wealth of insights about adjective amplification and added to our understanding of differences between NS and NNS in this domain (e.g. Lorenz 1999, 2002; Gill 2002; Granger 1998) but which factors correlate with differences between NS and NNS from diverse language backgrounds remains unclear.

The following section surveys previous research on adjective amplification with a focus on adjective amplification among NNS of English. The third section describes the corpus data, the data processing steps, and the statistical methods. The fourth section presents the results of the analysis while the fifth section discusses the results in light of previous research and highlights potential shortcomings of the present study.

1. **Previous Research**

Adjective amplification has received a substantial amount of attention (e.g. Bolinger 1972; Breban and Davidse 2016; D’Arcy 2015; Ito and Tagliamonte 2003; Lorenz 2002; Méndez-Naya 2003, 2008; Méndez-Naya and Pahta 2010; Nevalainen 2008; Paradis 2008; Partington 1993; Peters 1992; Pertejo and Martínez 2014; Rissanen 2008; Tagliamonte 2008; Tagliamonte and Denis 2014; Tagliamonte and Roberts 2005). In addition, various studies have produced a detailed picture of differences NS and NNS of English with respect to adjective amplification (e.g. Lorenz 1999, 2002; Gill 1998; Granger 1998; Hendrikx, Van Goethem, and Wulff 2019).

One of the most consistent findings of previous research relates to overuse of *very* and *really* in academic writing produced by intermediate to advanced learners of English (see Granger 1998; Lorenz 1999: 215; Philip 2007: 7). Philip (2007), for instance, found that *very* and *really* are used to the “virtual exclusion of any other adverb” (Philip 2007: 7). Similarly, Lorenz (1999: 198) found that German learners of English overuse the all-purpose intensifiers *very* and *really* which the L1-German leaners used like in speech rather than as would be appropriate in academic written discourse (1999: 215). Hinkel (2003) has argued in a similar vein stating that the patterning of adverbs in data coming from learners of English with Chinese, Japanese, Korean, and Indonesian language backgrounds mirrored the frequencies in conversational style, thus indicating that learners used colloquial style rather than a more appropriate formal style due to a limited lexical repertoire (2003: 1058).

With respect to the overall rate of amplification, previous research provides mixed results. Lorenz (1998) found that L1-German leaners amplify significantly more frequently than NS both relatively and in terms of absolute frequencies (Lorenz 1998: 64) whereas Granger (1998) and Hendrikx, Van Goethem, and Wulff (2019) found the opposite trend in their data. Lorenz (1998: 62-63) argues that the tendency among German learners to amplify more frequently represents a misapplication due to overall writing strategies which are grounded in NNS’s desire for expressiveness. The latter point echos analyses of the expression of stance in academic writing by NNS of English that found that use of boosters (amplifiers), together with stance-taking strategies, enabled students to develop a stronger voice (Fallas Escobar and Chaves Fernández 2017: 118).

Furthermore, both Lorenz (1998, 1999) and Granger (1998) found NNS show a higher degree of variability in amplifier-adjective collocations. Among the German NNS of English studied by Lorenz (1998) based on a subsection of the ICLE, this trend is captured as lower degree of cohesion among amplifier-adjective bigrams. Granger (1998) used both corpus data as well as word combination tasks to tap into intuitional differences concerning amplifier-adjective collocations between NS and French NNS of English. The word combination task indicated that learners showed significantly more variation in what they deemed acceptable pairs compared with NS, which, according to Granger (1998: 152), suggests that the sense of collocational salience was missing or misguided among (French) learners of English.

With respect to differences between NS and NNS based on NNS’ proficiency, several studies found that differences waned as NNS became more proficient. Results provided by Edmonds and Gudmestad (2014), who analyse intensifying adverbials in academic essays, indicated that advanced NNS of English preformed like English NS while less advanced NNS differed substantially from both advanced NNS and NS. This trend is corroborated by Forsberg (2010) who focused on the use of formulaic sequences among NNS of French and also found that differences between NNS and NS waned off as NNS became more proficient. Similarly, Hendrikx, Van Goethem, and Wulff (2019: 80-81) found that while NNS amplified significantly less than NS, more advanced NNS approximate the rate of amplification exhibited by NS.

Concerning the impact of *Content and Language Integrated Learning* (CLIL) on the acquisition of intensifying constructions among NS and NNS of French, Dutch, and English, including amplifier-adjective bigrams, Hendrikx, Van Goethem, and Wulff (2019) showed that students in CLIL produced more target-like intensification compared with students in non-CLIL. Target likeness was shown, e.g., by less overuse of all-round intensifiers and collostructions containing advanced vocabulary as well as informal intensifying language (Hendrikx, Van Goethem, and Wulff, 2019: 97). Despite being methodologically very well-designed, Hendrikx, Van Goethem, and Wulff (2019) is less relevant in the present context as the data is significantly less formal and because the outlook of that study differs substantially from the current analysis.

While the use of adjective amplifiers has been studied rather extensively, studies which use advanced multivariate statistical modeling to understand which factors contribute to native-like use of *very* are still lacking. The present paper addresses this research gap and adds to existing research by analyzing differences in the use of *very* by NS and NNS from diverse language backgrounds using MuPDARF. Specifically, the present study aims to answer which factors, in particular language-internal and cognitive factors, correlate with native-like use of *very* among NNS from diverse language backgrounds?

1. **Data and Methodology**

This section consists of two subsections: the first subsection describes the corpora that this study is based on as well as the data processing, while the subsequent subsection describes the statistical methods that are used.

3.1 Data sources and processing

This study is based on two data sources. The first is the *International Corpus of Learner English* (ICLE, Granger et al., 2002; see Granger, 1993). The ICLE was published in 2002 and consists of argumentative writing by intermediate to advanced learners of English with Bulgarian, Czech, Dutch, Finnish, French, German, Italian, Norwegian, Polish, Russian, Spanish, and Swedish language backgrounds.

The second corpus is the *Louvain Corpus of Native English Essays* (LOCNESS) which was also compiled at the Université catholique de Louvain by a team headed by Sylviane Granger and it has been specifically designed to provide data comparable to the ICLE. The LOCNESS represents native English essays and consists of British pupils’ A level essays (60,209 words), British university students’ essays (95,695 words), and American university students’ essays (168,400 words).

The data processing of both corpora was done in the programming environment R (R Core Team 2012) and was identical for both corpora to avoid incomparability issues. In a first step, the data were cleaned and harmonized by removing, for instance, meta-data such as file identifiers. In a next step, the token frequency of adjective types was calculated separately for each language to obtain a frequency measure which can be used to control for frequency effects. Since a first MuPDARF showed that the effect of frequency exhibited substantial curvature in its correlation with the respective amplifier types in the random forest models and with nativelike choices in the regression model, the fequency was transformed into a polynomial to the second degree to accommodate for curvature in effect of frequency (see Wulff and Gries to appear).

The cleaned data were then part-of-speech tagged by implementing a maximum entropy tagger provided in the *openNLP* package (Hornik 2016). After part-of-speech-tagging, all adjectives (tag JJ) were extracted and it was determined for each adjective whether it was amplified and which lexical form served as an amplifier. The classification of adjectives as being amplified relied on a list of amplifiers[[2]](#footnote-1) and was, in addition, checked manually in order to avoid missing instances of amplification or incorrect classification. It was crucial for the present study to avoid misclassification in this respect because the type of amplifier (in the visualizations and reports of the statistical analysis referred to as *Variant*) represents the dependent variable of the random forest models. Due to the low frequency of amplifiers other than *completely*, *extremely*, *really*, *so*, and *very*, all remaining amplifier types were collapsed into a single category (*other*).

In addition to determining the amplifier variant, the adjective type was determined. Given the size of the data, adjectives that were not used at least three times in any of the language background specific subsections were collapsed. This resulted in all adjectives except for *different*, *good*, *hard*, *difficult*, and *important* being collapsed into a single category (*other*).

Next, the language background (*Language*) of the speaker in which a given adjective occurred was added. This means that it was unambiguously determined for each adjective whether it occurred in a text written by a native speaker of English or, e.g., by a learner with a French language background.

In a next step, the syntactic function of the adjective (attributive or predicative) was determined. Also, annotation was added that showed whether the same amplifier type had occurred within a span of up to three previous pre-adjectival slots to test potential persistence or priming effects (cf. Tulving and Schacter 1990: 301; Szmrecsanyi 2005: 113; also Szmrecsanyi 2006). Priming in the present study refers to instances of production priming (cf. Szmrecsanyi 2005: 113) and can therefore be defined as re-use of linguistic material that was used in the preceding discourse (cf. Tulving and Schacter 1990: 301). While there exists a substantial amount of research on priming both in psycholinguistics and corpus linguistics, various issues remain unclear. One such issue relates to the duration of priming effects as the decay time varies from milliseconds, in the case of syntactic, form, and production priming, to months or even years in cases of semantic or conceptual priming (Althaus and Kim 2006: 962). The present study uses a scope of three adjectival slots as a window in which priming may occur due to the fact that form priming is short-lived and disappears soon after exposure to the stimulus prime (Althaus and Kim 2006: 962). The current study thus assumes that priming is present if the same amplifier is reused in at least one out of the subsequent three pre-adjectival slots.

After coding for priming, negated adjectives, misclassified items, as well as comparative and superlative forms were removed from the analysis. In addition, adjectives that were never amplified, or which were not amplified by at least two different amplifier types, were removed from the analysis to avoid including lexicalized forms such as the *right honourable*.

Then, a sentiment analysis was applied to the adjectives in the data using the *syuzhet* package in R (Jockers 2017). The sentiment analysis was implemented because the amplifier choice has been shown to be affected by the emotionality of the amplified adjective. The sentiment analysis used in the current study relies on the *Word-Emotion Association Lexicon* (Mohammad and Turney 2013; cf. http://www.purl.org/net/NRCemotionlexicon), which comprises 10,170 terms based on 38,726 ratings from 2,216 raters. The emotion coding of this lexicon is based on ratings gathered through the crowed-sourced Amazon Mechanical Turk service. In this Turk survey, raters were asked a sequence of questions relating to whether a given word was associated with one of eight emotions (ANGER, ANTICIPATION, DISGUST, FEAR, JOY, SADNESS, SURPRISE, TRUST). Each term was rated at least five times and for 85 percent of words, four or more raters provided identical ratings. According to the ratings, words like *dark* or *tragic* are more readily associated with SADNESS and words such as *happy* or *beautiful* are associated with JOY while words like *cruel* or *outraged* are associated with ANGER. In the present study, adjective associated with ANGER, DISGUST, FEAR, or SADNESS are coded as negative (e.g. *sad*, *angry*) while adjectives associated with ANTICIPATION, JOY, SURPRISE, or TRUST are coded as positive (e.g. *happy*, *nice*). Adjectives that are not associated with any emotional state are coded as non-emotional (e.g. *rusty* or *flat*).

Next, all remaining adjectives were classified semantically based on a simplified version of the classification provided by Dixon (1977, 2004; cf. also D’Arcy 2015; Tagliamonte 2008; Tagliamonte and Roberts 2005). The reason for simplifying Dixon’s (1977, 2004) categorization relates to the uneven distribution of category members. Because the data contained hardly any age and colour terms, such adjective types as well as adjectives that could not be categorized (e.g. *familiar*, *genuine*, or *inadequate*) were assigned the label NoSemType. As a result, the semantic classification consisted of five levels (*Difficulty, Dimension*, *HumanPropensity*, *NoSemType*, *PhysicalProperty*, and *Value*).

In a next step, a research assistant that had prior experience in coding adjectives for gradability (a semantic property of adjectives; cf. Quirk et al. 1985) annotated each adjective as being gradable (in case the adjective denoted a point on a scale as with the adjectives *cool, hot*, *low*, *big*, *narrow*, or *intelligent*), non-gradable (if the adjective denoted a state or limits of a scale as with the adjective *pregnant*, *wooden* or *married*), or whether the adjective was neither clearly gradable nor non-gradable (as with the adjectives *subjective* or *disputable*). In addition, the classification of gradeability was checked in a data-driven way by determining whether a give adjective had been amplified by either *very* or *extremely* (indicating gradability) or by *completely*, *total*, *totally*, *utterly*, or *absolutely* (indicating non-gradability). In cases where the human expert classification differed from the results of the data-driven annotation, the adjective was reinspected. This manual re-inspection was especially necessary because despite the fact that amplifiers are commonly considered to co-occur only with gradable adjectives, they can also co-occur with non-gradable adjectives for pragmatic reasons, e.g. for emphasis. Furthermore, during its grammaticalization, *very* has spread from non-gradable to gradable adjectival contexts (Adamson and González-Díaz 2004 cited in Tagliamonte 2008) indicating that co-occurrence with non-gradable adjectives may be indicative of innovative amplifier variants.

Next, a second data set was created that only contained amplified adjective tokens. While the full data set is used to show the overall rate of amplification, the reduced data set is used in the statistical analysis. The reasoning behind the decision to remove non-amplified tokens for the statistical analysis was that the linguistic variable can be defined as a situation in which “the speaker reaches a decision-point” (Wallenberg 2013; cited in Maddeaux and Dinkin 2017). The variable context in this present study reflects the decision which amplifier a speaker uses rather than the decision whether to amplify the adjective. The variable context thus encompasses only amplified adjectives while leaving out zero context, i.e. contexts where the speaker could have amplified an adjective but did not. This means that the dependent variable is the choice of amplifier type and accordingly all not amplified adjectives had to be removed.

As a final step of data processing, a new variable (*very*) was created for the data set that only contained amplified adjectives. This categorical variable contained a 1 if the adjective was amplified by *very* and 0 if another amplifier had been used. This variable represented the dependent variable of the MuPDARF analysis described below.

3.2 Statistical methods

The present study makes use of *Multifactorial Prediction and Deviation Analysis Using Regression/Random Forests* (MuPDARF; Gries and Deshors, 2014; Heller, Bernaisch, and Gries 2017) to determine which amplifier-adjective bigrams differ significantly between NS and NNS and to provide a fine-grained understanding of which factors contribute to non-native like choices among NNS from diverse L1-backgrounds. In the following, this statistical method is explained in order to facilitate the rationale behind this procedure.

A Multifactorial Prediction and Deviation Analysis Using Regression/Random Forests (MuPDARF), for short, was only recently introduced to SLA research (see Gries and Adelman 2014; Gries and Deshors 2014; Heller, Bernaisch, and Gries 2017). MuPDARF enable us to investigate the difference between the choices of NS and NNS in a very detailed manner as it controls for confounding factors and can incorporate multiple variables and interactions between variables simultaneously. In the present case, the MuPDARF allows us to detect combinations of amplifiers and adjectives that occur significantly more or less frequently in NNS data than would be expected given NS use. In addition, the MuPDARF enables the investigation of which factors correlate with NNS making non-nativelike choices and how strongly these factors affect the degree of non-nativeness. The MuPDARF analysis is fit to the data with the occurrence of *very* versus the occurrence of other amplifiers (*other*) being the dependent variable. This is particularly interesting in the because (i) *very* is the most frequently used amplifier and (ii) because it has the potential to show how much more detailed our understanding of what affects learners’ choices can become if we make use of advanced multivariate methods. A MuPDARF analysis consists of the following steps:

1. Fit a random forest to the native-speaker data to determine which factors contributed to what extend to the decision of a native speaker.
2. Apply the random forest obtained from the native speaker data to the learner data to predict which choice a native speaker would have made given the same conditions.
3. Save the predictions of the predicted choices in a vector and compare the choices actually made by the learners to the predicted choices.
4. Then, a second vector is created in which instances where the observed and the predicted choices are identical or not. Thereby, an additional variable was added to the data which serves as the dependent variable of the mixed-effects binomial logistic regression. The newly created variable reflects whether the observed behaviour of a NNS differs from the prediction of what a NS would have done in the respective context based on the random forest model. If the observed choice of the NNS agreed with the prediction, the instance is coded as 0 while it was coded as 1 if the observed choice and the prediction differ.
5. This second vector serves as the dependent variable of a mixed-effects binomial regression model. The results of that mixed-effects regression model then inform which factors contribute to learners making nativelike choices.

The regression modelling, that was implemented during the last step of MuPARF, used both adjective type and language background as random effects. Other random effect structures were evaluated but resulted in models with a substantially higher Akaike Information Criterion (AIC) value, indicating a suboptimal model fit. Model fitting was done by following a step-wise step up procedure during which predictors (variables and interactions between variables) are consecutively added to a model. Predictors were retained if they significantly improved model fit and if their inclusion did not result in unacceptable multicollinearity. Multicollinearity was measured via variance inflation factors (VIFs) and the cut-off point for VIFs was a value of 3 (see Zuur et al. 2010). In cases where the VIFs were unacceptable, model fitting continued without that predictor. The model fitting aims to find the final minimal regression model which explains a maximum of variance with a minimum number of predictors. Once a final minimal model is found, this model provides information about which factors contribute to what extent to learners making non-nativelike choices.

1. **Results**

This section presents the results of the analysis beginning with an overview of the rate of adjective amplification by language background (Table 1). According to Table 1, the vast majority of adjectives are not amplified with an average of only 6.43 percent of adjectives being amplified. In addition, the rate of amplification is relatively homogeneous with a range of 4.9 percent (French) to 9.5 percent (English). Thus, native English speakers amplify significantly more compared to the English learners.

Table 1: Overview of adjective tokens, amplified adjective tokens, and percentage of amplification across languages.

|  |  |  |  |
| --- | --- | --- | --- |
| **Language** | **Adjective tokens** | **Amplified adj. tokens** | **Percent amplified adj. tokens** |
| Bulgarian | 8,594 | 455 | 5,3 |
| Czech | 8,345 | 648 | 7,8 |
| Dutch | 6,126 | 395 | 6,4 |
| English | 3,201 | 303 | 9,5 |
| Finnish | 7,114 | 453 | 6,4 |
| Flemish | 3,616 | 277 | 7,7 |
| French | 11,055 | 540 | 4,9 |
| German | 8,457 | 575 | 6,8 |
| Italian | 9,611 | 514 | 5,3 |
| Polish | 9,919 | 655 | 6,6 |
| Russian | 9,295 | 674 | 7,3 |
| Spanish | 7,169 | 529 | 7,4 |
| Swedish | 12,020 | 731 | 6,1 |
| **Total** | 104,522 | 6,749 | 6.46 |

Figure 1 shows the rates of amplification across languages in descending order from left to right. Figure 1 confirms that NS have a significantly higher rate of amplification compared to the NNS.

|  |
| --- |
| Figure 1: Percentages of amplified adjectives in English by L1 in descending order. |

When taking the syntactic context into account, the trend that NS amplify more than NNS remains true for attributive adjectives but not predicative adjectives (see Figure 2).

|  |
| --- |
| Figure 2: Percentages of amplified adjectives in English by L1 in descending order across syntactic contexts. |

Before turning to the results which used only data of amplified adjectives, we inspect the proficiency scores across language backgrounds (see Figure 9 in the Appendix) confirms that NS have the highest proficiency score which is comforting as it suggests that the vocabulary richness-based proficiency scores do indeed reflect proficiency.

beginning with the relative frequency of amplifier types across L1-backgrounds (Figure 2). According to Figure 2 shows that *very* is the dominant and most frequent adjective amplifier across syntactic contexts and in all L1-backgrounds.

|  |
| --- |
| Figure 3: Percentages of amplifier types of all amplifiers by language background and syntactic function. |

Figure 2 shows that the use of amplifier variants across language backgrounds is very consistent with only minor differences in the percentages of *very*. After providing an overview of the data used in the present study, the following section describes the statistical procedures used in the present study.

To tap into more very fine-grained differences between NS and NNS of English, we now turn to the MuPDARF analysis. The initial random forest analysis used the NS data with *very* as the dependent variable and was fit with 1000 trees while considering two randomly selected variables at each split. The out-of-bag error rate was relatively high with 35.16 percent but still achieved an accuracy of 72.92 percent for the training set (70 percent of the data) and an accuracy of 71.94 for a test set (30 percent of the data). Also, the initial random forest model performed significantly better than a base-line model which only predicted 55.4 percent of the data correctly and improved prediction accuracy by 38.0 percent. An overview of the variable importance of the random forest model fit to the NS data is provided in Figure 3.

|  |
| --- |
| Figure 4: Variable importance plot of the random forest fit to the NS data with use of very as dependent variable. |

The prediction of random forest for NNS data with *very* being the dependent variable achieved an accuracy of 62.26 percent accuracy which represents a significant improvement compared to a base-line accuracy of 55.5 percent by a factor of 11.9 percent.

The model fitting process which included the additional variables *Language* and *NonNativeLike* (the dependent variable for the mixed-effects modelling) arrived at a final minimal model that performed significantly better than a base-line model (L.R. χ2: 230.95, DF 8, p < .001\*\*\*). As significant predictors, the final minimal model contains the syntactic context, priming, adjective type frequency, and emotionality as main effects and an interaction between syntactic function and frequency (see Table 5 in the Appendix). The model statistics, in particular the very low pseudo-R2, C, and Somers’ Dxy values, give reasons for concern and indicate a very poor, yet significant fit. The moderate model fit requires additional attention and is addressed in the discussion section.

In the following, the individual factors reported as significant by the regression model are visualized and briefly summarized. We begin with the intercept adjustments (random effects) of the regression model. Figure 4 shows that the intercept adjustments differ substantially across adjective types. This indicates that the proportion of nativelike choices by learners varies across adjective types with the most severe problems being associated with the adjective *different*. The most nativelike choices are made with respect to *important* and *difficult* while infrequent adjective types (*other*), *hard*, and *good* require relatively little adjustment.

|  |
| --- |
| Figure 5: Adjustment by adjective type. |

Figure 4 shows that the adjustments for language are remarkably homogenous with only minor adjustments being necessary. This indicates that the language background of speakers is substantially less important than the adjective that is amplified. With respect to the minor adjustment that are necessary, speakers with Czech, Polish, and Swedish language backgrounds differ most from native speakers while speakers with Spanish, Italian, and French language backgrounds make the most nativelike choices.

|  |
| --- |
| Figure 6: Adjustment by language background. |

With respect to fixed-effects, the model reports priming, the emotionality and frequency of the adjective, the syntactic context and an interaction between syntactic context and adjective frequency as significant predictors. Neither the semantic class of the adjective nor the proficiency of the speaker or any other interaction between predictors correlated significantly with the probability to make non-nativelike choices. Since adjective frequency and syntactic context are part of a significant interaction, their main effects will not be interpreted here.

According to the regression results, learners make significantly more non-native like choices in non-primed contexts (see Figure 5) as the predicted probabilities of making a non-nativelike choice is significantly lower in primed contexts. Also, learners are more likely to make non-nativelike choices if the adjective that is intensified is not associated with a core emotion compared with adjectives that are either associated with positive or negative core emotions (see Figure 6).

|  |  |
| --- | --- |
| Figure 7: Predicted probability of non-nativelike choice by priming. | Figure 8: Predicted probability of non-nativelike choice against emotionality of adjective. |

According to Figure 7, NNS are only more likely to make non-nativelike choices in predicative contexts when the intensified adjective is relatively infrequent while the syntactic context does not appear to matter for frequent adjectives. Among frequent adjectives, the rate of non-nativelike amplification is very high regardless of the syntactic context. It is also important to note that the increase in non-nativelike choices levels off among high-frequency adjectives (frequency values above 0.4).

|  |
| --- |
| Figure 9: Predicted probability of non-nativelike choice against the frequency of amplified adjective by syntactic function. |

After reviewing the results of the statistical analysis, the following section discusses the findings in light of the relevant literature and problematizes potential issues of the statistical procedures.

1. **Discussion and Outlook**

The analysis presented here offers surprising and interesting insights. For instance, the CFA results show that, when it comes to academic writing, NNS have problems with a rather specific set of adjectives, in particular *different* and *difficult*, rather than across the board (see Table 4). The same holds true for amplifiers: while NNS overuse *completely*, *extremely*, and *really*, they rarely overuse the general, all-purpose amplifier *very* – notable exceptions being speakers with Bulgarian and Spanish L1-backgrounds who overuse *very* with *little* and *important* respectively. This finding is unexpected because it contrasts with previous research that suggested that learners significantly overuse *very* (Lorenz 1999: 198). While the findings presented here cannot substantiate overuse of *very*, all significant divergences that are detected by the configural frequency analysis represent instances of overuse, i.e. cases in which the NNS used combinations of amplifiers and adjectives significantly more frequently than would be expected given the frequencies of these combinations among NS (see Table 2). This means that NNS have a significant tendency to overuse rather than underuse amplifier-adjective combinations compared to NS. This trend holds for all amplifier-adjective bigrams with the only exception being *very* cooccurring with infrequent adjectives among speakers with a Bulgarian L1-background.

What is striking about the significant divergences that the CFAs have detected is that they suggest that NNS do not simply misapply amplifiers but that they appear to either transfer usage patterns acquired in spoken conversation to academic discourse – as in the case of *really* – or overuse rather specific combinations (*extremely difficult*, *completely different*). Indeed, *really*, in particular, has been shown to occur predominantly in spoken rather than with written discourse (Biber et al. 2007: 565) and has replaced *very* as the dominant amplifier in informal spoken conversation across various geographically distinct varieties of English (cf. D’Arcy 2015 for NZE; Ito and Tagliamonte 2003 as well as Barnfield and Buchstaller 2010 for North East British English; Tagliamonte 2008 as well as Tagliamonte and Denis 2014 for Toronto English; and Tagliamonte and Denis 2014 for South Eastern Ontario English).

The MuPDARF results provide very detailed insights into which factors correlate with non-nativelike use of *very* among NNS of English from diverse L1-backgrounds. Indeed, these results highlight what can be gained from using advanced multivariate statistical methods in learner corpus research (see also Gries 2018). The results of the MuPDARF show that the language background is substantially less important compared with the effect of adjective type. This finding supports recent research on adjective amplification which has shifted the focus from the amplifiers themselves to an approach which more thoroughly considers the interdependency of amplifiers and adjectives (see, e.g., Wagner 2017 or Hendrikx, Van Goethem, and Wulff 2019). In addition, the results of the MuPDARF procedure exemplify how learner corpus research may profit from including cognitive factors such as priming. While being unsurprising in itself, arguably because both learners and native speakers are similarly affected by priming, the finding that learners are significantly more likely to behave like native speakers in primed contexts confirms the need for cognitive factors to be integrated into language learning and language acquisition research. To elaborate, if priming had not been part of the statistical modelling, then this would have led to an overestimation of non-cognitive factors. This is so because the amount of variance that is explained by priming would could been attributed to confounding factors that are part of the statistical model.

Concerning the association of adjectives with core emotional states, the statistical analysis indicates that learners appear to have more severe issues when dealing with non-emotional adjectives compared with emotive adjectives. This finding extends previous research that shows that emotive adjectives are intensified more frequently than non-emotive adjectives (Boucher & Charles, 1969; Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). The present study indicates that emotive adjectives are not only intensified more frequently than non-emotive ones but that learners acquire amplifying strategies more readily for emotive adjectives. It seems plausible that learners are better at targeting nativelike intensification strategies when dealing with emotional adjectives because these are likely to be more cognitively salient than non-emotional adjectives are. This is relevant from a theoretical perspective because it expands previous research which showed that physical form, learner attention, and instructional focus, rather than cognitive-emotional salience, positively affected L2-acquisiton (see Cintrón-Valentín and Ellis 2016).

What is interesting about the interaction between adjective frequency and syntactic context is that the effect of syntactic context only applies to low frequency adjectives. As adjective frequency increases, the rate of non-nativelike use of *very* increases regardless of syntactic context before the effect of frequency wanes off among high-frequency adjectives. The positive correlation between non-nativelike language use and adjective frequency appears puzzling because learners should be better at nativelike production of elements that they encounter frequently. However, the opposite appears to hold true. While the view that frequency information is crucial in both first and second language acquisition has gained notable traction over the past two decades or so (see, for instance, Gass and Mackey 2002; Tomasello 2009; or Bybee 2003), it is important to understand what is meant by “frequency” in this context. Frequency should not be understood as raw frequency but rather as conditional probability, i.e., the occurrence of something given something else. This means that mere frequency of occurrence is insufficient because learners acquire linguistic structures by exploiting statistical cues in the input (see Ellis 2002). Thus, learners require additional cues for statistical learning to take effect (see Saffran et al. 1997; Saffran 2003). In the present case, the cue is clearly the adjective that is modified. With this in mind, what appears to be happening is that that the frequency effect is piggybacking on the effect of adjective type. The learners struggle with adjectives that neither infrequent nor highly frequent because in this mid-frequency-range, the learners overgeneralize to avoid acquiring errors. The strategy causes issues only among mid-frequency-range combinations because learners did not have sufficient input to fully acquire the systematicity underlying these combinations. Once the learner has had more exposure to input – as with high frequency combinations- then the error rate stagnates and eventually wanes off because the learner encountered a sufficiently large number of cues to fully acquire the systematicity of the variation.

With respect to methodology, the present study draws attention to the advantages of making use of advanced multifactorial statistics. While previous research on amplification in learner data offers very thorough and detailed qualitative insights, multivariate statistics allow us to unearth quantitative differences between learners and native speakers that have so far gone unnoticed. In elaboration of Gries (2018), the present paper exemplifies that learner corpus research can profit from carefully defining the control against which learner output is evaluated. The present study serves as a case study on why the concepts of over- and underuse can be misleading if not used in the sense of significant deviations from expected frequencies given the native speaker output (see Gries 2018).

Despite their advantages, certain issues of the statistical analysis require additional attention. The mixed-effects model has only little explanatory power and model fit criteria are below common cut-off values. Somers’Dxy, for example, ranges from -1 to 1 where values around 0 indicate pure chance. A value of merely 0.234 thus shows that the model presented here performs rather poorly. This is substantiated by a Harroll’s C value of 0.62. Harroll’s C rangers from 0 to 1 and it has been suggested that models have real predictive capacity with values above .8 (Baayen 2008: 204). However, while model fit is clearly suboptimal, the model still performs significantly better than chance (L.R. χ2: 230.95, DF 8, p < .001\*\*\*) and increase the prediction accuracy by a factor of 10.89 percent. An additional issue that requires discussion is the operationalization of frequency in the mixed-effects modelling. Frequency was operationalized as the polynomial to the second degree of the relative frequency. This was necessary because the effect of frequency was non-linear and exhibited curvature (see Wulff and Gries to appear). If frequency had not been transformed, then the mixed-effects model would have ignored the curvature in the effect of frequency. While the transformation was justified given the data at hand, the interpretation of the effect is more difficult and does not allow simple inferences of the form “if the frequency of an adjective increases x this leads to a decrease in the error rate of the learner by y”. Thus, the interpretation of the interaction between adjective frequency and syntactic function remains relatively coarse-grained – the improved fit to the data hence comes at the cost of the interpretability of the model. Also, it should be kept in mind that the corpus data used in the present study were collected about 20 years ago and future research would profit from using more recent data as well as data representing learners of non-European languages. A potential remedy for these- issues could be to use the more recent, second version of the ICLE (ICLEv2).

**References**

Adamson, S., & González-Díaz, V. 2004. Back to the very beginning: The development of intensifiers in Early Modern English. Paper presented at the *Thirteenth International Conference on English Historical Linguistics*, University of Vienna, 23-28 August 2004.

Althaus, S. L., & Kim, Y. M. 2006. “Priming Effects in Complex Information Environments: Reassessing the Impact of News Discourse on Presidential Approval”, *Journal of Politics* 68(4), 960-976.

Baayen, R. H. 2008. *Analyzing Linguistic Data. A Practical Introduction to Statistics using R*. Cambridge: Cambridge University Press.

Barnfield, K., & Buchstaller, I. 2010. “Intensifiers on Tyneside: Longitudinal developments and new trends”, *English World-Wide* 31(3), 252-287.

Bauer, L., & Bauer, W. 2002. “Adjective boosters in the English of Young New Zealanders”, *Journal of English Linguistics* 30(3), 244–257.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C. & Vohs, K. D. 2001. „Bad Is Stronger than Good”, *Review of General Psychology* 5(4), 323–370.

Biber, D., Johansson, S., Leech, G. Conrad, S., & Finegan, E. 2007. *Longman grammar of spoken and written English*. London: Longman.

Bolinger, D. 1972. *Degree words*. The Hague: Mouton.

Bortz, J., Lienert, G. A., & Boehnke, K. 2008. *Verteilungsfreie Methoden in der Biostatistik*. Heidelberg: Springer.

Boucher, J. & Charles, E. O. 1969. “The Pollyana Hypothesis”, *Journal of verbal learning and verbal behavior* 8, 1–8.

Breban, T. & Davidse, K. 2016. “The history of very: the directionality of functional shift and (inter)subjectification”, *English Language and Linguistics* 20(2), 221–249.

Bybee, J. 2003. *Phonology and Language Use*. Cambridge: Cambridge University Press.

Cintrón-Valentín, M. C. & Ellis, N. C. 2016. “Salience in Second Language Acquisition: Physical Form, Learner Attention, and Instructional Focus”, *Frontiers in Psychology* 7, 1284.

D’Arcy, A. F. 2015. “Stability, stasis and change – the longue durée of intensification”, *Diachronica* 32(4), 449–493.

Dixon, R. M. W. 1977. “Where have all the adjectives gone?”, *Studies in Language* 1, 19-80.

Dixon, R. M. W. 2004. “Adjective classes in typological perspective”. In R. M. W. Dixon & A. Y. Aikhenvald (Eds.), *Adjective classes. A cross-linguistic typology*. Oxford: Oxford University Press, 1–49.

Edmonds, A., & Gudmestad, A. 2014. “Your Participation Is greatly/highly Appreciated: Amplifier Collocations in L2 English”, *The Canadian Modern Language Review*/*La Revue canadienne des* *langues vivantes* 70(1), 76–102.

Ellis, N. C. 2002. “Frequency Effects in Language Processing”, *Studies in Second Language* *Acquisition* 24(2): 143–188.

Fallas Escobar, C., & Chaves Fernández, L. 2017. “EFL Learners’ Development of Voice in Academic Writing: Lexical Bundles, Boosters/Hedges and Stance-Taking Strategies”, *Education and Learning Research Journal* 15, 96-124.

Forsberg, F. 2010. “Using conventional sequences in L2 French”, *International Review of Applied* *Linguistics in Language Teaching* 48(1), 25–51.

Friginal, E., Lee, J. J., Polat, B., & Roberson, A. 2017. *Exploring Spoken English Learner Language* *Using Corpora. Learner talk*. Cham: Palgrave Macmillan.

Fuchs, R. & Gut, U. 2016. “Register variation in intensifier usage across Asian Englishes”. In H. Pichler (Ed.), *Discourse-Pragmatic Variation and Change: Insights from English*. Cambridge: Cambridge University Press, 185–213.

Fuchs, R. 2017. “Do women (still) use more intensifiers than men? Recent change in the sociolinguistics of intensifiers in British English”, *International Journal of Corpus Linguistics* 22(3), 345–374.

Gass, S. M. & Mackey, A. 2002. “Frequency Effects and Second Language Acquisition”, *Studies in* *Second Language Acquisition* 24(2), 249–260.

Gill, P. 2008. “Adverb Use in EFL Student Writing: From Learner Dictionary to Text Production”. In E. Bernal & J. DeCesaris (Eds.), *Proceedings of the XIII EURALEX International Congress (Barcelona, 15-19 July 2008)*. Barcelona : Institut Universitari de Lingüística aplicada Universitat Pompeu Fabra, 1301-1310.

Granger, S. 1993. “The International Corpus of Learner English”. In J. Aarts, P. de Haan, & N. Oostdijk (Eds.), *English Language Corpora: Design, Analysis and Exploitation*. Amsterdam: Rodopi, 57-69.

Granger, S. 1998. “Prefabricated patterns in advanced EFL writing: Collocations and formulae”. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications*. Oxford, UK: Clarendon, 145–160.

Gries, S. T. 2018. “On over- and underuse in learner corpus research and multifactoriality in corpus linguistics more generally”, *Journal of Second Language Studies* 1(2), 277–309.

Gries, S. T. & Adelman, A. S. 2014. “Subject realization in Japanese conversation by native and non-native speakers: exemplifying a new paradigm for learner corpus research”. In J. Romero-Trillo (Ed.), *Yearbook of Corpus Linguistics and Pragmatics 2014: New empirical and theoretical paradigms*. Cham: Springer, 35-54.

Gries, S. T. & Deshors, S. C. 2014. “Using regressions to explore deviations between corpus data and a standard/target: two suggestions”. *Corpora* 9(1), 109-136.

Heller, B., Bernaisch, T., & Gries, S. T. 2017. “Empirical perspectives on two potential epicenters: The genitive alternation in Asian Englishes”. *ICAME Journal* 41, 111-144.

Hendrikx, I., Van Goethem, K., & Wulff, S. 2019. “Intensifying constructions in French speaking L2 learners of English and Dutch. Cross-linguistic influence and exposure effects”. *International Journal of Learner Corpus Research* 5(1), 63–103.

Hinkel, E. 2003. “Adverbial markers and tone in L1 and L2 students’ writing”, *Journal of Pragmatics* 35(7), 1049-1068.

Ito, R. & Tagliamonte, S. 2003. “Well weird, right dodgy, very strange, really cool: Layering and recycling in English intensifiers”, *Language in Society* 32, 257–279.

Krauth, J., & Lienert, G. A. 1973. *Die Konfigurationsfrequenzanalyse und ihre Anwendung in* *Psychologie und Medizin*. Freiburg: Alber.

Lorenz, G. R. 1998. “Overstatement in advanced learners’ writing: stylistic aspects of adjective intensification”. In S. Granger (Ed.), *Learner English on computer*. London: Longman, 53-66.

Lorenz, G. R. 1999. *Adjective Intensification – Learners Versus Native Speakers: A Corpus Study of Argumentative Writing.* Amsterdam: Rudopi.

Lorenz, G. R. 2002. “Really worthwhile or not really significant: A Corpus-based Approach to the Delexicalisation and Grammaticalisation of Adverbial Intensifiers in Modem English”. In I. Wischer & G. Diewald (Eds.), *New Reflections on Grammaticalization*. Amsterdam: John Benjamins, 143-161.

Maddeaux, R., & Dinkin, A. 2017. “Is like like like?: Evaluating the same variant across multiple variables”, *Linguistics Vanguard*, 3(1). Available at: https://www.degruyter.com/view/j/lingvan.2017.3.issue-1/lingvan-2015-0032/lingvan-2015-0032.xml?format=INT (accessed January 2019).

Macaulay, R. 2002. “Extremely interesting, very interesting, or only quite interesting: Adverbs and social class”, *Journal of Sociolinguistics* 6(3), 398–417.

Martínez, I. M. P., & Pertejo, P. N. 2012. „He’s absolutely massive. It’s a super day. Madonna, she is a wicked singer. Youth language and intensification: A corpus-based study”, *Text and Talk* 32(6). 773–796.

Méndez-Naya, B. 2003. “On intensifiers and grammaticalization: The case of SWIÞE”, *English Studies* 84(4), 372–391.

Méndez-Naya, B. 2008. “On the history of downright”, *English Language and Linguistics* 12(2), 267-287.

Méndez-Naya, B. & Pahta, P. 2010. “Intensifiers in competition: The picture from early English medical writing”. In I. Taavitsainen & P. Pahta (Eds.), *Early Modern English medical texts: Corpus* *description and studies*. Amsterdam: John Benjamins, 191–214.

Mohammad, S. M., & Turney, P. D. 2013. “Crowd sourcing a word-emotion association lexicon”, *Computational Intelligence* 29(3), 436-465.

Nevalainen, T. 2008. “Social variation in intensifier use: Constraint on -ly adverbialization in the past?”, *English Language and Linguistics* 12(2), 289–315.

Nevalainen, T. & Rissanen, M. 2002. “Fairly pretty or pretty fair? On the development and grammaticalization of English downtoners”, *Language Sciences* 24, 359–380.

Paradis, C. 2008. “Configurations, construals and change: Expressions of DEGREE”, *English Language and Linguistics* 12(2), 317-343.

Partington, A. 1993. “Corpus evidence of language change: The case of intensifiers”. In M. Baker, G. Francis, & E. Tognini-Bonelli (Eds.), *Text and technology: In honour of John Sinclair*. Amsterdam: John Benjamins, 177-192.

Pertejo, P. N. & Martínez, I. M. P. 2014. “That’s absolutely crap, totally rubbish. The use of intensifiers absolutely and totally in the spoken language of British adults and teenagers”, *Functions* *of Language* 21(2), 210-237.

Pérez-Paredes Díez-Bedmar 2012 The use of intensifying adverbs in learner writing

Tono Kawaguchi Minegishi Developmental and Crossslinguistic

Perspectives in Learner Corpus Research

Peters, H. 1992. “English boosters: Some synchronic and diachronic aspects”. In G. Kellermann & M. D. Morrissey (Eds.), *Diachrony within synchrony: Language history and cognition*. Frankfurt: Peter Lang, 529–545.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. 1985. *A Comprehensive Grammar of the English Language*. London and New York: Longman.

Rissanen, M. 2008. “From ‘quickly’ to ‘fairly’: On the history of *rather*”, *English Language and Linguistics* 12(2), 345–359.

Saffran, J. R. 2003. “Statistical language learning: mechanisms and constraints”, *Current Directions in Psychological Science* 12(4), 110–114.

Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. 1997. “Incidental Language Learning: Listening (and Learning) Out of the Corner of Your Ear”, *Psychological Science* 8(2), 101–105.

Stenström, A.-B. 1999. “He was really gormless – she’s bloody crap: Girls, boys and intensifiers”. In H. Hasselgård & S. Oksefjell (Eds.), *Out of corpora: Studies in honour of Stig Johansson*. Amsterdam: Rodopi, 69-78.

Szmrecsanyi, B. 2005. “Language Users as Creatures of Habit: A corpus-based Analysis of Persistence in Spoken English”, *Corpus Linguistics and Linguistic Theory* 1(1), 113-150.

Szmrecsanyi, B. 2006. *Morphosyntactic Persistence in Spoken English: A Corpus Study at the Intersection of Variationist Sociolinguistics, Psycholinguistics, and Discourse Analysis*. Berlin & New York: Walter de Gruyter.

Tagliamonte, S. 2008. “So different and pretty cool! Recycling intensifiers in Toronto, Canada”, *English Language and Linguistics* 12(2), 361-394.

Tagliamonte, S. & Denis, D. 2014. “Expanding the transmission/diffusion dichotomy: Evidence from Canada”, *Language* 90(1), 90-136.

Tagliamonte, S. & Roberts, C. 2005. “So weird; so cool; so innovative: The use of intensifiers in the television series Friends”, *American Speech* 80(3), 280-300.

Tomasello, M. 2009. “The Usage-based Theory of Language Acquisition”. In E. L. Bavin (Ed.), *The Cambridge Handbook of Child Language*. Cambridge: Cambridge University Press, 69–87.

Tulving, E., & Schacter, D. L. 1990. “Priming and Human Memory Systems”, *Science* 247(4940), 301-306.

Tweedie, F. J. & Baayen, H. R. 1998. „How Variable May a Constant Be? Measures of Lexical Richness in Perspective”, *Computers and the Humanities* 32(5), 323–352.

Wagner, S. 2017. “*Totally new* and *pretty awesome*: Amplifier-adjective bigrams in GloWbE”, *Lingua* 200, 63-83.

Wallenberg, J. C. 2013. A unified theory of stable variation, syntactic optionality, and syntactic change. Paper presented at *15th Diachronic Generative Syntax Conference*, University of Ottawa, 1-3 August 2013.

Wulff, S., & Gries, S. T. 2019. “Exploring individual variation in learner corpus research: some methodological suggestions”. In B. S. W. LeBruyn & M. Paquot (Eds.). *Learner corpus research and second language acquisition*. Cambridge: Cambridge University Press.

Yule, G. U. 1944. *The Statistical Study of Literary Vocabulary*. Cambridge: Cambridge University Press.

Zuur, A. F., Ieno, E. N., & Elphick, C. S. 2010. “A protocol for data exploration to avoid common statistical problems2, *Methods in Ecology and Evolution* 1(1), 3-14.

**Software**

Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., & Matsuo, A. 2018. “quanteda: An R package for the quantitative analysis of textual data”, *Journal of Open Source Software* 3(30), 774.

Hornik, K. 2016. “*openNLP: Apache OpenNLP Tools Interface. Version 0.2-6”*. Available at: https://cran.r-project.org/web/packages/openNLP/openNLP.pdf (accessed January 2019).

Jockers, Matthew. 2017. “*syuzhet: Extracts Sentiment and Sentiment-Derived Plot Arcs from Text*. Version 1.0.1”. Available at: https://github.com/mjockers/syuzhet (accessed January 2019).

R Core Team. 2008. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org/.

**Corpora**

Granger, S., Dagneaux, E., & Meunier, F. 2002. *The International Corpus of Learner English (ICLE)*. Louvain: Centre for English Corpus Linguistics, Université Catholique de Louvain.

The *Louvain Corpus of Native English Essays (LOCNESS)*. Centre for English Corpus Linguistics, Université Catholique de Louvain. Available at: https://uclouvain.be/en/research-institutes/ilc/cecl/locness.html (accessed January 2019).

**Appendix**

|  |
| --- |
| Figure 10:Proficiency score of NS and NNS. |

Table 2: Variables and variable levels with frequency and percentage of very included in the random forest analysis.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable** | **Scale** | **Level** | **very (N)** | **Adjectives (N)** | **very (%)** |
| **Dependent variable** | | | | | |
| Variant | categorical | completely | 3,781 | 6,749 | 56.0 |
| extremely |  |  |  |
| other |  |  |  |
| really |  |  |  |
| so |  |  |  |
| very |  |  |  |
| **Independent variables** | | | | | |
| Language | categorical | English (reference var.) | 174 | 303 | 57.4 |
| Bulgarian | 212 | 455 | 46.6 |
| Czech | 373 | 648 | 57.6 |
| Dutch | 237 | 395 | 60.0 |
| Finnish | 238 | 453 | 52.5 |
| Flemish | 182 | 277 | 65.7 |
| French | 326 | 540 | 60.4 |
| German | 279 | 575 | 48.5 |
| Italian | 297 | 514 | 57.8 |
| Polish | 336 | 655 | 51.3 |
| Russian | 373 | 674 | 55.3 |
| Spanish | 351 | 529 | 66.4 |
| Swedish | 403 | 731 | 55.1 |
| Priming | nominal | NoPrime (ref. var.) | 3,379 | 6,242 | 54.1 |
| Prime | 402 | 507 | 79.3 |
| Emotionality | categorical | Negative (ref. var.) | 451 | 1025 | 44.0 |
| NonEmotional | 2260 | 3914 | 57.7 |
| Positive | 1070 | 1810 | 59.1 |
| Function | nominal | Attributive (ref. var.) | 1735 | 2659 | 65.3 |
| Predicative | 2046 | 4090 | 50.0 |
| SemanticCategory | categorical | Difficulty | 344 | 505 | 40.5 |
| Dimension | 570 | 767 | 42.6 |
| HumanPropensity | 467 | 876 | 34.8 |
| NoSemType (ref. var.) | 932 | 2086 | 30.9 |
| PhysicalProperty | 359 | 708 | 33.6 |
| Value | 1109 | 1807 | 38.0 |
| Gradability | categorical | GradabilityUndetermined (ref. var.) | 77 | 213 | 26.6 |
| Gradable | 1813 | 2988 | 37.8 |
| NotGradable | 1891 | 3548 | 34.8 |
| Adjective | categorical | different | 115 | 325 | 26.1 |
| difficult | 221 | 312 | 41.5 |
| good | 140 | 207 | 40.3 |
| hard | 86 | 133 | 39.3 |
| important | 441 | 567 | 43.8 |
| other (ref. var.) | 2778 | 5205 | 34.8 |
| Frequency | numeric |  |  |  |  |

Table 3: Overview of the absolute token frequencies and percentages of amplifier types across languages with infrequent amplifier types being collapsed into the category “other”.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Language** |  | **completely** | **extremely** | **other** | **really** | **so** | **very** | **Total** |
| Bulgarian | N | 16 | 26 | 114 | 28 | 59 | 212 | 455 |
| % | 3.5 | 5.7 | 25.1 | 6.2 | 13.0 | 46.6 |  |
| Czech | N | 31 | 14 | 98 | 50 | 82 | 373 | 648 |
| % | 4.8 | 2.2 | 15.1 | 7.7 | 12.7 | 57.6 |  |
| Dutch | N | 15 | 13 | 77 | 14 | 39 | 237 | 395 |
| % | 3.8 | 3.3 | 19.5 | 3.5 | 9.9 | 60.0 |  |
| English | N | 4 | 14 | 66 | 5 | 40 | 174 | 303 |
| % | 1.3 | 4.6 | 21.8 | 1.7 | 13.2 | 57.4 |  |
| Finnish | N | 11 | 26 | 102 | 13 | 63 | 238 | 453 |
| % | 2.4 | 5.7 | 22.5 | 2.9 | 13.9 | 52.5 |  |
| Flemish | N | 9 | 8 | 33 | 16 | 29 | 182 | 277 |
| % | 3.2 | 2.9 | 11.9 | 5.8 | 10.5 | 65.7 |  |
| French | N | 14 | 17 | 90 | 21 | 72 | 326 | 540 |
| % | 2.6 | 3.1 | 16.7 | 3.9 | 13.3 | 60.4 |  |
| German | N | 23 | 25 | 121 | 52 | 75 | 279 | 575 |
| % | 4 | 4.3 | 21 | 9 | 13 | 48.5 |  |
| Italian | N | 29 | 8 | 74 | 44 | 62 | 297 | 514 |
| % | 5.6 | 1.6 | 14.4 | 8.6 | 12.1 | 57.8 |  |
| Polish | N | 19 | 51 | 124 | 39 | 86 | 336 | 655 |
| % | 2.9 | 7.8 | 18.9 | 6 | 13.1 | 51.3 |  |
| Russian | N | 16 | 27 | 111 | 54 | 93 | 373 | 674 |
| % | 2.4 | 4 | 16.5 | 8 | 13.8 | 55.3 |  |
| Spanish | N | 23 | 19 | 52 | 27 | 57 | 351 | 529 |
| % | 4.3 | 3.6 | 9.8 | 5.1 | 10.8 | 66.4 |  |
| Swedish | N | 18 | 45 | 122 | 30 | 113 | 403 | 731 |
| % | 2.5 | 6.2 | 16.7 | 4.1 | 15.5 | 55.1 |  |
| Total | N | 228 | 293 | 1,184 | 393 | 870 | 3,781 | 6,749 |
| % | 3.4 | 4.3 | 17.5 | 5.8 | 13.0 | 56.0 |  |

Table 4: Results of the final minimal mixed-effects binomial logistic regression model of the MuPDARF procedure.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Random effects** | **Group(s)** | **Variance** | **Std. Dev.** | **L.R. χ2** | **DF** | **Pr(>|z|)** |
| Random Effect 1 | Language | 0.12 | 0.14 | 128.67 | 2 | p < .001\*\*\* |
| Random Effect 2 | Adjective | 0.22 | 0.61 | p < .001\*\*\* |
| **Fixed effects** | **Estimate** | **VIF** | **OddsRatio** | **Std. Error** | **z value** | **Pr(>|z|)** |
| (Intercept) | -0.84 |  | 0.43 | 0.22 | -3.76 | p < .001\*\*\* |
| Function:Predicative | 0.28 | 1.01 | 1.32 | 0.06 | 5.09 | p < .001\*\*\* |
| Priming:Prime | -0.65 | 1.00 | 0.52 | 0.11 | -5.90 | p < .001\*\*\* |
| Frequency | 25.5 | 1.48 | >10 | 4.14 | 6.17 | p < .001\*\*\* |
| Emotionality:NonEmotional | 0.17 | 1.68 | 1.19 | 0.08 | 2.25 | p = .0244\* |
| Emotionality:PositiveEmotional | 0.22 | 1.65 | 1.25 | 0.09 | 2.37 | p = .0179\* |
| Function:Predicative::Frequency | -15.4 | 1.47 | 0.00 | 5.10 | -3.01 | p = .0026\*\* |
| **Model statistics** |  |  |  |  |  | **Value** |
| Number of Groups |  |  |  |  |  | 12 |
| Number of cases in model |  |  |  |  |  | 6,446 |
| Observed successes |  |  |  |  |  | 3,997 |
| Residual deviance |  |  |  |  |  | 8,329.7 |
| R2 (Nagelkerke) |  |  |  |  |  | 0.056 |
| R2 (Hosmer and Lemeshow) |  |  |  |  |  | 0.032 |
| R2 (Cox and Snell) |  |  |  |  |  | 0.041 |
| C |  |  |  |  |  | 0.620 |
| Somers’ Dxy |  |  |  |  |  | 0.234 |
| AIC |  |  |  |  |  | 8,347.7 |
| BIC |  |  |  |  |  | 8,408.6 |
| Prediction accuracy |  |  |  |  |  | 63.22% |
| **Model Likelihood Ratio Test** |  | **L.R. χ2: 230.95** | | | **DF: 8** | **p < .001\*\*\*** |

1. In examples, the code ICLE stands for the corpus (ICLE = International Corpus of Learner English). The next two capital letters represent a country or language code (BG = Bulgaria, CZ = Czech Republic, DB = Dutch (Belgium), DN = Dutch (Netherlands), FI = Finland, FR = France, GE = Germany, IT = Italia, PO = Poland, RU = Russia, SP = Spain, SW = Swedish). The next three capital letters refer to the city in which the data was collected, for example LND for Lund, NIJ for Nijmegen, POZ for Poznan, etc. The numbers at the end of the code identify the essay and the author of the essay. [↑](#endnote-ref-1)
2. The list of amplifiers contained the following elements: *absolutely, awful, awfully, bloody, certainly, clearly, completely, crazy, dead, decidedly, definitely, distinctly, dreadfully, enormously, entirely, exceedingly, exceptionally, excruciatingly, extraordinarily, extremely, fiercely, firmly, frightfully, fucking, fully, genuinely, greatly, grossly, heavily, highly, hopelessly, incredibly, mad, mega, mighty, obviously, openly, overwhelmingly, particularly, perfectly, plenty, positively, precisely, pretty, profoundly, purely, real, really, remarkably, seriously, significant, significantly, so, specifically, strikingly, strongly, substantially, super, surely, terribly, terrifically, total, totally, traditionally, #true, truly, ultra, utterly, very, viciously, wholly, wicked,* and *wildly*. [↑](#footnote-ref-1)