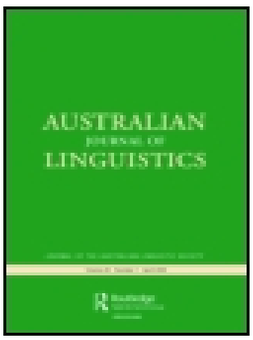

Routledge
Taylor & Francis Group

Australian Journal of Linguistics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/cajl20

Ongoing change in the Australian English amplifier system

Martin Schweinberger

To cite this article: Martin Schweinberger (2021): Ongoing change in the Australian English amplifier system, Australian Journal of Linguistics, DOI: 10.1080/07268602.2021.1931028

To link to this article: https://doi.org/10.1080/07268602.2021.1931028

Published online: 06 Sep 2021.

AUSTRALIAN JOURNAL OF LINGUISTICS
https://doi.org/10.1080/07268602.2021.1931028

Ongoing change in the Australian English amplifier system

Martin Schweinberger

The University of Queensland and The Arctic University of Norway

ABSTRACT

This study takes a corpus-based approach to investigating ongoing change in the Australian English adjective amplifier system based on the Australian component of the *International Corpus of English* (ICE). The paper analyzes changes in amplifiers across apparent time, with special attention being placed on amplifier–adjective–bigram frequencies, to provide insights into cognitive mechanisms underlying lexical replacement. Specifically, the paper analyzes why the innovative variant *really* was successful in replacing the traditional amplifier *very*, while other rivals (e.g. *so* or *pretty*) were not. Lexical diversity scores and distinctive collexeme analyses confirm that, in contrast to other rival variants, *really* specializes on and collocates with a few high-frequency adjectives (HFAs) while being dispreferred by low frequency adjectives. The results of a mixed-effects regression analysis show that the use of *really* is socially stratified with young speakers preferring *really* over other variants. In addition, the multivariate analysis shows that the replacement of *very* by *really* is a female-dominated change and that the use of *really* is enhanced by priming. The paper argues that collocating with HFAs leads to deeper entrenchment which, in turn, serves as an advantage in situations where speakers choose between rivalling innovative variants.

ARTICLE HISTORY
Accepted 11 May 2021

KEYWORDS
Intensification; amplifiers; lexical replacement; Australian English; Boruta

1. Introduction

This study takes a corpus-based approach to investigating ongoing change in the amplifier system of Australian English (AusE) based on the Australian component of the *International Corpus of English* (ICE-Aus) (Peters & Smith, forthcoming). The sentences in (1) show naturally occurring examples of adjective amplifiers taken from the Australian data used in this study.[1] Examples (1a–d) represent uses of adjective amplifiers in predicative contexts while the instances in (1e–g) represent attributive contexts.

(1) a. *It was **so interesting*** (ICE-Aus:S1A-004:1$A[2])
b. *He's **very bright*** (ICE-Aus:S1A-048:1$B)

CONTACT Martin Schweinberger m.schweinberger@uq.edu.au

[1]The examples provided here have been stripped of additional annotation such as symbols indicating overlap or pauses.

[2]In the identifier code 'ICE' stands for *International Corps of English*, 'Aus' for *Australian component*, 'S1A' for *spoken private dialogue*; the three-digit number identifies the file, the number after the second colon shows the subfile, and the final capital letter identifies the speaker. So 'ICE-Aus:S1A-004:1$A' was uttered by speaker A in file 004, subfile 1, in the spoken private dialogue section of the Australian component of the International Corpus of English.

c. *It's* ***pretty bizarre*** (ICE-Aus:S1A-061:1$B)
d. *It was* ***really good*** (ICE-Aus:S1A-070:1$B)
e. *You're a* ***very unusual*** *child* (ICE-Aus:S1A-079:1$B)
f. *I do like* ***pretty weird*** *films* (ICE-Aus:S1A-043:1$B)
g. *But I like him he's he's a* ***really great*** *character* (ICE-Aus:S1A-072:1$B)

The aim of the current study is to provide an overview of the amplifier system of standard AusE as spoken by educated speakers and to add to our understanding of lexical replacement, one of the core mechanisms of language change. While adjective amplification in English has been thoroughly studied (as will be shown in the subsequent section), issues relating to why certain amplifier variants, rather than their rivals, become temporarily dominant remain unanswered. As such, the present paper addresses the following four research questions:

1. Did *really*, as the successful amplifier variant in AusE, broaden or specialize when it increased in use?
2. Did *really* collocate predominantly with high-frequency adjectives (HFAs) in AusE?
3. What correlations between social, cognitive and linguistic factors accompanied the increase in the use of *really*?
4. Do the findings point to a general mechanism which is potentially underpinning the observed processes?

The following elaborates on adjective amplification as a linguistic phenomenon and provides the conceptual framework that motivates the research questions and the interpretation of the statistical results.

Amplification is an intensification strategy, and it is related to the semantic category of degree. Due to the connection between intensification and degree, amplifying intensifiers are also referred to as degree adverbs or adverbs of degree (cf., e.g. Biber et al., 2007, p. 554). The degree of intensity ranges between very low (downtoning) and very high (amplifying) (Quirk et al., 1985, pp. 589–590). According to Quirk et al. (1985, pp. 589–590), amplifiers "scale upwards from an assumed norm [while] downtoners have a lowering effect, usually scaling downwards from an assumed norm" (1985, p. 590). In accordance with other variationist studies (e.g. Tagliamonte, 2008; Tagliamonte & Denis, 2014), the current paper restricts itself to the analysis of amplifiers while leaving aside downtoners (which include approximators such as *almost*, compromisers such as *more or less*, diminishers such as *partly* and minimizers such as *hardly*). Within the category of amplifiers, Quirk et al. (1985, pp. 589–590) differentiate between maximizers such as *completely* which denote the upper extreme of a scale (Quirk et al., 1985, p. 590) and boosters such as *very* which denote a high degree or a high point on a scale. Boosters, in particular, form an open class which adopts new members to replace forms which have lost their expressiveness due to frequent use (cf. Quirk et al., 1985, p. 590).

Amplifier systems are prone to change (Ito & Tagliamonte, 2003, p. 257; Quirk et al., 1985, p. 590) because "new expressions are frequently created to replace older ones whose impact follows the trend of hyperbole in rapidly growing ineffectual" (Quirk et al., 1985, p. 590). The loss of expressivity of traditional forms and the constant need for fresh variants with higher degrees of expressivity causes intensification to be a

domain of "fevered invention" (Bolinger, 1972, p. 18) and an area of waxing and waning that is characterized by continuous invention and renewal (D'Arcy, 2015, p. 450). The perpetual waxing and waning of forms in this domain is particularly intriguing from a language variation and change perspective as their changing nature predestines amplifier systems to be an ideal case for testing mechanisms of change. As such, amplification lends itself to studies which aim to understand factors that determine which forms temporarily come to dominate the amplifier system, which is the topic of this paper.

The present study adds to existing research in that it focuses specifically on the relationship and interdependencies between amplifiers and adjectives. This is relevant because recent research has shown that adjectives play a critical role in facilitating changes in amplifier use (cf. Tagliamonte, 2012, p. 144; Wagner, 2017). In addition, only a few studies have investigated the use of amplifiers in AusE (Sowa, 2009) and the results thus provide insights into the amplifier system of this regionally distinct variety of English.

The next section surveys findings from previous research with a special focus on those relevant for the current topic. Section 3 provides an overview of the data used in the current study and discusses issues relating to methodology, such as data processing and the statistical tools. Section 4 contains the results of the current study including graphs and tables to summarize the findings. Section 5 discusses these results in light of previous research.

2. Previous research on intensification

Intensification has been the subject of much linguistic research for more than a century and various studies have been dedicated to investigating degree modification from functional (e.g. Paradis, 2008) and historical perspectives (cf., e.g. Bolinger, 1972; Breban & Davidse, 2016; Lorenz, 2002; Méndez-Naya, 2003, 2008; Méndez-Naya & Pahta, 2010; Nevalainen, 2008; Nevalainen & Rissanen, 2002; Partington, 1993; Peters, 1992, 1993, 1994; Rissanen, 2008). Due to the substantial amount of research, the developmental pathways of individual intensifiers (e.g. Aijmer, 2011, 2018a; Macaulay, 2006; Pertejo & Martínez, 2014; Rickford et al., 2007; Tao, 2007), their use by native and non-native English speakers (e.g. Fuchs, 2016; Lorenz, 1999; Su, 2016), as well as differences based on the age, gender and social class of speakers are well understood (Bauer & Bauer, 2002; Fuchs, 2017; Ito & Tagliamonte, 2003; Macaulay, 2002; Martínez & Pertejo, 2012; Murphy, 2010; Pertejo & Martínez, 2014; Stenström, 1999; Tagliamonte, 2008; Tagliamonte & Denis, 2014; Tagliamonte & Roberts, 2005). In addition, the distribution of selected intensifiers across registers (e.g. Brown & Tagliamonte, 2012; Fuchs, 2016) and their use within and across varieties of English is well documented (cf., e.g. Aijmer, 2018b; Bauer & Bauer, 2002; Bulgin et al., 2008; Calle-Martín, 2014; de Klerk, 2005; Fuchs, 2016; Paradis, 1997; Sowa, 2009). Furthermore, various fine-grained variationist analyses have provided detailed descriptions of linguistic and social layering (e.g. D'Arcy, 2015; Ito & Tagliamonte, 2003; Tagliamonte, 2008; Tagliamonte & Denis, 2014; Tagliamonte & Roberts, 2005). It has also been shown that amplifiers play a crucial role in how speakers express themselves socially and emotionally (Ito & Tagliamonte, 2003, p. 258; Labov, 1985, p. 43). Thus, intensifiers are part of an inventory on which speakers rely to create and mark their social identity (Tagliamonte, 2012, p. 30).

Various apparent time studies of adjective amplification in geographically distinct varieties of English have found that in most of these varieties *very* is declining while *really* is increasing (cf. D'Arcy (2015) for NZE; Ito and Tagliamonte (2003) and Barnfield and Buchstaller (2010) for North East British English; Tagliamonte (2008) and Tagliamonte and Denis (2014) for Toronto English; and Tagliamonte and Denis (2014) for South Eastern Ontario English). In addition, D'Arcy (2015), Ito and Tagliamonte (2003), Tagliamonte (2008), as well as Tagliamonte and Roberts (2005) show that change in intensifier systems does not proceed in a haphazard manner but that it is highly systematic. The systematicity of the observed changes is reflected by the fact that changes exhibit extra- as well as intra-linguistic stratification.

With respect to intra-linguistic stratification, which represents language internal constraints, the syntactic context or function of the adjectives is among the most consistent factors (cf. Mustanoja, 1960, pp. 326–327; Tagliamonte, 2008, p. 373) as collocation with adjectives in predicative function can be regarded as an indication of a later stage of change (cf. Tagliamonte & Denis, 2014, p. 116). In contrast, initial stages of change are typically associated with overproportionate use of innovative amplifier variants in attributive positions[3] (D'Arcy, 2015, pp. 471–472; Mustanoja, 1960, pp. 326–327). This patterning has been explained by grammaticalization processes and, more specifically, by delexicalization. During the delexicalization process, forms enter syntactic contexts due to a loss of semantic content (semantic bleaching/delexicalization) which these forms were previously barred from entering. One of the best attested clines for delexicalization is the grammaticalization of *very* from an adjective to intensifying adverb. According to Mustanoja (1960, pp. 326–327), the delexicalizaition of *very* proceeded in four steps, as illustrated in (2a–d) below, taken from Tagliamonte (2008, p. 363). Initially, *very* functioned as an adjective meaning 'genuine/true', as in (2a). In a next step, it was used in coordinate constructions with a following attributive adjective, as in (2b). At this stage, the demarcation between adjective and adverb is difficult, if not impossible, as the syntactic context does not allow for an unambiguous interpretation. In a next stage, *very* is used as an intensifier with attributive adjectives, as in (2c). Only in a last stage has *very* entered predicative contexts and its semantic meaning of 'genuine/true' has been entirely lost, leaving solely the intensifying function, as in (2d).

(2) a. *Grant me confort this day, As thow art God* ***verray***! (c.1470, Gol. & Gaw 957; OED *very a., adv. n.1 A.I.1.a)*
b. *He was a* ***verray*** *parfit gentil knyght.* (Chaucer, Canterbury Tales, A Prol. 72)
c. *I was a* ***very interested*** *and anxious spectator.* (1782, R. Cumberland, Anecd. Painters (1787) II. 90; OED very a., adv. n.1 B.2.c)
d. *He was sike … and was* ***verray contrite*** *and sorwful in his herte.* (Trev. Higd. VI 93; cited in Mustanoja, 1960, p. 326)

In analogy to the trajectory of *very*, modification of predicative adjectives has been deemed a sign of advanced delexicalization (Tagliamonte, 2008, p. 373) and it appears

[3]This, however, does not apply to amplifiers that are syntactically restricted such as *so*, which almost exclusively occurs in predicative contexts.

to be a general trend that innovative amplifier variants occur with attributive adjectives first (Mustanoja, 1960, pp. 326–327). Only during later stages do innovative forms occur with adjectives in predicative function (Tagliamonte, 2008, p. 363).

Another factor which accompanies changes in intensifier systems is the association of innovative variants with negative polarity items or emotional adjectives more generally (cf., e.g. Peters, 1994). Partington (1993, p. 184) argues that the reason for this tendency lies in the negative domain from which intensifiers are often recruited (*terrible*, *horrible*, etc.) and that these negative items must first undergo delexicalization before modifying positive polarity items such as *good* or *nice* (cf. also Lorenz, 2002). In other words, during early stages of change, intensifiers that are recruited from negative semantic domains tend to associate with semantically negative adjectives and only during later stages, when the incoming forms have been bleached, do they collocate with positive adjectives as semantic restrictions wane. Furthermore, Tagliamonte and Roberts (2005) found that the innovative intensifier *so* significantly collocates with emotional adjectives (particularly among female speakers) (Tagliamonte & Roberts, 2005, p. 289) as does *really* among speakers between the ages of 20 and 29 in Toronto (Pertejo & Martínez, 2014, p. 230; Tagliamonte, 2008, p. 383). One of the most intriguing aspects that has been described in the respective literature relates to lexical restrictions on collocations. Tagliamonte (2008) found that innovative forms are restricted to a relatively small and fixed set of adjectives and that this set expands once a form becomes more frequent as collocational restrictions erode (cf. Méndez-Naya, 2003, p. 377; Tagliamonte, 2008, p. 376).

The situation is complex when it comes to extra-linguistic stratification. 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 strongly prefer *very* (D'Arcy, 2015; Ito & Tagliamonte, 2003; Tagliamonte, 2008). In contrast, findings for gender preferences in intensification are less coherent. With respect to general gender differences in intensification, Fuchs (2017) found that intensification has increased from the 1990s to the 2010s and that men consistently use intensifiers less frequently than women. Tagliamonte and D'Arcy (2009) also found a significant but weak gender difference which was, however, restricted to the use of the innovative amplifier *so*. D'Arcy (2015, p. 477) found that ongoing change in the amplifier system of NZE did not show significant gender differences. Similarly, Ito and Tagliamonte (2003) as well as Tagliamonte (2008) did not find a consistent effect of gender as gender differences were 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, p. 385). Adding to the complexity is that, in their study of intensification based on the *British National Corpus* (BNC, 2007), Xiao and Tao (2007) found that men prefer maximizers (e.g. *completely*) while women prefer boosters (e.g. *very*).

3. Data

This section describes the corpus data, how they were processed, and the coding of the variables that were included in the statistical analyses.

3.1 ICE Australia

The current study draws on data from the spoken private dialogue section of the Australian component of the *International Corpus of English* (ICE) (Peters & Smith, forthcoming). All ICE components share a common design with the spoken section of any ICE component consisting of 300 files, each encompassing around 2,000 words, for a total of approximately 600,000 words. Of the spoken ICE files, 180 represent dialogues of which 100 files contain private dialogues (90 files with face-to-face conversations and 10 files representing phone calls). For the present analysis, only private dialogues are considered in order to control for potential register effects.

A major advantage that comes with using ICE data is the fact that each component is accompanied by extensive socio-demographic details about the speakers so that the ICE represents a valuable resource for studying variation along the lines of social dimensions.

3.2 Data processing

To extract all adjectives, the corpus data were split into turns based on the annotation present in the Australian ICE data. Then, meta-data such as comments and non-linguistic tags (e.g. <laughter>) were removed. In a next step, information about the file and the speaker were added to each turn. Next, the cleaned turns were part-of-speech tagged in the programming environment R (R Core Team, 2019) 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 transformed into a separate variable. Next, it was determined for each adjective whether it was amplified, and which lexical form served as an amplifier.

3.2.1 Text-related information

Contextual information (audience size and the conversation type) was then added to the data. AudienceSize (the name of the variable) refers to the number of interlocutors that are present in a given dialogue. In this respect, the current study distinguishes between dyads (two interlocutors) and group conversations (multiple interlocutors). ConversationType (again the name of the variable) represents the type of conversation, or more precisely, whether the interlocutors in a given conversation had the same sex (same-sex conversation) or differed in their sex (mixed-sex conversation).

3.2.2 Socio-demographic information

Then, the socio-demographic details of speakers (age, gender, etc.) were added to the data. The age classification of the Australian component of the ICE is very fine-grained but not consistent as, in some cases, age ranges are provided (also overlapping age ranges) while, in other cases, the exact age of speakers is provided. Thus, to retain as much information as possible, the age of speakers was coded in two different ways: one age variable represented age groups (17–25, 26–40, 41–80) while the other age variable represented the exact age of speakers in years. For the latter, if the age of a speaker was provided as a category, e.g. '30–40', the mean age was calculated, i.e. 35. During the statistical analysis, the lexical diversity scores and the covarying collexeme analysis relied on the age groups while the Boruta and mixed-effects models used the numeric age as an

independent variable. Data points for which the age of a speaker was not available were removed from the analysis. This is important as previous research has shown that amplifier use differs across age cohorts (see, for instance, D'Arcy, 2015).

3.2.3 Frequency

Next, the frequency of each adjective by age group was calculated so that changes in adjective use across apparent time could be controlled for. The raw frequency was then transformed into the percentage of uses of a specific adjective within each age group. This percentage value was then logged and scaled to minimize the effect of overly frequent or infrequent adjectives during the regression modelling.

3.2.4 Priming

It was next determined whether the use of an amplifier type may have been the result of potential persistence effects (cf. Tulving & Schacter, 1990, p. 301) – also referred to as production priming (cf. Szmrecsanyi, 2005, p. 113; also Szmrecsanyi, 2006). Priming refers to the re-use of material that was used in previous utterances (cf. Tulving & Schacter, 1990, p. 301). The fact that speakers re-use material is widely acknowledged; however, while there is a growing body of research dedicated to, for example, structural priming which builds on both psycholinguistic experimentation and corpus-linguistic analyses, various issues remain unsolved. For instance, there is much debate concerning the duration of priming effects as the decay time may vary between milliseconds and months or even years (Althaus & Kim, 2006, p. 962). One of the factors determining the durability of priming is the type of priming: semantic or conceptual priming may last for extended periods of time (but not necessarily so), while effects of syntactic, form and production priming typically decay very swiftly (within seconds). The current study assumes that priming is present if the same amplifier is reused in at least one out of the subsequent three pre-adjectival slots. The scope of three adjectival slots is based on the fact that form priming is short-lived and disappears soon after exposure to the stimulus (Althaus & Kim, 2006, p. 962).

3.2.5 Negated and comparative adjectives

After coding for priming, negated adjectives, misclassified items, as well as comparative and superlative forms were removed from the analysis. Furthermore, adjectives that were never amplified, or which were not intensified by at least two different amplifier types, were removed from the analysis to weed out lexicalizations, such as *right honourable*, and tagging errors.

3.2.6 Syntactic context

All adjectives that occurred at the end of utterances or that did not occur before nouns were coded as predicative while adjectives that occurred before nouns or before either adverbs or adjectives and then nouns were coded as being attributive.

3.2.7 Emotionality

Next, the emotionality of adjectives was coded by implementing a sentiment analysis using the syuzhet package in R (Jockers, 2017). The sentiment analysis performed for the current study uses the *Word-Emotion Association Lexicon* (Mohammad & Turney,

2013; cf. http://www.purl.org/net/NRCemotionlexicon), which comprises 10,170 terms, in which lexical elements are assigned scores based on ratings gathered through the crowd-sourced Amazon Mechanical Turk service. For the Word-Emotion Association Lexicon raters were asked whether a given word was associated with one of eight emotions according to their judgement. The concept of emotion in the present study rests on Plutchik (1980, 1994) who proposes eight basic emotions (ANGER, ANTICIPATION, DISGUST, FEAR, JOY, SADNESS, SURPRISE, TRUST). The resulting associations between terms and emotions are based on 38,726 ratings from 2,216 raters who answered a sequence of questions for each word which were then fed into the emotion association rating (cf. Mohammad & Turney, 2013). Each term was rated five times. For 85% of words, at least four raters provided identical ratings. For instance, the words *dark* or *tragic* are more readily associated with SADNESS while words such as *happy* or *beautiful* are indicative of JOY and words like *cruel* or *outraged* may indicate ANGER. If an adjective was associated with ANGER, DISGUST, FEAR or SADNESS, the adjective was categorized as NegativeEmotional. In contrast, if an adjective was associated with ANTICIPATION, JOY, SURPRISE or TRUST, the adjective was categorized as PositiveEmotional. If a word was not associated with any emotion, it was categorized as NonEmotional.

3.2.8 Gradability

Each adjective was assigned a gradability score. Gradability represents a semantic property of adjectives (cf. Quirk et al., 1985) which encodes the degree of a characteristic (Biber et al., 2007, p. 521) and thus refers to a more or less compared to a baseline. In contrast to non- or ungradable adjectives, gradable adjectives can occur in contexts that denote comparative or superlative states. In such contexts, gradable adjectives either take morphological marking with <-er> or <-est> or periphrastic marking with 'more' or 'most' while the base forms that do not occur in comparative contexts do not have any special marking. Some adjectives rarely form a comparative or superlative as, for example, **more dental* or **? most motionless*. While amplifiers are typically not used with non- or ungradable adjectives, such adjectives can be amplified for pragmatic purposes such as emphasis. This suggests that gradability reflects a quantitative rather than a qualitative property of adjectives. To accommodate the fact that gradability is not categorical, gradability is operationalized as the logit of the probability of an adjective being used in a comparative context versus the overall rate of comparison among adjectives. Since the data used in the present study are inadequate to warrant such an operationalization, the gradability score was calculated based on the *Corpus of Contemporary American English* (COCA) (Davies, 2010). In the present study, a gradability score of zero represents a neutral state where an adjective occurs with an average rate in comparative contexts while negative values indicate a tendency towards non-gradability (the adjective occurs less often in comparative contexts than the average adjective) and positive scores indicate a tendency towards gradability (the adjective occurs more frequently in comparative contexts than the average adjective).

3.2.9 Semantic classification

The semantics of all remaining adjectives were annotated based on the classification provided by Dixon (1977, 2004; see also D'Arcy, 2015; Tagliamonte, 2006, 2008; Tagliamonte & Roberts, 2005). Thus, all adjectives were coded as belonging to one of the following

Table 1 Absolute frequencies and percentages of amplifiers in the private dialogue section of ICE Australia

Amplifier	Frequency (*N*)	Percent (all)	Percent (amplifiers)
Ø (not amplified)	1,914	76.7	
Really	244	9.8	41.9
Very	125	5.0	21.5
So	116	4.6	19.9
Pretty	53	2.1	9.1
Bloody	9	0.4	1.5
Absolutely, totally	7	0.3 (0.6)	1.2 (2.4)
Completely	5	0.2	0.9
Extremely, particularly	3	0.1 (0.2)	0.5 (1.0)
Terribly, true	2	0.1 (0.2)	0.3 (0.6)
Actually, awfully, genuinely, incredibly, real, strongly	1	0.1 (0.6)	0.2 (1.2)
Total	2,496 (582)	100 (23.3)	100

semantic groups: *Age* (e.g. *old* or *young*), *Colour* (e.g. *white* or *green*), *Difficulty* (e.g. *easy* or *difficult*), *Dimension* (e.g. *high* or *left*), *Human Propensity* (e.g. *intelligent* or *stubborn*), *Physical Property* (e.g. *hard* or *soft*) and *Value* (e.g. *good* or *bad*). Adjectives which could not be categorized (e.g. *familiar*, *genuine* or *inadequate*) were assigned the label '*other*'.

Then, the data were manually cross-evaluated and checked to minimize erroneous annotation. As a result, the data processing produced a pre-final data set, an overview of which is provided in Table 1.

Table 1 shows that only a few amplifier types are responsible for the vast majority of overall amplification: the most frequent amplifier in the private dialogue section of ICE Australia is *really* (244), followed by *very* (125), *so* (116) and *pretty* (53). Combined, these four types are used in 92.4% of cases, leaving the remaining amplifiers a mere 7.6% of amplification.

3.2.10 Removing all non-amplified adjectives

Before we can turn to the final data set, there is an important caveat to acknowledge. Since the linguistic variable can be defined as a situation in which "the speaker reaches a decision-point" (Wallenberg, 2013; cited in Maddeaux & Dinkin, 2017), the variable context in the present study is the decision of which amplifier to use once the speaker has already decided to amplify an adjective. The variable context thus encompasses only amplified adjectives while leaving out zero contexts, i.e. contexts where the speaker could have amplified an adjective but did not. This means that all instances of adjective types that are not amplified are removed from further analysis in order to only focus on contexts that are amplified and thus represent a variable context.

3.2.11 Multiple variants versus binary choice

Finally, for the lexical diversity scores and the covarying collexeme analysis, all amplifiers except for *really*, *very*, *so* and *pretty* were collapsed into the bin category *other* while, for the Boruta enhanced mixed-effects regression modelling, all variants except for *really* were collapsed – leaving only a binary distinction between *really* and *other*.

3.3 Overview of the final data set and the variables

After the data processing, the data were annotated for the linguistic, psychological and social variables shown in Table 2.

Table 2 Overview of annotated variables with variable levels and description in the final data set

Variable	Scaling	Levels	Description
Dependent variables			
Variant (LD, CCLA)	Categorical	*Other*, *pretty*, *really*, *so*, *very*	Frequent variant types
really (Boruta, GLMM)	Nominal	0,1	*Really* vs. *other* amplifier
Independent variables			
Age (LD, CCLA)	Categorical	17–25, 26–40, 41–80	Age groups
Age (Boruta, GLMM)	Numeric	Range: 17, 70; median: 21, SD: 10.7	Age in years
Gender	Nominal	Man, Woman	Self-defined gender
AudienceSize	Nominal	Dyad, group	Size of audience
ConversationType	Nominal	MixedGender vs. SameGender	Type of conversation
FileSpeaker	Nominal	Speaker id	Individual speaker
Frequency	Numeric	Range: −0.77, 2.09; median: 0.0; SD: 1.0	Scaled and logged percentage of adjective types by age group
Function	Nominal	Attributive vs. Predicative	Syntactic context of adjective
Emotionality	Nominal	Positive vs. Neutral vs. Negative	Emotionality of the adjective
Gradability	Numeric	Range: −3.6, 2.1; median: −1.1; SD: 1.8	Likelihood of adjective type occurring in a comparative context
SemanticCategory	Categorical	Age, Colour, Difficulty, Dimension, HumanPropensity, Other, PhysicalProperty, Value	Semantic classification of adjectives
Adjective	Categorical	Adjective type	Adjective type in final data set
Priming	Nominal	Primed vs. NotPrimed	A specific amplifier has occurred within three preceding adj. slots

An overview of the final data set with respect to the frequency of speakers, adjectives, instances of *really*, and the percentages of *really* by age and gender in ICE-Aus is shown in Table 3.

The final data set comprises 167 speakers, 582 amplified adjectives and 244 instances of *really*, which amounts to 41.9% of all amplification. Table 3 also shows that young speakers between the ages of 17 and 25 and young women, in particular, are notably overrepresented in the data while speakers older than 40 are substantially underrepresented. The following section focuses on methodology and introduces the statistical procedures applied in this study.

3.4 Methodology

This study makes use of three types of analyses:

1. Lexical diversity scores to address research question 1 (Did *really*, as the successful amplifier variant in AusE, broaden or specialize when it increased in use?);
2. Covarying collexeme analysis to address research question 2 (Did *really* collocate predominantly with high-frequency adjectives (HFAs) in AusE?); and
3. Mixed-effects binomial logistic regression (with a Boruta analysis serving as variable selection procedure) to address research question 3 (What correlations between social, cognitive and linguistic factors accompanied the increase in the use of *really*?).

Both the lexical diversity scores (LD) and the covarying collexeme analysis tap into changes in the collocational profile of amplifiers. LD scores and collostructional analysis are complementary because the former tap into the number of adjective types that

Table 3 Number of speakers by age and gender as well as absolute frequencies and percentages of amplifiable adjectives and really in the processed private dialogue section of ICE Australia

Age	Gender	Speakers (*N*)	Adjectives (*N*)	*Really* (*N*)	*Really* (%)
17–25	Man	27	79	25	31.6
17–25	Woman	79	325	160	49.2
26–40	Man	15	65	19	29.2
26–40	Woman	22	56	26	46.4
41–80	Man	8	14	4	28.6
41–80	Woman	16	43	10	23.3
Total		167	582	244	41.9

amplifiers co-occur with while the latter focuses on the collocation strength between individual amplifiers and adjectives. Also, only the collostructional analysis is able to control for frequency effects that are likely to confound and skew LD scores. The mixed-effects regression analysis tests the correlations between the use of the innovative variant *really* and the language-internal and -external variables discussed above. The Boruta analysis serves as a variable selection procedure used to streamline the model fitting process of the regression analysis.

3.4.1 Lexical diversity

To analyze changes in the collocational profile of amplifiers, the current study uses lexical diversity scores (LD). The LD score is calculated by dividing the number of adjective types a given amplifier co-occurs with by the number of tokens of that amplifier:

(3) $\text{LD} = N_{\text{Adj.Types}}/N_{\text{Amp. Tokens}}$

The maximum score of LD is 1 which indicates high lexical diversity. The lower the LD score, the lower the degree of lexical diversity (cf. Table 4).

As Table 4 shows, *very* has a lower LD score compared with *pretty* because there are 125 tokens of *very* that co-occur with 48 adjective types while there are only 53 tokens of *pretty* which also co-occur with 33 adjective types. The LD score allows testing of whether the lexical diversity of an amplifier variant correlates with its use. If variants became successful via broadening, then a successful variant should have a relatively high LD score. In contrast, if successful variants specialize on a few high-frequency adjectives, then this would predict that successful variants should have a comparatively low LD score and an increase in token frequency should be accompanied by a drop in the variant's LD score. It should be noted though that the LD scores are substantively affected and skewed by frequency effects as amplifiers that do not occur frequently are likely to have high LD scores. Hence, the LD scores should be treated as indications of trends and interpreted with care.

Table 4 Table exemplifying the calculation of LD

Amplifier	Amp. Tokens (*N*)	Adj. Types (*N*)	Calculation	LD score
Variant_A	10	1	1/10	0.1
Variant_B	10	5	5/10	0.5
Variant_C	10	10	10/10	1
Very	125	48	48/125	0.384
Pretty	53	33	33/53	0.623
Extremely	3	2	2/3	0.667

3.4.2 Covarying collexeme analysis

Covarying collexeme analysis is part of the collostructional family of analyses (Gries & Stefanowitsch, 2004; Stefanowitsch & Gries, 2003, 2005; cf. also Hilpert, 2006). Covarying collexeme analyses evaluate the attraction between elements that occur in two distinct slots within a specified construction. In the present case, the first slot is the amplifier slot and the second slot is the adjective slot in the amplifier–adjective construction. Each slot can be occupied by a variant from a set of potential candidates and the covarying collexeme analysis determines if the use of an amplifier in the first slot affects the likelihood of an adjective in the second slot. In other words, the analysis tests if it is more likely that, e.g. *nice* occurs in the second slot given that, e.g. *really* occurs in the first slot compared with another amplifier in the first slot. If the likelihood increases, then the forms attract each other while they repel each other if the likelihood decreases.

The covarying collexeme analyses are applied within age groups to test for changes in apparent time. The effect size measure reported here is φ (lower case phi). As such, attraction is indicated by positive values while negative values indicate that the amplifier and the adjective occur less frequently together than would be expected by chance. The advantage of covarying collexeme analyses over other methods that evaluate collocational attraction is that it is very robust as it does not rely on distributional assumptions like tests from the χ^2-family of tests.

3.4.3 Mixed-effects binomial logistic regression modelling

The study uses a type of multivariate analysis (mixed-effects binomial logistic regression) to statistically test if any of the independent variables or interactions between them correlate with the use of *really* in the ICE-Aus data.

To streamline the step-wise step-up model fitting procedure (cf. Field et al., 2012; Gries, 2009; independent variables and their interactions are added consecutively, i.e. the model is built up), the present study uses Boruta. Boruta (Kursa & Rudnicki, 2010, 2018) is a variable selection procedure and it represents an extension of random forest analyses (cf. Breiman, 2001; Tagliamonte & Baayen, 2012). Boruta uses a distributional approach during which hundreds of (random) forests are grown from permutated data sets. The Boruta analysis then assesses the performance of variables by comparing them against a control, so-called shadow variables.

As Boruta is a variable selection procedure, it is limited in the sense that it provides information on which predictors to include and how good these predictors are (compared to the shadow variables) but it is neither able to take hierarchical data structure into account, nor does it provide information about how one level of a factor compares to other levels. In order to retrieve this information, additional mixed-effects regression modelling is used.

Thus, if the Boruta confirmed variables as significant, these variables were used in the regression analysis which then used model fitting based on these variables to arrive at the final minimal adequate model, i.e. the best model in the sense that a minimum of predictors explains a maximum of variation. After each addition of a predictor during the model fitting, the ANOVA-function (with both models as arguments) was used to test if the inclusion of a predictor was justified.[4] In addition to including main effects, the statistical

[4]The predictor was retained (i) if the variance inflation factors were acceptable, i.e. having a value smaller than or equal to 3 (Zuur et al., 2010); (ii) if the more saturated model had a lower BIC value compared with the smaller model; and (iii) if the ANOVA reported a significant improvement of model fit.

analyses tested all secondary or two-way interactions (interaction between two main effects) as well as tertiary or three-way interactions (interaction between three main effects). Interactions were not included in cases where this would have led to complete separation, cases of incomplete information (cf. Field et al., 2012, pp. 322–323), failures to converge or unacceptable (multi-)collinearity (variance inflation factors greater than 5). The step-wise step-up model fitting arrived at a final minimal adequate model and the significant steps are displayed in Table A1 in the Appendix.

As a final note on methodology, at least one issue – the relatively moderate size of the data sets used in the current study – requires additional attention. To ascertain if the sample size of the data is sufficient to derive meaningful conclusions, additional power analyses were performed using the SIMR package in R (Green & MacLeod, 2016). Power analyses are used to test if the sample size of a study is sufficient to detect an effect with a certain accuracy. Sample sizes are deemed sufficient, if the statistical models that are fit to the data detect a medium sized effect with 80% accuracy based on bootstrapped samples of the data (Green & MacLeod, 2016). The results of the power analyses showed that the data were sufficient for all main effects and two-way interactions (with the exception of the interaction between Gender and Priming which only achieved an accuracy of 55% instead of the required 80%). However, the sample size was not sufficient to detect three-way interactions with a satisfactory accuracy if the interactions had the weakest medium effect size. The results of the power analyses are also included in Table A1 in the Appendix.

4. Results

To ascertain the trajectory of change in AusE, the percentage of amplifier types are tabulated and plotted against speaker age to assess potential ongoing apparent time change (see Table 5 and Figure 1). Showing the apparent time distribution also allows us to compare the trajectory of change in the Australian ICE data with findings of previous research on the NZE amplifier system. D'Arcy (2015) showed that *really* has replaced *very* in terms of frequency in NZE and that *pretty* has experienced a substantial increase in use in NZE since the 1970s. The apparent time trajectories in AusE mirror this finding (Table 5, Figure 1).

Table 5 shows that in both syntactic contexts, *very* and *other* amplifier variants decreased across apparent time while *really* increased notably – as did *pretty*, although to a lesser degree. In predicative contexts, *so* also increased substantially, thus mirroring the increase of *really*. These trends are visualized in Figure 1.

Figure 1 shows that, across apparent time, *really*, as an innovative upcoming variant, has replaced the traditional variant *very* as the dominant form in both syntactic contexts. The replacement of *very* by *really* appears to have happened in predicative contexts first as *really* had already replaced *very* among speakers between the ages of 26 and 40 while in that same age bracket *very* was still dominant in attributive contexts. In predicative contexts, *so* is the second most frequently used amplifier among speakers between the ages of 17 and 25 showing that it is another upcoming variant. In contrast to changes in the amplifier system of NZE, *pretty* does not appear to be a serious contester for *really*, although the data show a notable increase in use among younger speakers. Thus, the data taken from the private dialogue section of the Australian component of the ICE show that *really* has successfully replaced *very* as the dominant amplifier in standard AusE.

Table 5 Tabulation of percentages of amplifier variants in AusE across apparent time

Variant	Function	41–80	26–40	17–25
Other	Attributive	18.2	33.3	4.0
Pretty	Attributive	0	0	12.0
Really	Attributive	18.2	20	61.3
So	Attributive	0	0	0.0
Very	Attributive	63.6	46.7	22.7
Other	Predicative	10.9	9.9	4.6
Pretty	Predicative	2.2	8.8	10.7
Really	Predicative	26.1	42.9	42.6
So	Predicative	13	14.3	28.8
Very	Predicative	47.8	24.2	13.2

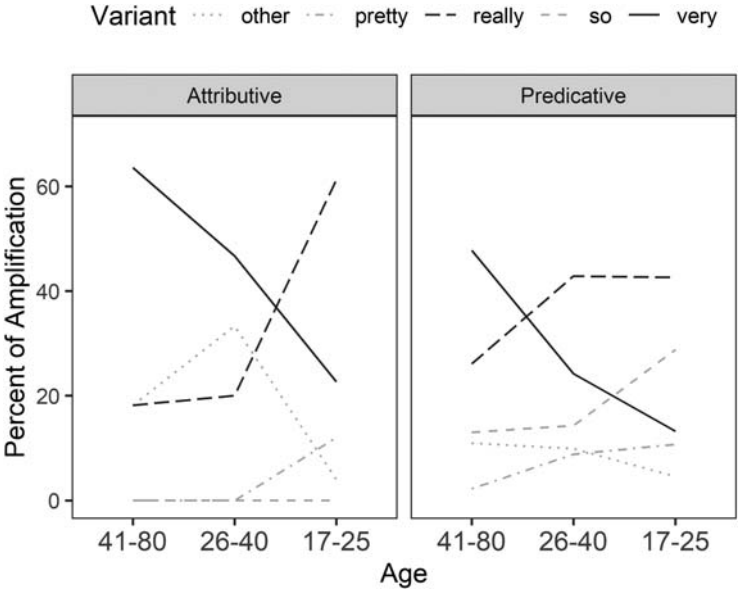

Figure 1 Trajectories of amplifier variants in AusE across apparent time

4.1 Lexical diversity results

The LD scores that have been calculated for *pretty*, *really*, *so*, *very* and *other* amplifiers across apparent time are shown in Table 6 and Figure 2.

Both Table 6 and Figure 2 show that all variants exhibit a drop in LD scores among younger speakers. The drop in lexical diversity by *other*, *pretty*, *really* and *so* is almost parallel. However, only *really* and *other* amplifiers show a consistent trend across age groups with *really* showing the most dramatic and substantial drop in lexical diversity. *Very*, in contrast, only shows a moderate decline in the lexical diversity after an initial increase among middle-aged speakers. Among speakers aged between 17 and 25, *really* has the lowest lexical diversity. It needs to be noted though that the results for the oldest speakers in the data are less reliable given the moderate frequency of amplifiers used by these speakers (see Table 6). Also, it has to be borne in mind that lower frequencies lead to skewed LD scores as lower frequencies are likely to result in higher LD scores. Thus, the high values among middle-aged and older speakers have to be interpreted with caution.

Table 6 Lexical diversity scores in AusE across apparent time

Age	Variant	Adjective types (*N*)	Variant frequency (*N*)	LD score
17–25	*Other*	11	18	0.6
17–25	*Pretty*	28	44	0.6
17–25	*Really*	50	185	0.3
17–25	*So*	47	97	0.5
17–25	*Very*	27	60	0.4
26–40	*Other*	14	19	0.7
26–40	*Pretty*	8	8	1.0
26–40	*Really*	27	45	0.6
26–40	*So*	12	13	0.9
26–40	*Very*	21	36	0.6
41–80	*Other*	6	7	0.9
41–80	*Pretty*	1	1	1.0
41–80	*Really*	10	14	0.7
41–80	*So*	5	6	0.8
41–80	*Very*	15	29	0.5

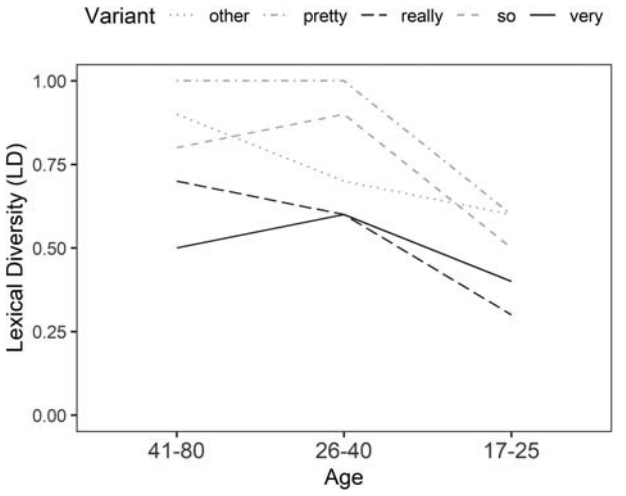

Figure 2 Apparent time distribution of the lexical diversity scores of amplifiers in standard AusE

Nonetheless, what both Table 6 and Figure 2 suggest is that the lexical diversity scores decrease as the frequencies of amplifiers increase with dominant forms (*very* for old speakers and *really* for young speakers) having the lowest scores. We now turn to the results of the distinctive collexeme analyses to probe deeper into changes in the collocation patterns of amplifiers.

4.2 Results of the collostructional analysis

The covarying collexeme analysis determines whether certain amplifier types prefer (Type) or reject (Antitype) certain adjectives. Statistically significant results of this analysis are provided in Table 7.

Table 7 shows that all significant results are confined to the youngest age group. Only three results were still significant once the *p*-values were adjusted for multiple, repeated testing (significance after Benjamini–Hochberg correction). The most interesting of these

Table 7 Results of the covarying collexeme analysis of amplifiers and adjectives in AusE

Age	Variant (*N*)	Adjective (*N*)	Uncorrected *p*-value	X^2	φ (phi)	expected	Significance$_{\text{Benjamini–Hochberg corrected}}$	Type
17–25	*Other* (212)	*Other* (227)	0.0261	5.6	0.12	216.9	n.s.	Antitype
17–25	*Really* (53)	*Good* (84)	0.0005	12.8	0.18	38.5	$p < .01$	Type
17–25	*Really* (83)	*Other* (227)	0.0000	17.8	0.21	103.9	$p < .001$	Antitype
17–25	*So* (10)	*Funny* (24)	0.0478	4.4	0.1	5.8	n.s.	Type
17–25	*So* (7)	*Good* (84)	0.0001	14.3	0.19	20.2	$p < .01$	Antitype
17–25	*So* (63)	*Other* (227)	0.0471	4	0.1	54.5	n.s.	Type
17–25	*Very* (0)	*Funny* (24)	0.0343	4.5	0.1	3.6	n.s.	Antitype

corrected findings are that *really* occurred significantly more often with *good* and significantly less often with *other* adjectives than would be expected by chance. In addition, *so* co-occurred significantly less frequently with *good* than would be expected. This shows that *really* and *good* (the most frequent adjectives in the AusE data) significantly collocate which indicates specialization with a focus on high-frequency adjectives on the part of *really*.

The apparent time trajectories of collocation strengths of variants and adjectives that had effect sizes of above 0.1 or below −0.1 are visualized in Figure 3.

Figure 3 shows that the attraction between *really*, and *good*, *nice* and *funny* increased across apparent time. This is relevant as *good*, *nice* and *funny* are the three most frequent adjectives in the AusE data (*good* 118, *nice* 49, *funny* 27).

In contrast, *so* and *good* occurred less frequently together than would be expected by chance indicating that *so* is repelled by *good* among young speakers. Also, *really* is increasingly repelled by adjectives other than *good*, *nice* and *funny* across apparent time which substantiates that *really* is specializing rather than broadening during its increase in use. Interestingly, the opposite trend, broadening, appears to be the case

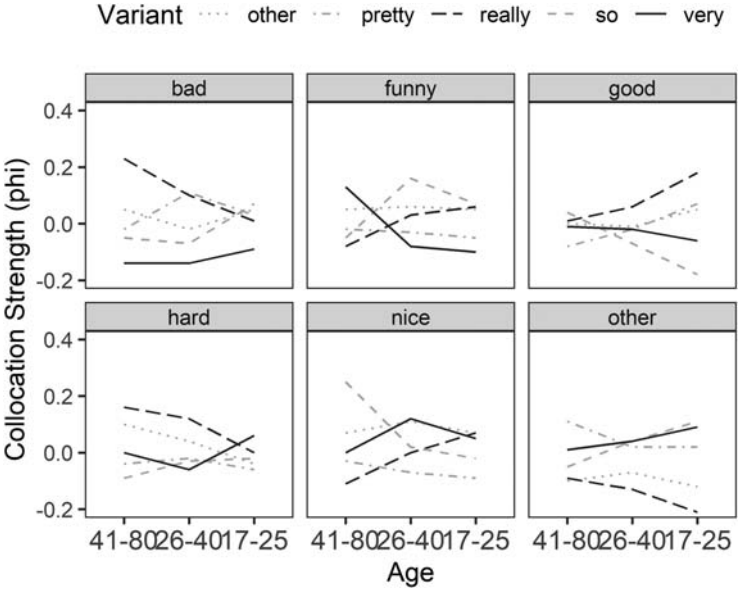

Figure 3 Apparent time changes in attraction and repulsion of amplifiers and adjectives based on covarying collexeme analyses

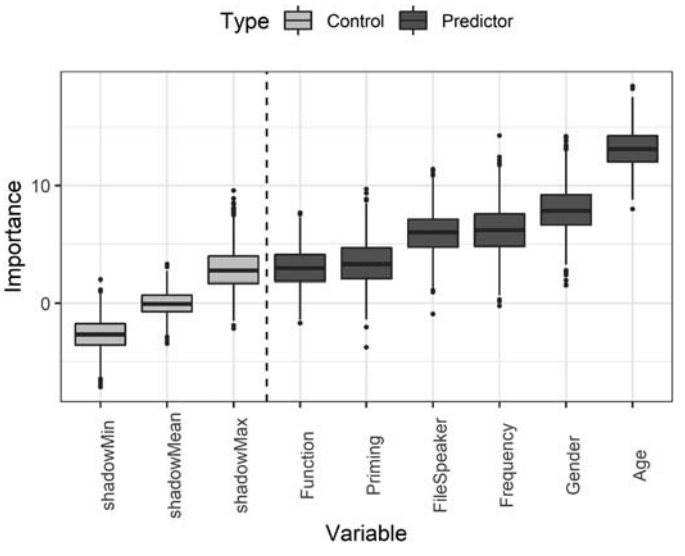

Figure 4 Results of the Boruta analysis

for *very* while it is on its way out (see the increase in collocation strength with other adjectives in the lower right panel). We will now turn to the results of the Boruta variable selection procedure and the regression analysis in order to ascertain which factors correlate with the use of *really* during its rise to dominance.

4.3 Boruta and mixed-effect regression results

To streamline the regression modelling, the analysis used a Boruta variable selection procedure to determine which variables had any non-random relationship with the dependent variable (use of *really* versus any *other* amplifier). The results of the Boruta analysis which showed which variables have any impact on the use of *really* are displayed in Figure 4.

The results show that the age (in years) and the gender of speakers, as well as the frequency of adjectives by age group, and priming, are relevant predictors for the use of *really*. The absence of other predictors also shows that neither audience size or conversation type, nor the emotionality, gradability or semantic category of adjectives significantly correlated with the use of *really*. To evaluate the direction of the effects of the variables and to determine whether the variables impact the use of *really* as main effects or as part of interactions, the analysis used mixed-effects binomial logistic regression modelling with a step-wise step-up model fitting procedure. The results of the final minimal adequate model which contained varying intercepts for speakers (FileSpeaker)[5] are presented in Table 8.

[5]It should be noted that most of the speakers are associated with only one observation and concerns have been raised that it may therefore be problematic to include varying intercepts for speakers. There has indeed been an ongoing discussion about the minimum number of observations per random effect level. The literature on this issue suggests, however, that, given that the number of levels (speakers) is sufficiently large, even large proportions of levels with only a single observation do not lead to overdispersion or inaccurate estimates (see Bell et al., 2008; Clarke, 2008; Clarke & Wheaton, 2007; Maas & Hox, 2005).

Table 8 Results of the final minimal adequate mixed-effects binomial logistic regression model

	Intercept only baseline model			Final minimal adequate model		
Predictors	Odds ratios	CI	*p*	Odds ratios	CI	*p*
(Intercept)	0.64	0.50–0.81	<0.001***	0.80	0.39–1.66	0.601
Age				0.97	0.95–0.99	0.004**
GenderWoman				1.96	1.17–3.28	0.009**
Frequency				1.47	1.21–1.79	<0.001***
PrimingPrimed				1.88	1.17–3.04	0.016*
Random effects						
σ^2	3.29					3.29
τ_{00}	$0.72_{FileSpeaker}$					$0.54_{FileSpeaker}$
ICC	0.18					0.14
N	$167_{FileSpeaker}$					$167_{FileSpeaker}$
Observations	582					582
Marginal R^2	0.000					0.105
Conditional R^2	0.179					0.231
C						0.8
Somers' D_{xy}						0.6
AIC	776.96					748.473
BIC	785.69					772.66
Likelihood ratio test				X^2_{DF5}: 57.11	$p < 0.001$***	

The results of the regression model in Table 8 confirm that speaker (FileSpeaker; varying intercepts), the age (in years) and the gender of speakers, as well as the frequency of adjectives by age group, and priming, significantly correlate with the use of *really* as an adjective amplifier in standard AusE based on the ICE-Aus data. The final minimal model performed significantly better than an intercept-only baseline model and has excellent model fit parameters which means that it possesses real explanatory and predictive capacity as indicated by the high C, Somers' D_{xy} and R^2 values (see Baayen, 2008, p. 204). The power analysis confirmed that the sample was sufficient to detect the weakest medium effect for all main effects as well as two-way interactions with at least 80% accuracy[6] which is the standard criterion for adequately sized samples in clinical trials (see Green & MacLeod, 2016, p. 495). However, the sample proved to be insufficient to detect three-way interactions (see Table A1).

The results of the multivariate regression modelling confirm that the probability of using *really* rather than another adjective amplifier decreases with age (Figure 5) and correlates positively with adjective frequency (Figure 6). The latter substantiates the results of the covarying collexeme analysis in that *really* associates with high-frequency adjectives.

In addition, women are more likely to amplify adjectives with *really* compared to their male peers regardless of age which suggests that the replacement of *very* by *really* is a female-dominated change (Figure 7). In addition, the use of *really* is facilitated by priming because the use of *really* is significantly more likely if *really* has been used in the preceding discourse (Figure 8).

We will now interpret these results with respect to the research questions that this paper aimed to answer and in light of relevant research literature.

[6]The interaction between Gender and Priming being the only exception with a power of 55%.

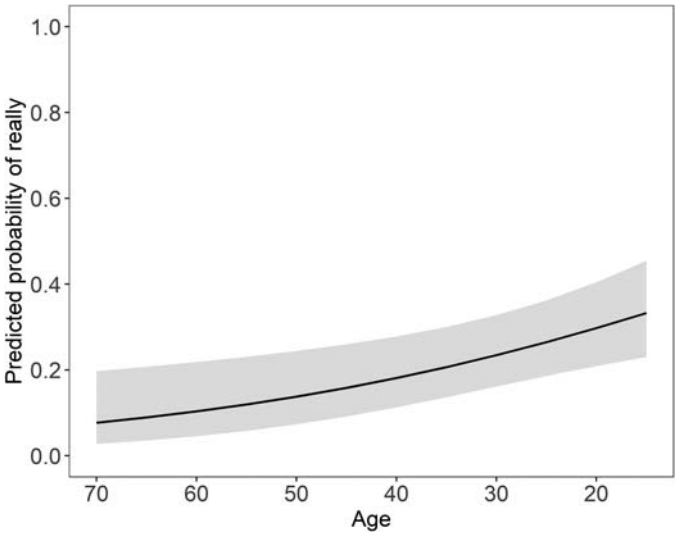

Figure 5 Probability of *really* by age of speaker as predicted by the final minimal adequate model

5. Discussion

The current analysis of the amplifier system in standard AusE based on the Australian ICE data has unearthed intriguing interdependencies between amplifier use and adjective frequency. The analysis shows a steady decline in use of *very* as the dominant amplifier in this reginal variety. The analysis confirms that the frequency of the host-adjective is a determining factor in the replacement of *very* by *really* and that this replacement process can be characterized as a situation of ongoing struggle for dominance and rivalry among competing variants (see Figure 1). Also, the results highlight that in

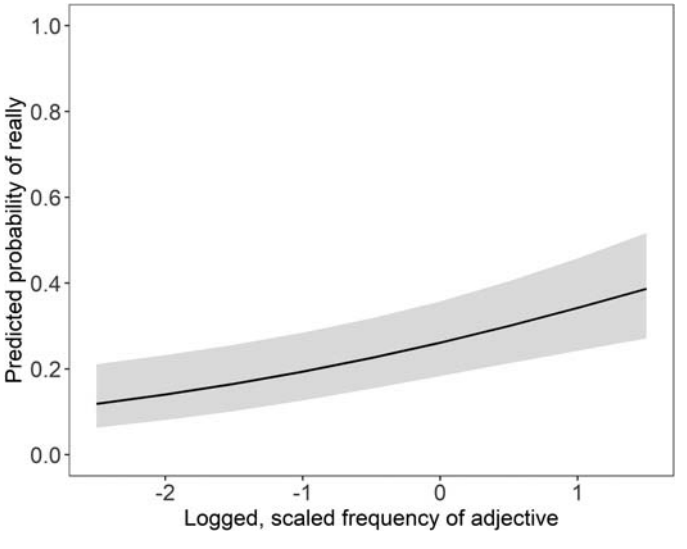

Figure 6 Probability of *really* by adjective frequency as predicted by the final minimal adequate model

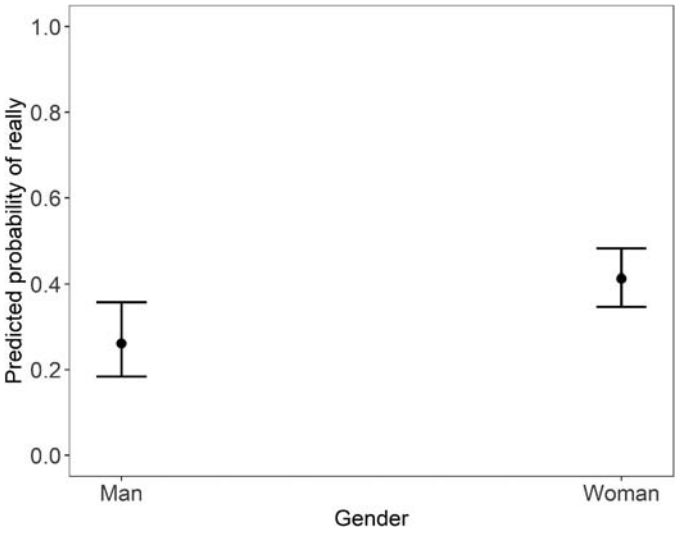

Figure 7 Probability of *really* by gender as predicted by the final minimal adequate model

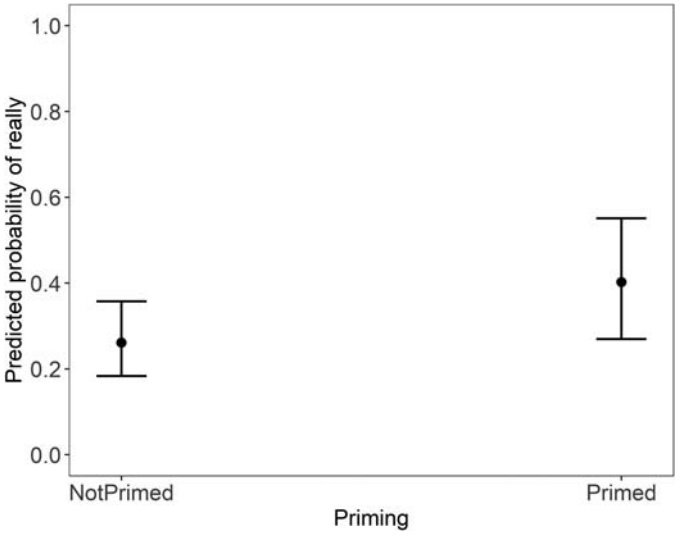

Figure 8 Probability of *really* by priming as predicted by the final minimal adequate model

order to understand changes in amplifier systems one has to take adjectives and their collocational preferences into account (cf. also Wagner, 2017, in support of this argument).

The traditionally dominant variant *very* is being replaced by *really* in both attributive and predicative syntactic contexts which aligns with previous research that has shown that *really* is replacing *very* in Toronto English (Tagliamonte, 2008; Tagliamonte & Denis, 2014), South Eastern Ontario (Tagliamonte & Denis, 2014), North East British English (Barnfield & Buchstaller, 2010; Ito & Tagliamonte, 2003) and New Zealand English (D'Arcy, 2015). As such, the trajectory of change in the AusE adjective amplifier system represents a global or supra-varietal trend rather than a local or regionalized

phenomenon. We will now turn to the research questions that this paper aimed to address.

The first research question, which asked if *really* broadened or specialized when it increased in use, can be answered based on the apparent time trajectories of the lexical diversity scores and the results of the covarying collexeme analysis: the results consistently show that the replacement of *very* by *really* is accompanied by specialization of the waxing variant (*really*) and broadening of the waning form (*very*). The lexical diversity scores showed a steady decrease for *really* while no clear trend emerged for *very*. The covarying collexeme analysis confirmed that *really* significantly collocated with *good* while it was repelled by *other*, infrequent adjectives. This suggests that waxing variants specialize, rather than broaden, while they become dominant.

The second research question, which asked if *really* collocates predominantly with high-frequency adjectives, was also confirmed by the results of the present analysis. The covarying collexeme analysis showed that the attraction between *really* and the three most frequent adjectives in the ICE-Aus data (*good*, *nice* and *funny*) steadily increased across apparent time while its collocation strength with *other*, infrequent adjectives decreased. This indicates that waxing variants may specialize on a small set of high-frequency adjectives when becoming dominant. However, cross-linguistic research looking at a variety of replacement phenomena would be required to ascertain if this correlation that we have observed for *really* in AusE is representative of a general mechanism or merely a localized correlation.

The third research question, which asked what correlations between social, cognitive and linguistic factors accompany the increase in use of *really*, was addressed by the regression analysis. The regression analysis shows that the replacement of *very* by *really* is socially stratified and a female-dominated change. Both the stratification by age and the female-led change align with common trends that have emerged in variationist research. According to Labov (2002), the majority of changes that have been studied in the twentieth century are female led (see also Labov, 2001, 2010) and differences in apparent time have been shown to reliably reflect real-time trajectories of change (Bailey et al., 1991).

The impact of priming on the use of *really* is, while expected, interesting from a theoretical perspective. So far, only a few studies of ongoing change have considered priming as a factor – notable exceptions are Estival (1985), Gries (2005, 2013), Hilpert (2013), Poplack (1980), Poplack and Tagliamonte (1993, 1996), Schweinberger (2018, 2020a), Szmrecsanyi (2006), and Weiner and Labov (1983). As such, the present analysis lends support to the hypothesis that "priming, which strongly influences language processing, may also play an important role in language change" (Pickering & Garrod, 2017, p. 190).

The frequency effect that is confirmed by the regression analysis also aligns with previous research that has drawn attention to the role that frequency plays in language change (see, e.g. Bybee, 2007, 2010, 2015). However, the present study adds to this research in a novel way as it confirms that it is not only the frequency of the form itself that is predictive but also the frequency of the host construction (the adjective). In the present data, predominantly *good*, *nice* and *funny* are responsible for the drastic increase of *really* among post-adolescent speakers (see Figure 3). The results presented here thus substantiate the hypothesis that variants that become dominant initially specialize on a few high-frequency adjectives and expand their collocational profile only after

experiencing a substantial increase in frequency. In fact, it appears likely that it is precisely this specialization on a few high-frequency adjectives that sets *really* apart from other rivalling variants. This finding, while predicted by the hypothesis that specialization is a precondition for successful lexical replacement (Schweinberger, 2020b), is intriguing as it could be indicative of a more universal mechanism of how frequency affects language change – a mechanism that has so far escaped thorough analysis.

What is interesting about the finding that the use of *really* is accompanied by specialization is that it contrasts with findings from very similar research that has analyzed amplifier use in Canadian English: Tagliamonte (2008), as well as Ito and Tagliamonte (2003, p. 276), who found that *really* expanded first and only subsequently increased in frequency. In contrast, the present analysis of adjective amplification in AusE suggests that the increase in use precedes broadening. A likely explanation for the different findings stems from the fact that the type of expansion investigated here is different compared with the studies mentioned above. In the present study, expansion refers to the mere number of adjective types that a variant co-occurs with. In Tagliamonte (2008) as well as in Ito and Tagliamonte (2003), on the other hand, expansion refers to the ability of amplifiers to co-occur with adjectives that belong to distinct semantic classes. Since the current analysis has included semantic categories of adjective types, it would have been expected that the investigation would have produced a similar result. This is however not the case – the Boruta analysis rejected the semantic categories of adjectives as a meaningful predictor for the use of *really*. I would like to argue that the effect of the semantic class piggybacked on the frequency of certain adjective types. As a result, the semantic class could have been a confound in previous research. Support for this interpretation can be found in Tagliamonte and Denis (2014, p. 118, p. 120) who showed that *really* increases uniformly across semantic classes as it increases in frequency.

As such, the findings presented here align with Bybee et al. (1994, p. 8) who suggest that the frequency of innovative or intrusive forms increases, or even escalates, when the respective form enters new contexts (in the current study, this new context represents the modifier position of the adjectival heads *good*, *nice* and *funny*). An additional advantage of this hypothesis is that it is also able to accommodate the observation that the expansion of the collocational profile of *really* in AusE appears to go hand-in-hand with the correlation with the frequency of the adjectives reported by both the Boruta and the regression analysis.

An explanation for why this increase in frequency is dependent on only a few high-frequency adjectives builds on the concept of entrenchment (cf. Langacker, 1987, 2008; Stefanowitsch, 2008, 2011; Stefanowitsch & Flach, 2016). Entrenchment is a concept in cognitive science that forms a continuous scale (Langacker, 1987, p. 59) and relates to repetition priming. The underlying idea is that linguistic structures become entrenched by repeated use and less entrenched by disuse (cf. Stefanowitsch & Flach, 2016, p. 104). While Langacker (1987) did not define entrenchment in psychological or psycholinguistic terms, frequently used elements, i.e. entrenched elements, are easier to retrieve from memory compared to less entrenched elements (cf. Schmid, 2000). The idea that frequency instantiates entrenchment is furthermore captured in Schmid's from-corpus-to-cognition principle (Schmid, 2000, p. 39; cf. also Stefanowitsch & Flach, 2016, p. 104). As corpora represent large collections of natural language that provide information about the frequency of constructs, a corpus-based study

appears to be particularly appealing for analyzing cognitive mechanisms that rely on entrenchment.

If an innovative variant increasingly co-occurs with high-frequency adjectives – as is the case with *really* in the present data – it will increase in overall use and thereby become more deeply entrenched over time. Deeper entrenchment would then facilitate easier retrieval and thus another increase in use. If this proposed mechanism were indeed correct, then this would not only help explain why *really* has replaced *very* in standard educated AusE, but it would also have more general implications as it would predict that the frequency of host constructions is a crucial factor in similar cases of replacement.

In conclusion, the current study shows that observable changes are multifactorial in nature and highly context-dependent which, when applied to adjective amplification, means that changes in the collocational preferences of adjectives or in the frequency of adjective types can affect – if not trigger – changes in the use of individual amplifiers or even entire amplifier systems. The results presented here suggest that successful amplifiers specialize on a few high-frequency adjectives and that successful variants broaden only after they have experienced a substantial increase in frequency due to becoming more deeply entrenched.

6. Outlook

The results presented here pose the question whether the proposed cognitive mechanism could potentially be more general in nature. It would be interesting to know if this mechanism were indeed applicable to other linguistic phenomena such as the replacement of the allomorph <(e)th> by the allomorph <(e)s> as the dominant third person singular morpheme in the verbal domain of English. If applicable, then this could mean that the mechanism proposed here could indeed represent a universal cognitive mechanism that underlies replacement processes where rivalling modifiers compete for dominance in a linguistic subsystem. Before adopting this view, however, sufficient typological and cross-linguistic evidence in support of the current findings are needed.

Acknowledgements

I am indebted to Pam Peters and Adam Smith for providing me with a preliminary version of the Australian component of the International Corpus of English (ICE-Aus) and without whom this current study would not have been possible. In addition, I would like to thank my colleagues at the University of Queensland, Australia, for their comments and feedback on earlier versions of this paper.

Disclosure statement

No potential conflict of interest was reported by the author.

Notes on contributor

Martin Schweinberger is Associate Professor in the AcqVA-Aurora Center at the Arctic University of Norway in Tromsø. In addition, Martin holds an appointment as Postdoctoral Research Fellow in

Language Technology at the University of Queensland, Australia where he has been establishing the *Language Technology and Data Analysis Laboratory* (LADAL). Martin is a language data scientist with a PhD in English linguistics and his research focuses on analyses of language acquisition as well as language variation and change using quantitative computational methods.

ORCID

Martin Schweinberger http://orcid.org/0000-0003-1923-9153

References

Aijmer, K. (2011). Are you totally spy? A new intensifier in present-day American English. In S. Hancil (Ed.), *Marqueurs discursifs et subjectivité* (pp. 155–172). Universités de Rouen and Havre.

Aijmer, K. (2018a). That's well bad. Some new intensifiers in spoken in British English. In V. Brezina, R. Love, & K. Aijmer (Eds.), *Corpus approaches to contemporary British English* (pp. 60–95). Routledge.

Aijmer, K. (2018b). Intensification with 'very', 'really' and 'so' in selected varieties of English. In S. Hoffmann, A. Sand, S. Arndt-Lappe, & L. M. Dillmann (Eds.), *Corpora and lexis* (pp. 106–139). Brill Rodopi.

Althaus, S. L., & Kim, Y. M. (2006). Priming effects in complex information environments: Reassessing the impact of news discourse on presidential approval. *The Journal of Politics*, *68*(4), 960–976. https://doi.org/10.1111/j.1468-2508.2006.00483.x

Baayen, R. H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge University Press.

Bailey, G., Wikle, T., Tillery, J., & Sand, L. (1991). The apparent time construct. *Language Variation and Change*, *3*(3), 241–264. https://doi.org/10.1017/S0954394500000569

Barnfield, K., & Buchstaller, I. (2010). Intensifiers on tyneside: Longitudinal developments and new trends. *English World-Wide. A Journal of Varieties of English*, *31*(3), 252–287. https://doi.org/10.1075/eww.31.3.02bar

Bauer, L., & Bauer, W. (2002). Adjective boosters in the English of young New Zealanders. *Journal of English Linguistics*, *30*(3), 244–257. https://doi.org/10.1177/007524202030003002

Bell, B. A., Ferron, J. M., & Kromrey, J. D. (2008). Cluster size in multilevel models: The impact of sparse data structures on point and interval estimates in two-level models. In *American statistical association, joint statistical meetings proceedings. Section on survey research methods* (pp. 1122–1129). American Statistical Association.

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

Bolinger, D. (1972). *Degree words*. Mouton.

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. https://doi.org/10.1017/S1360674315000428

Breiman, L. (2001). Random forests. *Machine Learning*, *45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

Brown, L., & Tagliamonte, S. (2012). A really interesting story: The influence of narrative in linguistic change. *University of Pennsylvania Working Papers in Linguistics*, *18*(2), [Selected Papers from NWAV 40] Article 2. http://repository.upenn.edu/pwpl/vol18/iss2/2

Bulgin, J., Elford, N., Harding, L., Henley, B., Power, S., & Walters, C. (2008). So very really variable: Social patterning of intensifier use by Newfoundlanders online. *Linguistica Atlantica*, *29*, 101–115.

Bybee, J. (2007). *Frequency of use and the organization of language*. Oxford University Press.

Bybee, J. (2010). *Language, usage and cognition*. Cambridge University Press.

Bybee, J. (2015). *Language change*. Cambridge University Press.

Bybee, J., Perkins, R. D., & Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect and modality in the languages of the world*. University of Chicago Press.

Calle-Martín, J. (2014). On the history of the intensifier wonder in English. *Australian Journal of Linguistics*, *34*(3), 399–419. https://doi.org/10.1080/07268602.2014.898224

Clarke, P. (2008). When can group level clustering be ignored? Multilevel models versus single-level models with sparse data. *Journal of Epidemiology & Community Health*, *62*(8), 752–758. https://doi.org/10.1136/jech.2007.060798

Clarke, P., & Wheaton, B. (2007). Addressing data sparseness in contextual population research. Using cluster analysis to create synthetic neighborhoods. *Sociological Methods & Research*, 35(3), 311–351.

D'Arcy, A. F. (2015). Stability, stasis and change – The longue durée of intensification. *Diachronica*, *32*(4), 449–493. https://doi.org/10.1075/dia.32.4.01dar

de Klerk, V. (2005). Expressing levels of intensity in Xhosa English. *English World-Wide. A Journal of Varieties of English*, *26*(1), 77–95. https://doi.org/10.1075/eww.26.1.04dek

Dixon, R. M. W. (1977). Where have all the adjectives gone? *Studies in Language*, *1*(1), 19–80. https://doi.org/10.1075/sl.1.1.04dix

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* (pp. 1–49). Oxford University Press.

Estival, D. (1985). Syntactic priming of the passive in English. *Text*, *5*, 7–21.

Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. Sage.

Fuchs, R. (2016). Register variation in intensifier usage across Asian Englishes. In H. Pichler (Ed.), *Discourse-pragmatic variation and change: Insights from English* (pp. 185–213). Cambridge University Press.

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. https://doi.org/10.1075/ijcl.22.3.03fuc

Gries, S. T. (2005). Syntactic priming: A corpus-based approach. *Journal of Psycholinguistic Research*, *34*(4), 365–399. https://doi.org/10.1007/s10936-005-6139-3

Gries, S. T. (2009). *Statistics for Linguistics with R*. Mouton de Gruyter.

Gries, S. T. (2013). Sources of variability relevant to the cognitive sociolinguist, and corpus – as well as psycholinguistic methods and notions to handle them. *Journal of Pragmatics*, *52*, 5–16. https://doi.org/10.1016/j.pragma.2012.12.011

Gries, S. T., & Stefanowitsch, A. (2004). Extending collostructional analysis: A corpus-based perspectives on 'alternations'. *International Journal of Corpus Linguistics*, *9*(1), 97–129. https://doi.org/10.1075/ijcl.9.1.06gri

Hilpert, M. (2006). Discussion note: Distinctive collexeme analysis and diachrony. *Corpus Linguistics and Linguistic Theory*, *2*(2), 243–256. https://doi.org/10.1515/CLLT.2006.012

Hilpert, M. (2013). *Constructional change in English: Developments in allomorphy, word formation, and syntax*. Cambridge University Press.

Ito, R., & Tagliamonte, S. (2003). *Well* weird, *right* dodgy, *very* strange, *really* cool: Layering and recycling in English intensifiers. *Language in Society*, *32*(2), 257–279. https://doi.org/10.1017/S0047404503322055

Kursa, M. B., & Rudnicki, W. R. (2010). Feature selection with the Boruta package. *Journal of Statistical Software*, *36*(11), 1–13. https://doi.org/10.18637/jss.v036.i11

Labov, W. (1985). Intensity. In D. Schiffrin (Ed.), *Meaning, form and use in context: Linguistic applications* (pp. 43–70). Georgetown University Press.

Labov, W. (2001). *Principles of linguistic change: Volume 2: Social factors*. Blackwell.

Labov, W. (2002, August 2). *Driving forces in linguistic change* [Paper presentation]. 2002 international conference on Korean linguistics, Seoul National University. Retrieved January 1, 2019, from https://www.ling.upenn.edu/~wlabov/Papers/DFLC.htm

Labov, W. (2010). *Principles of linguistic change: Volume 3: Cognitive and cultural factors*. Wiley-Blackwell.

Langacker, R. W. (1987). *Foundations of cognitive grammar: Vol. 1. Theoretical prerequisites*. Stanford University Press.

Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. Oxford University Press.

Lorenz, G. R. (1999). *Adjective intensification – Learners versus native speakers: A corpus study of argumentative writing*. 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* (pp. 143–161). John Benjamins.

Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology*, *1*(3), 86–92. https://doi.org/10.1027/1614-2241.1.3.86

Macaulay, R. (2002). Extremely interesting, very interesting, or only quite interesting: Adverbs and social class. *Journal of Sociolinguistics*, *6*(3), 398–417. https://doi.org/10.1111/1467-9481.00194

Macaulay, R. (2006). Pure grammaticalization: The development of a teenage intensifier. *Language Variation and Change*, *18*(3), 267–283. https://doi.org/10.1017/S0954394506060133

Maddeaux, R., & Dinkin, A. (2017). Is like like like?: Evaluating the same variant across multiple variables. *Linguistics Vanguard*, *3*(1). Retrieved January 1, 2019, from, https://www.degruyter.com/view/j/lingvan.2017.3.issue-1/lingvan-2015-0032/lingvan-2015-0032.xml?format=INT https://doi.org/10.1515/lingvan-2015-0032

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. https://doi.org/10.1515/text-2012-0036

Méndez-Naya, B. (2003). On intensifiers and grammaticalization: The case of SWIÞE. *English Studies*, *84*(4), 372–391. https://doi.org/10.1076/enst.84.4.372.17388

Méndez-Naya, B. (2008). On the history of downright. *English Language and Linguistics*, *12*(2), 267–287. https://doi.org/10.1017/S1360674308002621

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* (pp. 191–214). John Benjamins.

Mohammad, S. M., & Turney, P. D. (2013). Crowd sourcing a word-emotion association lexicon. *Computational Intelligence*, *29*(3), 436–465. https://doi.org/10.1111/j.1467-8640.2012.00460.x

Murphy, B. (2010). *Corpus and sociolinguistics: Investigating age and gender in female talk*. John Benjamins.

Mustanoja, T. F. (1960). *A middle English syntax, part 1, parts of speech*. Société Néophilologique.

Nevalainen, T. (2008). Social variation in intensifier use: Constraint on -ly adverbialization in the past? *English Language and Linguistics*, *12*(2), 289–315. https://doi.org/10.1017/S1360674308002633

Nevalainen, T., & Rissanen, M. (2002). Fairly pretty or pretty fair? On the development and grammaticalization of English downtoners. *Language Sciences*, *24*(3–4), 359–380. https://doi.org/10.1016/S0388-0001(01)00038-9

Paradis, C. (1997). *Degree modifiers of adjectives in spoken British English*. Lund University Press.

Paradis, C. (2008). Configurations, construals and change: Expressions of DEGREE. *English Language and Linguistics*, *12*(2), 317–343. https://doi.org/10.1017/S1360674308002645

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* (pp. 177–192). John Benjamins.

Pertejo, P. N., & Martínez, I. 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. https://doi.org/10.1075/fol.21.2.03pal

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

Peters, H. (1993). *Die englischen Gradadverbien der Kategorie booster*. Narr.

Peters, H. (1994). Degree adverbs in early modern English. In D. Kastovsky (Ed.), *Studies in early modern English* (pp. 269–288). Mouton de Gruyter.

Pickering, M. J., & Garrod, S. (2017). Priming and language change. In M. Hundt, S. Mollin, & S. E. Pfenninger (Eds.), *The changing English language. Psycholinguistic perspectives* (pp. 173–190). Cambridge University Press.

Plutchik, R. (1980). A general psychoevolutionary theory of emotion. *Emotion: Theory, Research, and Experience*, *1*(3), 3–33. https://doi.org/10.1016/B978-0-12-558701-3.50007-7

Plutchik, R. (1994). *The psychology and biology of emotion*. Harper Collins.

Poplack, S. (1980). The notion of the plural in Puerto Rican English: Competing constraints on (s) deletion. In W. Labov (Ed.), *Locating language in time and space* (pp. 55–68). Academic Press.
Poplack, S., & Tagliamonte, S. (1993). The zero-marked verb: Testing the creole hypothesis. *Journal of Pidgin and Creole Languages*, *8*(2), 171–206. https://doi.org/10.1075/jpcl.8.2.02tag
Poplack, S., & Tagliamonte, S. (1996). Nothing in context: Variation, grammaticization and past time marking in Nigerian Pidgin English. In P. Baker & A. Syea (Eds.), *Changing meanings, changing functions* (pp. 71–94). Westminster University Press.
Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (Eds.). (1985). *A comprehensive grammar of the English language*. Longman.
Rickford, J., Wasow, T., Zwicky, A., & Buchstaller, I. (2007). Intensive and quotative *all*: Something old, something new. *American Speech*, *82*(1), 3–31. https://doi.org/10.1215/00031283-2007-001
Rissanen, M. (2008). From 'quickly' to 'fairly': On the history of *rather*. *English Language and Linguistics*, *12*(2), 345–359. https://doi.org/10.1017/S1360674308002657
Schmid, H.-J. (2000). *English abstract nouns as conceptual shells: From corpus to cognition* [Topics in English Linguistics 34]. Mouton de Gruyter.
Schweinberger, M. (2018). The discourse particle eh in New Zealand English. *Australian Journal of Linguistics*, *38*(3), 395–420. https://doi.org/10.1080/07268602.2018.1470458
Schweinberger, M. (2020a). Speech-unit final *like* in Irish English. *English World-Wide. A Journal of Varieties of English*, *41*(1), 89–117. https://doi.org/10.1075/eww.00041.sch
Schweinberger, M. (2020b). Using semantic vector space models to investigate lexical replacement – A corpus based study of ongoing changes in intensifier systems. In Y. Asahi (Ed.), *Proceedings of methods in dialectology XVI. Papers from the sixteenth international conference on methods in dialectology 2017* (pp. 241–249). Peter Lang.
Sowa, J. (2009). 'Sweet as!': The intensifier as in New Zealand and Australian English. *English Today*, *25*(2), 58–61. https://doi.org/10.1017/S0266078409000212
Stefanowitsch, A. (2008). Negative entrenchment: A usage-based approach to negative evidence. *Cognitive Linguistics*, *19*(3), 513–531. https://doi.org/10.1515/COGL.2008.020
Stefanowitsch, A. (2011). Cognitive linguistics meets the corpus. In M. Brdar, S. T. Gries, & M. Žic Fuchs (Eds.), *Cognitive linguistics: Convergence and expansion* (pp. 257–289). John Benjamins.
Stefanowitsch, A., & Flach, S. (2016). The corpus-based perspective on entrenchment. In H.-J. Schmid (Ed.), *Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge* (pp. 101–128). Walter de Gruyter.
Stefanowitsch, A., & Gries, S. T. (2003). Collostructions: Investigating the interaction between words and constructions. *International Journal of Corpus Linguistics*, *8*(2), 209–243. https://doi.org/10.1075/ijcl.8.2.03ste
Stefanowitsch, A., & Gries, S. T. (2005). Covarying collexemes. *Corpus Linguistics and Linguistic Theory*, *1*(1), 1–43. https://doi.org/10.1515/cllt.2005.1.1.1
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* (pp. 69–78). Rodopi.
Su, Y. (2016). Corpus-based comparative study of intensifiers: Quite, pretty, rather and fairly. *Journal of World Languages*, *3*(3), 224–236. https://doi.org/10.1080/21698252.2017.1308306
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. https://doi.org/10.1515/cllt.2005.1.1.113
Szmrecsanyi, B. (2006). *Morphosyntactic persistence in spoken English: A corpus study at the intersection of variationist sociolinguistics, psycholinguistics, and discourse analysis*. Walter de Gruyter.
Tagliamonte, S. (2006). *Analysing sociolinguistic variation*. Cambridge University Press.
Tagliamonte, S. (2008). So different and pretty cool! Recycling intensifiers in Toronto, Canada. *English Language & Linguistics*, *12*(2), 361–394. https://doi.org/10.1017/S1360674308002669
Tagliamonte, S. (2012). *Variationist sociolinguistics: Change, observation, interpretation*. Wiley-Blackwell.
Tagliamonte, S., & Baayen, H. R. (2012). *Models, forests and trees of York English: Was/were variation as a case study for statistical practice*. Potsdam Mind Research Repository. Retrieved January 1, 2019,

from, http://read.psych.uni-potsdam.de/index.php?option=com_content&view=article&id=78:tagliamonte-and-baayen-2012-models-forests-and-trees-of-york-english-was-were-variation-as-a-case-study-for-statistical-practice&catid=24:publications&Itemid=32
Tagliamonte, S., & D'Arcy, A. F. (2009). Peaks beyond phonology: Adolescence, incrementation, and language change. *Language*, *85*(1), 58–108. https://doi.org/10.1353/lan.0.0084
Tagliamonte, S., & Denis, D. (2014). Expanding the transmission/diffusion dichotomy: Evidence from Canada. *Language*, *90*(1), 90–136. https://doi.org/10.1353/lan.2014.0016
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. https://doi.org/10.1215/00031283-80-3-280
Tao, H. (2007). A corpus-based investigation of absolutely and related phenomena in spoken American English. *Journal of English Linguistics*, *35*(1), 5–29. https://doi.org/10.1177/0075424206296615
Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science*, *247*(4940), 301–306. https://doi.org/10.1126/science.2296719
Wagner, S. (2017). *Totally new* and *pretty awesome*: Amplifier-adjective bigrams in GloWbE. *Lingua. International Review of General Linguistics. Revue internationale De Linguistique Generale*, *200*, 63–83. https://doi.org/10.1016/j.lingua.2017.08.004
Wallenberg, J. C. (2013). A unified theory of stable variation, syntactic optionality, and syntactic change. In *Paper presented at the 15th diachronic generative syntax (DiGS) conference*. University of Ottawa.
Weiner, J., & Labov, W. (1983). Constraints on the agentless passive. *Journal of Linguistics*, *19*(1), 29–58. https://doi.org/10.1017/S0022226700007441
Xiao, R., & Tao, H. (2007). A corpus-based sociolinguistic study of amplifiers in British English. *Sociolinguistic Studies*, *1*(2), 241–273. https://doi.org/10.1558/sols.v1i2.241
Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, *1*(1), 3–14. https://doi.org/10.1111/j.2041-210X.2009.00001.x

Corpora

Davies, M. (2010). *The corpus of historical American English: 400 million words, 1810–*2009. Retrieved January 1, 2019, from http://corpus.byu.edu/coha/
ICE Website. Retrieved January 1, 2019, from https://www.ice-corpora.uzh.ch/en.html
The British National Corpus, version 3 (BNC XML Edition) (2007). *Distributed by Bodleian Libraries*. University of Oxford, on behalf of the BNC Consortium. http://www.natcorp.ox.ac.uk/
Peters, P., & Smith, A. (forthcoming). *ICE Australia*. Retrieved January 1, 2019, from https://www.ausnc.org.au/corpora/ice; https://www.ice-corpora.uzh.ch/en/joinice/Teams/iceaus.html

Software

Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalised linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504
Hornik, K. (2016). *openNLP: Apache OpenNLP tools interface* (Version 0.2-6). Retrieved January 1, 2019, from https://cran.r-project.org/web/packages/openNLP/openNLP.pdf
Jockers, M. (2017). *syuzhet: Extracts sentiment and sentiment-derived plot arcs from text* (Version 1.0.1). Retrieved January 1, 2019, from https://github.com/mjockers/syuzhet
Kursa, M. B., & Rudnicki, W. R. (2018). *Boruta: Wrapper algorithm for all relevant feature selection* (Version 6.0.0). Retrieved January 1, 2019, from https://notabug.org/mbq/Boruta/
R Core Team. (2019). *R: A language and environment for statistical computing. R foundation for statistical computing*. https://www.R-project.org/

Appendix

Table A1 Overview of the step-wise step-up model fitting procedure and results of the power analyses

Model	Term added	Compared to …	AIC (BIC)	X^2	*p*-value	Power (%)
m0.glmer	1 (Intercept)	–	777.0 (785.7)	–	–	–
m1.glmer	Age	m0.glmer	769.2 (782.3)	9.76	0.00178**	100.0 (96.38, 100.0)
m2.glmer	Gender	m1.glmer	763.84 (781.3)	7.36	0.00667**	100.0 (96.38, 100.0)
m3.glmer	Frequency	m2.glmer	752.07 (773.9)	13.77	0.00021***	100.0 (96.38, 100.0)
m4.glmer	Priming	m3.glmer	748.23 (774.43)	5.84	0.01568*	100.0 (96.38, 100.0)
m5.glmer	Age:Gender	m4.glmer	749.63 (780.2)	0.59	0.44051	97.0 (91.48, 99.38)
m6.glmer	Age:Frequency	m4.glmer	746.11 (776.68)	4.11	0.04251*	100.0 (96.38, 100.0)
m7.glmer	Age:Priming	m4.glmer	747.9 (778.46)	2.33	0.12684	92.0 (84.84, 96.48)
m8.glmer	Gender:Frequency	m4.glmer	749.34 (779.9)	0.89	0.34504	100.0 (96.38, 100.0)
m9.glmer	Gender:Priming	m4.glmer	749.12 (779.69)	1.11	0.29287	55.00 (44.73, 64.97)
m10.glmer	Frequency:Priming	m4.glmer	750.09 (780.66)	0.13	0.71351	100.0 (96.38, 100.0)
m11.glmer	Age*Gender*Frequency	m4.glmer	750.8 (794.47)	5.43	0.24634	79.0 (69.71, 86.51)
m12.glmer	Age*Gender*Priming	m4.glmer	752.48 (796.15)	3.75	0.44146	35.0 (25.73, 45.18)
m13.glmer	Age*Frequency*Priming	m4.glmer	749.81 (793.47)	6.42	0.16984	63.0 (52.76, 72.44)
m14.glmer	Gender*Frequency*Priming	m4.glmer	754.06 (797.72)	2.17	0.70385	44.0 (34.08, 54.28)