Martin Schweinberger

**Analyzing Historical Changes in the Irish English Amplifier System**

**Abstract**

This study takes a corpus-based approach to examining the diachronic development of adjective amplification in Irish English (IrE) based on a corpus of Irish emigrant letters (*The Hamburg Corpus of Irish English; HCIE*). The results show that amplifier use in attributive contexts remains remarkably stable in the HCIE, while *so* has experienced a significant increase in frequency in predicative contexts, thereby replacing *very* as the dominant adjective amplifier in the HCIE during the first decades of the 20th-century. While this development aligns with analogous historical changes in American English, it contrasts starkly with a cross-varietal trend in spoken registers where *really*, rather than *so*, has become the dominant adjective amplifier during the 20th-century. The results presented here show that *so* and *very* are semantically highly similar and that the increase in *so* is neither triggered by significant changes in its collocational profile nor by semantic broadening or an increase in its lexical diversity. The results are therefore interpreted to suggest that the increase in *so* will remain local rather than expanding to a wider variety of contexts and genres.

**Keywords**

Adjective Amplification, Lexical Replacement, Diachronic Linguistics, Irish English

1. **Introduction**

From a language variation and change perspective, adjective amplification is particularly interesting, as this domain is prone to change (Brinton and Arnovik 2006, 441; Ito and Tagliamonte 2003, 257; Quirk et al. 1985, 590). The waxing and waning of forms, alongside invention and renewal (D’Arcy 2015, 450) in the domain of adjective amplification thus represents an area of grammar that undergoes "fevered invention" (Bolinger 1972, 18). The continuous change that is observable in the domain of adjective amplification is particularly intriguing because their changing nature predestines amplifier systems to be an ideal opportunity for testing assumptions about the underpinnings of language change. From the point of view of pragmatics, adjective amplifiers are intriguing because they play a crucial part in how speakers express themselves socially and emotionally (Labov 1985, 43; Ito and Tagliamonte 2003, 258). Thus, adjective amplifiers form part of an inventory on which speakers rely on to create, index, and mark their social identity (Tagliamonte 2011, 30).

Adjective amplification is a subtype of intensification and is related to the semantic category of degree. Accordingly, 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 (which encompasses 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*,which 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). In the present study, boosters and maximizers in both predicative and attributive contexts are considered. Differentiating between these two syntactic contexts is crucial because certain amplifier variants, for example *so*, are substantially less likely to occur in attributive contexts (although this tendency is quantitative rather than qualitative as can be seen from example 2c).

(1) Adjective amplification in predicative contexts

a. they were *very kind* to me on bord (HCIE\_Alderd01\_1880)[[1]](#footnote-1)

b. John was *prety sick* all spring (HCIE\_CarroA36\_1887)

c. I am *so glad* (HCIE\_Burke\_03\_1884)

d. I was *really glad* to hear that (HCIE\_Spratt01\_1869)

(2) Adjective amplification in attributive contexts

a. I have *very good* health (HCIE\_GambJ\_06\_1876)

b. I have *pretty good* wages at present (HCIE\_ButleJ02\_1850)

c. we have *so large* grounds to pley (HCIE\_Sloane04\_1892)

d. I got a *real nice* Passage (HCIE\_Markey01\_1889)

Before turning to an overview of previous research on adjective amplification, a note on the historicity of amplifier variants and terminology is in order. Although variants which become dominant are commonly referred to as innovative variants, the term "innovative" does not refer to a recent origin but to changes in use. Indeed, most – if not all – amplifiers which become highly frequent have a long-standing history. The most prominent amplifiers in present-day English (*very*, *really*, and *so*) have been in use as amplifiers for centuries. *So*, and its original form *swa*, has even been in use since *Beowulf* (written between 975 and 1010; see the instances provided in Tagliamonte and Roberts 2005, 293). Furthermore, *very* and *really* became the notable adjective amplifiers in the 16th and the 18th century respectively (Mustanoja 1960, 323-324; see also Ito and Tagliamonte 2003, 260). This is to say that innovative variants are not innovative in the sense that they are ‘new’ but in having undergone a functional or other usage-related change.

The next section presents previous research on amplifiers and changes in amplifier systems. Section 3 provides information about the corpus data used in the current study, elaborates on the steps undertaken during data processing, and describes the statistical analyses that were applied to the data. Section 4 presents the results of the statistical analysis while section 5 discusses the results in light of previous research and evaluates shortcomings of the present analysis.

1. **Previous Research**

Intensification has been subject to linguistic analyses for more than a century and various studies have been dedicated to investigate degree modification from historical perspective (see, for instance, Bolinger 1972; Breban and Davidse 2016; Fettig 1934; Lorenz 2002; Méndez-Naya 2003, 2008; Méndez-Naya and Pahta 2010; Nevalainen 2008; Nevalainen and Rissanen 2002; Partington 1993; Peters 1992, 1993, 1994; Rissanen 2008). Indeed, given the substantial amount of attention that intensification has attracted, the developmental pathways of individual or selected intensifiers has been subject to intensive research (e.g. Aijmer 2011, 2018a; Macaulay 2006; Méndez-Naya 2003; Pertejo and Martínez 2014; Rickford et al. 2007; Tao 2007). Furthermore, various studies which focus on Present-Day English (PDE) have investigated adjective amplifiers and their use across different age groups, genders, milieus, and regions (e.g. Aijmer 2018b; Bauer and Bauer 2002; Calle-Martín 2014; de Klerk 2005; Fuchs and Gut 2016; Fuchs 2017; Macaulay 2002; Martínez and Pertejo 2012; Paradis 1997; Pertejo and Martínez 2014; Stenström 1999; Xiao and Tao 2007). Similarly, various fine-grained variationist studies have provided detailed analyses of the trajectories of ongoing changes in intensifier systems (e.g. D’Arcy 2015; Ito and Tagliamonte 2003; Tagliamonte 2008; Tagliamonte and Denis 2014; Tagliamonte and Roberts 2005). One common pattern that has emerged from these sociolinguistic and variationist studies of PDE in that *very* is declining in both apparent and real time while *really* is increasing (see Barnfield and Buchstaller 2010 as well as Ito and Tagliamonte 2003 for North East British English; D’Arcy 2015 for NZE; Tagliamonte 2008 and Tagliamonte and Denis 2014 for Toronto and South Eastern Ontario English). In addition, these studies showed that re-arrangements in amplifier systems are accompanied by ordered extra- as well as intra-linguistic stratification (e.g. D’Arcy 2015; Lorenz 2002; Tagliamonte 2008; Tagliamonte and Roberts 2005).

With respect to intra-linguistic stratification and language internal constraints on the replacement of a dominant form by a more innovative variant, syntactic function appears to be among the most consistent factors (see Tagliamonte 2008, 373; Mustanoja 1960, 326–327). Indeed, collocation with adjectives in predicative function has been claimed to indicate later stages of change (see Tagliamonte 2008, 363 as well as Tagliamonte and Denis 2014, 116). In initial stages of change, innovative variants are typically associated with attributive contexts (Mustanoja 1960, 326-327; Tagliamonte and Denis 2014, 116) while they enter predicative contexts only during later stages of change (D’Arcy 2015, 471-472). However, with the exception of Mustanoja (1960), all studies that claimed that innovative variants enter predicative contexts only during later stages of change are based on PDE. In contrast to this, diachronic studies which used data that has a greater time-depth have shown that degree modification in attributive position arises much later than degree modification in predicative position and motivates this change in grammaticalization theory (see de Smet 2012 as well as Breban and Davidse 2016). This argument is strengthened by the fact that degree modification in predicative position is much more common than degree modification in attributive position (D’Arcy 2015: 472) which points to the earlier occurrence in predicative contexts. In addition, the fact that there is more change over time in predicative position (see e.g. Schweinberger 2020) and that predicative positions are less restrictive, for example with respect to the use of so as an adjective amplifier, also argue for the hypothesis that earlier change is bound to occur in predicative contexts.

Another factor which frequently accompanies changes in intensifier use is the association of innovative variants with negative polarity items or emotional adjectives more generally (see, for example, Peters 1994). Tagliamonte and Roberts (2005), for instance, found that that innovative intensifiers (*so* and *really*) significantly collocates with emotional adjectives (this trend is particularly strong for *so* among female speakers and restricted to speakers between the ages of 20 and 29 for *really*). According to Partington (1993, 184), the tendency to collocate with negative polarity elements is for the most part restricted to elements which are recruited from negative domains and it does not extend to elements which are recruited from a positive domain such as *very*.

With respect to collocation preferences of intensifiers, previous research has found that innovative forms are restricted to a relatively small and fixed set of adjectives. This limited set only expands once a form becomes more frequent and as collocational restrictions erode (see Lorenz 2002, 144; Méndez-Naya 2003, 377; and Tagliamonte 2008, 376). Tagliamonte and Roberts (2005) show that *really* collocates with high-frequency adjectives (and thus presumably correlates positively with adjective frequency) whereas *so* – as the most innovative intensifier in their data – prefers adjectives that are neither particularly frequent nor infrequent. The traditional form *very*, however, is the most general form with respect to its distribution and only shows a marked underuse for the three most frequent adjectives in the data of Tagliamonte and Roberts (2005). In addition, Tagliamonte and Denis (2014) show, in their research on the trajectory by which innovative variants come to dominate a linguistic subsystem, that successful variants first associate with few forms and only expand across semantic classes of adjectives, i.e. show an increase in their lexical diversity, after they have become well established.

Concerning extra-linguistic stratification, the situation is somewhat complex. Apparent-time distributions as well as multivariate statistics confirm consistent trends for age – with younger speakers preferring *really* and other innovative forms such as *dead*, *pretty* or *so* whereas older speakers exhibit a marked preference for *very* (Bauer and Bauer 2002; D’Arcy 2015; Ito and Tagliamonte 2003; Tagliamonte 2008). Ito and Tagliamonte (2003) as well as Tagliamonte (2008) found that gender differences were restricted and dependent upon speaker age, which shows that "the use of intensifiers by male and female speakers of different ages is intimately tied to the stages of intensifier renewal in the community grammar" (Tagliamonte 2008, 385).

Yet, while having been studied intensively, remarkably little research has focused on adjective amplification within and across written, but speech-like, genres. In a brief discussion, Biber et al. (2007) point out differences in the use of amplifiers between British and American English conversation and academic discourse. Regarding differences between spoken conversational and written academic texts, Biber et al. (2007, 565) found that *entirely*, *extremely*, *fully*, *highly*, and *strongly* are used more frequently in academic discourse, while the use of *very* was remarkably stable across registers. In addition, amplifiers had fewer collocations in academic prose compared with conversational data. Schweinberger (2020) analyses the fiction section of the *Corpus of Historical American English* (COHA) and finds that *so* has replaced *very* as the dominant adjective amplifier in predicative contexts in that genre during the latter half of the 20th-century. Adjective amplification in attributive contexts, however, remains notably stable. According to Schweinberger (2020), the replacement of *very* by *so* is conditioned on their semantic similarity, independent of lexical broadening, and triggered an increase in collocation strength between *so* and the two high-frequency adjectives *sorry* and *beautiful*.

In addition to the relative scarcity of research on adjective amplification in written but speech-like texts, three other research gaps exist with respect to adjective amplification. Firstly, while previous research has scrutinized changes in the use of individual or a selected subset of amplifiers, only a few studies have analyzed historical changes in an amplifier system as a whole (notable exceptions are D’Arcy 2015; Schweinberger 2020; Tagliamonte 2008; Tagliamonte and Roberts 2005; Tagliamonte and Denis 2014). This is problematic as changes in the use of one amplifier variant are depended on changes in the use of other, rivaling, variants. Secondly, to-date, no research exists which has analyzed historical changes in the Irish English amplifier system. Such a study would, however, be particularly intriguing because Irish English represents one of the few varieties of English for which there exists a sufficient amount of data to conduct fine-grained, diachronic analyses of long-term changes in adjective amplification. Thirdly, only few studies have analyzed the role of the adjectives in adjective amplification. However, recent research has shown that adjectives play a critical role in facilitating changes in amplifier use (see Schweinberger 2020; Tagliamonte 2011, 144; Tagliamonte and Denis 2014; Tagliamonte and Roberts 2005; Wagner 2017).

The present study addresses these research gaps and adds to existing research by focusing on long-term change in the Irish English amplifier system and by specifically zeroing in on interdependencies between amplifiers and adjectives. In addition, the current study extends existing research by zeroing in on a well-circumscribed and restrictive variable context (similar approaches have been undertaken by D’Arcy 2015; Tagliamonte 2008; Tagliamonte and Roberts 2005; Tagliamonte and Denis 2014), namely, the variation of amplifier variants in pre-adjectival slots. Restricting the analysis to a narrowly defined variable context allows the precise evaluation of factors contributing to variation, which would not be possible if the analysis built on relative frequencies (see Labov 1966, 49; Tagliamonte 2011, 9-10).

1. **Data and Methodology**

The following section consists of two subsections: the first subsection provides information about the HCIE corpus and describes the data processing while the second subsection describes the statistical methods that have been applied to the data.

* 1. **Corpus Description and Data Processing**

To study adjective amplification in Irish English, the present analysis takes a corpus-based approach using *The Hamburg Corpus of Irish English* (HCIE; Siemund and Pietsch unpublished)[[2]](#footnote-2) which mainly represents private letters written back to Ireland from America and Australia by approximately 400 Irish emigrants between 1675 and 1931 (Pietsch 2009, 545) (see Table 1). Due to the semi-literate status of the writers, orthography in the HCIE is highly variable and mirrors the writers’ or scribes’ pronunciation (see, for instance, the examples in (1) and (2)). In addition, the style of the letters represents a hybrid of speech-like writing and formal phraseology. In terms of style, however, the data is clearly more speech-like or speech-based than written. As such, the material is more similar to transcribed court proceedings which have been analyzed, for instance, by Claridge and Kytö (2014) compared to analyses of written texts proper. The relatively higher rate of poverty and, thus, semi-literate writers in and from Ireland – compared to authors of private letters from England, for instance – make the study of IrE particularly intriguing. This is so because texts produced by semi-literate writers are more speech-like than letters produced by authors who were taught to write in school. As such, semi-literate IrE resembles historical spoken language more closely than other sources of historical data (an exception being the Old Bailey Corpus which represents transcripts of spoken court proceedings, see again, for instance, Claridge and Kytö 2014). Table 1 provides an overview of the HCIE data.

Table 1: Overview of number of letters and words per period in the HCIE.

|  |  |  |
| --- | --- | --- |
| **Period** | **Letters (N)** | **Words (N)** |
| 1675-1750 | 16 | 9,818 |
| 1751-1850 | 355 | 225,032 |
| 1901-1935 | 608 | 322,107 |
| Unknown | 1 | 223 |
| **Total** | **980** | **557,180** |

To extract all adjectives, the data was part-of-speech tagged in the programming environment R (R Development Core Team 2012) by implementing a maximum entropy tagger provided in the openNLP package (Hornik 2016). After part-of-speech-tagging, adjectives that occurred in texts which could not unambiguously be assigned to a date were removed and, subsequently, all adjectives (part-of-speech tag "JJ") were extracted – this encompassed base forms and participles while excluding comparative and superlative forms. Next, it was determined if an adjective was amplified and, if so, by which type of amplifier. At this stage, the data was spell-checked and orthography of adjectives and amplifiers was harmonized, i.e. spelling variants were subsumed under a single label. For instance, the variants *veery*, *verey*, *verrey*, *verrry*, *verry*, and *very* were all subsumed under the label *very*. The next step consisted in determining the date when the text was written and the syntactic function of the adjective (predicative vs. attributive). If the syntactic function could not be determined, the adjective was removed from the analysis.

Furthermore, if an adjective token was preceded by a downtoning intensifier (e.g. *sort of*, *kind of*, *a little bit*, *somewhat*, *slightly*, *barely*, *hardly*, etc.), by a negative form (e.g. *not*, *never*, etc.), by forms indicating comparison (*more* and *most*), or by a special form (i.e. *too*, *quite*, etc.), the respective adjective token was removed from the analysis. By excluding negated adjectives, the present study follows Tagliamonte (2008) and D’Arcy (2015) who argue that such contexts require removal as negation affects the semantics of the amplified adjective. For example, *very good* and *not very good* or *very tasty* and *not very tasty* are not semantically symmetric antonyms as *not very good* does not necessarily mean *bad* and *not very tasty* does not necessarily imply that something is *inedible*. Next, adjective types that were never amplified or that were not amplified by at least two distinct amplifier types were removed from the analysis to prevent the inclusion of lexicalized forms (e.g. *right honorable*) and to allow the researcher to zoom in on a restrictive but well-circumscribed variable context. As a final step, the data were manually cross-checked to minimize the amount of errors arising from the (semi-)automatic data processing.

* 1. **Statistical Methods**

The present study makes use of three types of analyses: (1) semantic vector space modelling to investigate semantic similarities of individual amplifier variants, (2) lexical diversity scores (LD) to investigate range of adjectives that amplifier variants co-occur with, and (3) covarying collexeme analysis to investigate changes in the collocational profiles of amplifiers. The latter two analyses are complementary in that lexical diversity scores provide information about the general profile of amplifier variants as it indicates their breath by showing how many different adjective types amplifier variants co-occur with while covarying collexeme analysis evaluate the collocational strength between individual amplifier variants and specific adjective types.

**3.2.1 Semantic Vector Space Modelling**

Semantic vector space models allow for the inspection of the relative similarity of amplifier types based on their co-occurrence frequencies with adjectives (also referred to as vectors or embeddings) (see Levshina 2015). The underlying reasoning is founded in distributional semantics, according to which, words that share many collocates are semantically similar (Stefanowitsch 2010, 368-370). In this view, semantic similarity can be measured in terms of the similarity of collocation profiles. The results of vector space models are commonly displayed in the form of dendrograms in which semantically similar elements inhabit the same branches.

**3.2.2 Lexical Diversity**

To analyse changes in the diversity of the collocational profiles of amplifiers in the HCIE, the current study uses a simple lexical diversity measure (LD). LD scores are calculated by dividing the number of adjective types a given amplifier co-occurs with by the number of tokens of that amplifier type (see 3).

(3) LD = NAdj.Types/NAmp. Tokens

LD can reach a maximum value of 1 in which case it indicates high lexical diversity. The lower the LD value, the lower the degree of lexical diversity (see Table 2).

Table 2: Example of LD calculation.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Amplifier** | **Amp. Tokens (N)** | **Adj. Types (N)** | **Calculation** | **LD value** |
| variantA | 10 | 1 | 1/10 | .1 |
| variantB | 10 | 5 | 5/10 | .5 |
| variantC | 10 | 10 | 10/10 | 1 |
| very1675-1750 | 11 | 17 | 11/18 | .647 |
| very1751-1850 | 86 | 368 | 86/368 | .234 |
| very1851-1930 | 102 | 845 | 102/845 | .121 |

As Table 2 shows, the LD values of *very* decrease over time because there are 11 tokens of *very* that co-occur with merely 17 adjective types in letters written between 1675 and 1750 while there are 102 tokens of *very* that co-occur with 845 distinct adjective types in letters written between 1851 and 1930. The LD value allows testing whether the lexical diversity of an amplifier variant changes over time and thus whether amplifier variants are semantically broadening, i.e. becoming more general, or narrowing, i.e. specializing on modifying fewer adjectives across real time.

**3.2.3 Covarying Collexeme Analysis**

To analyse changes in the attraction between the most common amplifier variants and selected individual adjectives, the present study makes use of covarying collexeme analysis which is part of the collostructional family of analyses (Gries and Stefanowitsch 2004; Hilpert 2006; Stefanowitsch and Gries 2003, 2005). Covarying collexeme analyses allow for the quantification and evaluation of attraction between elements that occur in two distinct slots within a specified construction. In the present case, the first slot represents the amplifier slot and the second slot represents the adjective slot. Each slot can be occupied by a variant from a set of potential candidates – the set of amplifier variants for the first slot and the set of adjectives for the second slot. Covarying collexeme analysis provides information about whether the likelihood of a certain variant in the first slot affects the likelihood of another variant from another set occuring in the second slot. In other words, it is more likely that *nice* occurs in the second slot given that *really* occurs in the first slot compared with another amplifier taking the first slot. The p-values reported by the covarying collexeme analysis are Bonferroni corrected to control the inflation of α-error rates. This is necessary because a huge number of tests were run which would have resulted in an inflation of false positive results (α-error) if the corrections had not been applied. The covarying collexeme analyses were performed for each period (1675 to 1750, 1751 to 1850, and 1851 to 1930). The effect size measure reported here is the logged p-value (Stefanowitsch and Gries 2005). The values of this effect size measure inform about whether the amplifier and adjective repel or attract each other. Values below 0 indicate rejection while values above 0 indicate attraction and values around 0 do neither indicate preference nor rejection. In statistical terms, values below 0 show that an amplifier and an adjective occur less frequently together than would be expected by chance while values above 0 show that an amplifier and an adjective occur more often together than expected by chance. For the analysis, all adjectives other than *dangerous*, *different*, *difficult*, *good*, *important*, *little*, *necessary*, and *new* were collapsed into one category (*other*). The advantage of covarying collexeme analyses over similar methods to evaluate collocational attraction is that it is a very robust method as it is an extension of Fisher’s Exact Test that does not rely on distributional assumptions as tests form the χ2-family of tests do. The following section presents the results of the quantitative analysis.

1. **Results**

The final data consisted of 11,947 adjectives of which 13.98% were amplified (see Table 1). The most frequent adjective amplifier in the data is *very* with 1,230 instances which amounts to 73.65 percent of all instances of amplification. The second most frequent adjective amplifier is *so* with 283 instances representing 16.95 percent of overall amplification which goes to show not only that *very* is the dominant adjective amplifier in the HCIE data but that it is pronouncedly so as all other amplifiers combined make up merely 26.35% of all instances of amplification.

Table 3: Overview of the cleaned and processed data.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variant** | **N** | **%** | **% (of Amp.)** |
| Ø (not amplified) | 10,277 | 86.02 |  |
| very | 1,230 | 10.30 | 73.65 |
| so | 283 | 2.37 | 16.95 |
| pretty | 62 | 0.52 | 3.71 |
| real | 17 | 0.14 | 1.02 |
| truly | 14 | 0.12 | 0.84 |
| awefully | 13 | 0.11 | 0.78 |
| extremely | 9 | 0.08 | 0.54 |
| perfectly, really | 5 (10) | 0.08 (0.04) | 0.6 (0.3) |
| exceedingly | 4 | 0.03 | 0.24 |
| dreadfully, remarkably, surely, terribly, unusually | 2 (10) | 0.2 (0.02) | 0.12 (1.2) |
| certainly, dead, deeply, excessively, excellently, fully, greatly, imminently, miserably, perfect, sincerely, vastly | 1 (12) | 0.01 (0.12) | 0.06 (0.72) |
| **Total** | 11,947 | 100.0 | 100.0 |

Figure 1 displays the rate of amplification among adjectives in the HCIE data and shows that the rate of amplification in attributive contexts is lower compared to the rate of amplification in predicative contexts and that rate of amplification in attributive contexts is highly stable while the rate of amplification in predicative contexts has steadily increased since 1750 after having experienced a drop in the earliest subsection of the data.

|  |
| --- |
| Figure 1: Percent of amplified adjectives across syntactic contexts and time in the HCIE. |

It should be borne in mind, however, that the percentage values depicted in Figure 1 are restricted to the variable context, i.e. not all adjective slots are considered. The dendrogram shown in Figure 2 displays the results of the Semantic Vector Space model. According to Figure 2, *very* and *so* are highly semantically similar based on their co-occurrence profiles with adjectives and stand apart from all other amplifiers. In addition, variants which are phonetically and orthographically similar (*true* and *truly* as well as *real* and *really*) are also reported to behave very similarly with respect to their collocational preferences.

|  |
| --- |
| Figure 2: Horizontal cluster dendrogram showing the semantic similarity of adjective amplifiers in the HCIE based on their co-occurrence frequencies with adjectives. |

Figure 3 shows the distribution of *very*, *so*, and *other* amplifier types in attributive and predicative contexts in the HCIE across real time to enable a more fine-grained understanding of changes in amplifier use.

|  |
| --- |
| Figure 3: Distribution of amplifiers across syntactic contexts and time in the HCIE. |

According to Figure 3, the distribution of amplifier types in attributive contexts is remarkably stable. In predicative contexts, *so* has replaced *very* as the dominant adjective amplifier during the first decades of the 20th-century. In addition, *so* – which appears to be the sole rival of *very* in the Irish data – is almost categorically restricted to predicative contexts.

We now turn to the LD scores of adjective amplifiers to gain a better understanding of changes in the range of adjectives that variants co-occur with. The distribution of LD scores across real time is shown in Figure 4. In attributive contexts, the LD scores of *very* decrease over time while the LD scores of *so* and *other* amplifiers remain stable and only show some minor fluctuation. In predicative contexts, LD scores of *very*, *so*, and *other* adjective amplifiers decrease almost uniformly over time with *very* consistently having the lowest LD scores. Interestingly, there is no noticeable change in the patterning between the 1851 and 1930, i.e. during the time period when *so* has replaced *very* in predicative contexts.

|  |
| --- |
| Figure 4: Changes in LD scores of very, so, and other adjective amplifiers across syntactic contexts and time in the HCIE.[[3]](#footnote-3) |

There are at least two issues related to the current measurement and display of lexical diversity. Firstly, the LD scores do not control for frequency and, thus, the most frequent adjective amplifier is bound to have the lowest LD sore by chance alone. However, this also implies that the LD scores for *so* should dip when *so* experienced an increase in frequency which is not the case.

Secondly, the distributions shown in Figure 4 could be affected by changes in the token frequency of adjectives. In other words, the changes in amplifier use could be confounded by changes in the frequency of adjectives. Indeed, a variant such as *so* could become dominant, not because it undergoes some kind of change, but merely because the adjectives it collocates with become more frequent. To control for changes in adjective frequency as a confounding factor, the percentages of adjectives across real time are displayed in Figure 5.

|  |
| --- |
| Figure 5: Changes in the use of adjective types across syntactic contexts and time in the HCIE. |

Figure 5 indicates that the adjective system in HCIE data is remarkably stable (adjective types other than the seven most frequent adjectives (*dear, few, good, great, last, little,* and *old*) were merged collapsed into the category *other*). The only trend that emerges from Figure 5 is a slow and steady increase in the use of infrequency adjectives. The stability of the adjectival system suggests that changes in the frequency of individual adjective types cannot be responsible for the consistency in the patterning of LD scores shown in Figure 4.

The covarying collexeme analysis did not detect any significant collocations after being Bonferroni-corrected, i.e. significance levels were adapted to control for repeated testing. This means that none of the combinations of amplifiers and adjectives occurred significantly more or less frequently than would be expected by chance. However, the analyses did reveal interesting changes in collocation strength (see Figure 6).

|  |
| --- |
| Figure 6: Collocation strength of adjective amplifiers by adjective type across syntactic contexts and time in the HCIE.[[4]](#footnote-4) |

The panels in Figure 6 reveal that even though the amplifier system in attributive contexts is highly stable, there have been collocational changes below the apparently conservative surface. With respect to individual patterns, *very* is most strongly associated with *good* in both syntactic contexts and preferred by *other* adjectives in predicative contexts in the latest subsection of the data – as indicated by the trends in the solid lines. *Other* amplifiers strongly associate with *bad* and *glad* in predicative contexts and with *other* adjectives in attributive contexts during the period from 1751 to 1850. *So* has experienced an increase in collocational strength with *glad* and *good* in predicative contextsas well as with *other* adjectives in attributive contexts as shown by the upward trends in the dotted line. Given the results of the covarying collexeme analyses, the attraction between *so* and *good* in predicative contexts, although appearing to be only a minor increase in attraction, may quite likely be the reason why *so* has replaced *very* as the dominant amplifier in the most recent data due to *good*'s sheer frequency. The fact that *so* increasingly associates with *other* adjectives in attributive contexts is also interesting as it poses the question of whether broadening in in one syntactic context could affect the occurrence of a variant in another context. A more thorough discussion of the findings is provided below.

1. **Discussion**

The current analysis of the historical development of adjective amplification in the HCIE has unearthed intriguing and unexpected findings. The analysis of Irish emigrant letters shows a remarkably stable amplifier system in attributive contexts and the replacement of *very* by *so* as the dominant adjective amplifier in predicative contexts during the first decades of the 20th-century (see Figure 3). The stability of adjective amplification in attributive positions in itself appears to be a remarkable finding because it stands in stark contrast to previous claims. In fact, the stability of the adjective amplifier system in the HCIE does not conform with the notion that the domain of adjective amplification, or the domain of intensification more generally, is prone to change (Brinton and Arnovik 2006, 441; Ito and Tagliamonte 2003, 257; Quirk et al. 1985, 590), a site of invention and renewal (D’Arcy 2015, 450), and an area of grammar that undergoes "fevered invention" (Bolinger 1972, 18). Since Schweinberger (2020) found a very similar trend in historical fiction data in the COHA, the question arises if such statements can be maintained in their universality or whether the idea of constant renewal is constrained by extra-linguistic restrictions regarding genre or text type. This finding loses its surprising aspect, however, once we consider previous research that has focused on long-term diachronic change which showed that it is precisely predicative contexts that are the locus of innovation and the position of earlier stages of change (see de Smet 2012). As such, it is even expected that predicative contexts show change before innovations diffuse to the more conservative attributive contexts – not last because attributive contexts are more restrictive, for instance with respect to innovative variants such as *so*.

This study found that only 13.98 percent of adjectives are amplified with *very* being the most frequent adjective amplifier in the HCIE. The 1,230 instances of *very* in the Irish emigrant letters amount to 73.65 percent of all instances of amplification while *so*, the second most frequent variant, occurs only 283 times (16.95 percent; see Table 2). Overall, more adjectives are amplified in predicative contexts compared with attributive contexts (Figure 1) – which again substantiates that predicative contexts are more likely to allow change due to higher frequency of use. While the dominance of *very* across both syntactic contexts is somewhat expected given the fact that very represents the formal adjective amplifier *par excellence*. However, this finding is still highly relevant as it draws additional attention to the fact that adjective amplification differs substantially across modalities: while there are very clear indications that *really* and *pretty*, rather than *so*, have replaced *very* in spoken dialogues and narratives (see Barnfield and Buchstaller 2010; D’Arcy 2015; Ito and Tagliamonte 2003; Schweinberger 2020; Tagliamonte 2008; Tagliamonte and Denis 2014), this cross linguistic trend does not hold true for written data.

A semantic vector space model confirmed that *so* and *very* are semantically very similar (Figure 2). This finding lends support to previous research which has proposed that semantic similarity is a necessary yet not sufficient condition for innovative variants to replace dominant forms (see Schweinberger 2020). To elaborate, this finding substantiates the hypothesis that an innovative variant can only become dominant if its collocational profile is sufficiently similar to the traditional variant.

Another particularly puzzling result of the present study relates to the observation that the increase in *so* in the present data is not accompanied by a broadening of its collocational profile. The lexical diversity values for *so* in predicative contexts show a steady decrease and remain consistently above the values of *very* when *so* increased in frequency (see Figure 3). This finding is independent of frequency fluctuations within the adjective system (see Figure 5). This lack of broadening of *so* contrasts with hypotheses of language change which assume that increases in frequency are dependent upon either syntactic or lexico-semantic broadening. For instance, Mair’s (2004) "delayed increase of discourse frequency" hypothesis but also Bybee, Perkins and Pagliuca (1994, 8) suggest that the frequency of innovative or intrusive forms "escalates", only once such forms enter new contexts. Hence, the lack of broadening of *so* in the HCIE data requires an explanation. One explanation which is supported by the covarying collexeme analyses is that the increase in *so* is caused by its attraction to *good* and *glad.* This explanation ties in with the interdependency between high-frequency adjectives and amplifiers shown by Schweinberger (2020) and it highlights the pivotal role that adjectives play as vehicles that drive the restructuring of amplifier systems (see also Tagliamonte 2011, 144; Wagner 2017). Indeed, the increase in frequency of *so* can be seen as a consequence of its increased co-occurrence with *good* as the most frequent adjective in predicative contexts. From this perspective, the increase in *so* was not accompanied by a broadening of contexts because it expanded in a context that it had already entered. Also, the few instances of *so* that associated with *other* adjectives in attributive contexts additionally could have weakened existing collocational ties which then resulted in *so* being able to attach to *good* in its preferred context of use, i.e. in predicative position.

This is also supported by the fact that the decrease in lexical diversity of *so* in predicative contexts is, indeed, indicative of specialization rather than broadening. An expansion of its collocational profile would then occur after the form has already become highly frequent. This assumption that broadening only occurs after an initial phase of specialization is also supported by previous research (Schweinberger 2020; Tagliamonte and Roberts 2005; Tagliamonte and Denis 2014).

However, despite their general semantic similarity, the covarying collexeme analysis has shown that *so* and *very* still exhibit distinct semantic profiles with respect to their semantic specialization (see Figure 6). Accordingly, it would not be quite accurate to consider the increase in the use of *so* and the simultaneous decrease of *very* as representative of lexical replacement proper. This is so because lexical replacement requires not only replacement in terms of frequency but also in terms of function (D’Arcy 2015).

One issue which remains unanswered in the present study relates to the role that social factors have played in the restructuring of the adjective amplifier system of Irish English. The inclusion of social factors into the analysis of adjective amplification is warranted because the "use of linguistic forms to increase distinctiveness of particular groups is a driving force for the acceleration of change" (Labov 2002, 19). Thus, future research should go beyond focusing on the interdependencies between amplifiers and adjectives and take a multivariate approach which simultaneously takes various intra- and extra-linguistic factors into account. In this respect, the current analysis represents a starting point for more in-depth studies on factors that cause *so*, rather than *really*, to become dominant in predicative contexts in the written domain.

1. **References**

Aijmer, Karin. "Are you totally spy? A new intensifier in present-day American English." *Marqueurs discursifs et subjectivité*. Ed. Sylvie Hancil. Rouen: Universités de Rouen and Havre, 2011. 155-172.

Aijmer, Karin. "That’s well bad. Some new intensifiers in spoken in British English." *Corpus Approaches to Contemporary British English*. Eds. Vaclav Brezina, Robbie Love, and Karin Aijmer. New York and London: Routledge, 2018a. 60-95.

Aijmer, Karin. “Intensification with 'very', 'really' and 'so' in selected varieties of English.” *Corpora and Lexis.* Eds*.* Sebastian Hoffmann, Andrea Sand, Sabine Arndt-Lappe, and Lisa Marie Dillmann. Leiden and Boston: Brill Rodopi, 2018b. 106-139.

Barnfield, Katie and Isabelle Buchstaller. "Intensifiers on Tyneside: Longitudinal developments and new trends." *English World-Wide* 31.3 (2010): 252-287.

Bauer, Laurie and Winifred Bauer. "Adjective boosters in the English of Young New Zealanders." *Journal of English Linguistics* 30.3 (2002): 244-257.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan. *Longman Grammar of Spoken and Written English*. 6th edition. Pearson: Harlow, UK, 2007.

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

Breban, Tine and Kristin Davidse. "The history of very: the directionality of functional shift and (inter)subjectification." *English Language and Linguistics* 20.2 (2016): 221-249.

Brinton, Laurel J. and Leslie K. Arnovik. *The English language. A linguistic history*. Oxford: Oxford University Press, 2006.

Bybee, Joan L., Revere D. Perkins, and William Pagliuca. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press, 1994.

Calle-Martín, Javier. "On the history of the intensifier *wonder* in English." *Australian Journal of Linguistics* 34.3 (2014): 399-419.

Claridge, Claudia and Kytö, Merja. “‘You are a bit of a sneak’: Exploring a degree modifier in the Old Bailey Corpus”. *Late Modern English Syntax*. Ed. Marianne Hundt. Cambridge: Cambridge University Press, 2014. 239–268.

D’Arcy, Alexandra F. "Stability, stasis and change – the *longue durée* of intensification." *Diachronica* 32.4 (2015): 449-493.

de Klerk, Vivian. "Expressing levels of intensity in Xhosa English." *English World-Wide* 26.1 (2005): 77-95.

Fettig, Adolf. *Die Gradadverbien im Mittelenglischen*. Heidelberg: Winter, 1934.

Fuchs, Robert and Ulrike Gut. "Register variation in intensifier usage across Asian Englishes." *Discourse-Pragmatic Variation and Change: Insights from English*. Ed. Heike Pichler. Cambridge: Cambridge University Press, 2016. 185-213.

Fuchs, Robert. "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 (2017): 345-374.

Gries, Stefan Th. and Anatol Stefanowitsch. "Extending collostructional analysis: A corpus-based perspectives on 'alternations'." *International Journal of Corpus Linguistics* 9.1 (2004): 97-129.

Hilpert, Martin. "Discussion note: Distinctive collexeme analysis and diachrony." *Corpus Linguistics and Linguistic Theory* 2.2 (2006): 243-256.

Ito, Rika and Sali Tagliamonte. "*Well weird*, *right dodgy*, *very strange*, *really cool*: Layering and recycling in English intensifiers." *Language in Society* 32 (2003): 257-279.

Labov, William. *The Social Stratification of English in New York City*. Washington, DC: Center for Applied Linguistics, 1966.

Labov, William. "Intensity." *Meaning, form and use in context: Linguistic applications*. Ed. Deborah Schiffrin. Washington, DC: Georgetown University Press, 1985. 43-70

Labov, William. "Driving forces in linguistic change." 2002. *Paper presented at the International Conference on Korean Linguistics*. Web. 9 October 2019.

Levshina, Natalia. *How to do linguistics with R: Data exploration and statistical analysis*. Amsterdam: John Benjamins, 2015.

Lorenz, Gunter R. "Really worthwhile or not really significant: A corpus-based approach to the delexicalisation and grammaticalisation of adverbial intensifiers in Modem English." *New Reflections on Grammaticalization*. Eds. Ilse Wischer and Gabriele Diewald. Amsterdam: John Benjamins, 2002. 143-161.

Macaulay, Ronald. "Extremely interesting, very interesting, or only quite interesting: Adverbs and social class." *Journal of Sociolinguistics* 6.3 (2002): 398-417.

Macaulay, Ronald. "Pure grammaticalization: The development of a teenage intensifier." *Language Variation and Change* 18.3 (2006): 267-283.

Mair, Christian. "Corpus linguistics and grammaticalisation theory. Statistics, frequencies and beyond." *Corpus approaches to grammaticalization in English*. Eds. Hans Lindquist and Christian Mair. Amsterdam and Philadelphia: John Benjamins, 2004. 121-150.

Martínez, Ignacio M. Palacios and Paloma Núñez Pertejo. "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 (2012): 773–796.

Méndez-Naya, Belén. "On intensifiers and grammaticalization: The case of swiþe." *English Studies* 84.4 (2003): 372-391.

Méndez-Naya, Belén. "Introduction. Special issue on English intensifiers." *English Language and Linguistics* 12.2 (2008): 213-219.

Méndez-Naya, Belén and Pahta, Päivi. "Intensifiers in competition: The picture from early English medical writing." *Early Modern English medical texts: Corpus description and studies*. Eds. Irma Taavitsainen and Päivi Pahta. Amsterdam: John Benjamins, 2010. 191-214.

Mustanoja, Tauno F. *A Middle English syntax, part 1, Parts of speech*. Helsinki: Société Néophilologique, 1960.

Nevalainen, Terttu. "Social variation in intensifier use: constraint on -ly adverbialization in the past?" *English Language and Linguistics* 12.2 (2008): 289-315.

Nevalainen, Terttu and Matti Rissanen. "Fairly pretty or pretty fair? On the development and grammaticalization of English downtoners." *Language Sciences* 24 (2002): 359-380.

Paradis, Carita. *Degree modifiers of adjectives in spoken British English*. Lund: Lund University Press, 1997.

Partington, Alan. "Corpus evidence of language change: The case of intensifiers." *Text and technology: In honour of John Sinclair*. Eds. Mona Baker, Gill Francis, and Elena Tognini-Bonelli. Amsterdam and Philadelphia: John Benjamins, 1993. 177-192.

Pertejo, Paloma Núñez and Ignacio M. Palacios Martínez. "*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 (2014): 210-237.

Peters, Hans. "English boosters: Some synchronic and diachronic aspects." *Diachrony within synchrony: Language history and cognition*. Eds. Günter Kellermann and Michael D. Morrissey. Frankfurt: Peter Lang, 1992. 529-545.

Peters, Hans. *Die englischen Gradadverbien der Kategorie booster*. Tübingen: Narr, 1993.

Peters, Hans. "Degree adverbs in Early Modern English." *Studies in Early Modern English*. Ed. Dieter Kastovsky. Berlin and New York: Mouton de Gruyter, 1994. 269-288.

Pietsch, Lukas. "Hiberno-English medial-object perfects reconsidered. A case of contact-induced grammaticalisation." *Studies in Language* 33.3 (2009): 528-568.

Quirk, Randolph, Sydney Greenbaum, Geoffrey Leech, and Jan Svartvik. *A Comprehensive Grammar of the English Language*. London and New York: Longman, 1985.

Rickford, John, Thomas Wasow, Arnold Zwicky, and Isabelle Buchstaller. "Intensive and quotative *all*: something old, something new." *American Speech* 82.1 (2007): 3-31.

Rissanen, Matti. "From 'quickly' to 'fairly': On the history of *rather*." *English Language and Linguistics* 12.2 (2008): 345–359.

Schweinberger, Martin. "Analyzing change in the American English amplifier system in the fiction genre." *Corpora and the Changing Society. Studies in the evolution of English, Part II*. Eds. Paula Rautionaho, Arja Nurmi, and Juhani Klemola. Amsterdam: John Benjamins, 2020. *223–249.*

Stefanowitsch, Anatol. "Empirical cognitive semantics: Some thoughts." *Quantitative Cognitive Semantics: Corpus-driven approaches*. Eds. Dylan Glynn and Kerstin Fischer. Berlin: Mouton de Gruyter, 2010. 355-380.

Stefanowitsch, Anatol and Stefan T. Gries. "Collostructions: Investigating the interaction between words and constructions." *International Journal of Corpus Linguistics* 8.2 (2003): 209-243.

Stefanowitsch, Anatol and Stefan T. Gries. "Covarying collexemes." *Corpus Linguistics and Linguistic Theory* 1.1 (2005): 1-43.

Stenström, Anna-Brita. "He was really gormless – she’s bloody crap: Girls, boys and intensifiers." *Out of corpora: Studies in honour of Stig Johansson*. Eds. Hilde Hasselgård and Signe Oksefjell. Amsterdam: Rodopi, 1999. 69-78.

Tagliamonte, Sali A. "So different and pretty cool! Recycling intensifiers in Toronto, Canada1." *English Language & Linguistics* 12.2 (2008): 361-394.

Tagliamonte, Sali A. *Variationist Sociolinguistics. Change, Observation, Interpretation*. Oxford and Malden, MA: Wiley & Sons, 2011.

Tagliamonte, Sali A. and Derek Denis. "Expanding the transmission/diffusion dichotomy: Evidence from Canada." *Language* 90.1 (2014): 90-136.

Tagliamonte, Sali and Chris Roberts. "So weird; so cool; so innovative: The use of intensifiers in the television series Friends." *American Speech* 80.3 (2005): 280-300.

Tao, Hongyin. "A corpus-based investigation of absolutely and related phenomena in spoken American English." *Journal of English Linguistics* 35.1 (2007): 5–29.

Wagner, Susanne. "Totally new and pretty awesome: Amplifier-adjective bigrams in GloWbE." *Lingua* 200 (2017): 63-83.

Xiao, Richard and Hongyn Tao. "A corpus-based sociolinguistic study of amplifiers in British English." *Sociolinguistic Studies* 1.2 (2007): 241-273.

1. **Corpora and Software**

Hornik, Kurt. Package "OpenNLP – Apache OpenNLP Tools Interface." 2016. Web. 10 October 2019.

R Development Core Team. "R: A language and environment for statistical computing." 2012. *R Foundation for Statistical Computing. Vienna, Austria*. Web. 10 October 2019.

Siemund, Peter and Lukas Pietsch. "The Hamburg Corpus of Irish English." unpublished. *Universität Hamburg. Collaborative Research Center: Multilingualism. Project H5: Hiberno-English: Variation and Universals in Contact-Induced Language Change*. Web. 10 October 2019.

1. The annotation of the examples indicates the corpus (HCIE), followed by a sequence with represents the author combined with a number that identifies a letter (if an author is associated with several texts in the corpus). Alderd01, thus, indicates that this text is the first letter written by *Alderd*. Finally, the code consisting of four numbers stands for the year in which the letter was written. [↑](#footnote-ref-1)
2. I am indebted and grateful to the project leader who was responsible for the compilation of *The Hamburg Corpus of Irish English*, Peter Siemund, for granting me access and the right to use the HCIE for the present study. [↑](#footnote-ref-2)
3. The table, that Figure 4 is based on, contained missing values for *so* and *other* amplifiers for the time period 1675 to 1750 as all amplified adjectives in this period were amplified by *very*. To compensate these missing values without distorting or misrepresenting the trends shown in Figure 4, the missing values were replaced with values of 0.5. [↑](#footnote-ref-3)
4. The table. that Figure 6 is based on. contained missing values for *so* and *other* amplifiers, amounting to twelve missing values out of 120 possible combinations. Nine of the missing values were present in attributive positions as the data did not contain (1) instances of *so* with *good* or *other* adjectives during the period 1675 to 1750; (2) instances of *other* amplifiers with *good* or *other* adjectives during the period 1675 to 1750; (3) instances of *so* with *bad* in the periods 1751 to 1850 and 1851 to 1930; (4) instances of *so* with *glad* in the period 1751 to 1850; (5) instances of *other* amplifiers with *glad* during the periods 1751 to 1850 and 1851 to 1930. Only three missing values occurred in predicative contexts: the *glad* did neither occur with *so* nor with *other* amplifiers in the period 1751 to 1850 and the data did not contain any instances of *so glad* in the period 1675 to 1750. To compensate these missing values without distorting or misrepresenting the trends shown in Figure 6, the missing values were replaced with values of 0. [↑](#footnote-ref-4)