

**Sacred That and Wicked Which:
Prescriptivism and Change in the Use of English Relativizers**

Axel Bohmann

The University of Texas at Austin

Patrick Schultz³¹

The University of Texas at Austin

1. Introduction

Prescriptivism, “the belief that the grammar of a language should lay down rules to which usage must conform” (Oxford English Dictionary n.d.) as the *Oxford English Dictionary* has it, is generally held in low esteem among most contemporary linguists. Prescriptive grammars or usage guides are commonly seen as unscientific and old-fashioned endeavors, indeed “the term [prescriptivist] is pejorative in linguistic contexts” (Crystal 2008). It is therefore not surprising that little research has been done to investigate the linguistic influence of prescriptive grammars and usage guides such as H.W. Fowler’s (1965) *Dictionary of Modern English Usage* or William Strunk’s (1999) *Elements of Style*. There is, as Donald Mackay (1980, p.349) puts it, “a long-standing rift between prescriptive and theoretical linguistics” due to the fact that

prescriptive linguists view theoretical linguistics as irrelevant to their goal of teaching language use, whereas theoretical linguists view prescriptive problems as unfundamental and irrelevant to their goal of describing the principles underlying language use.

Examples of this attitude towards prescriptivist writings abound; in his review of *The Elements of Style* Geoffrey Pullum (2009) for example, calls the authors “grammatical incompetents” and “idiosyncratic bumbler” and the book “a toxic mix of purism, atavism and personal eccentricity”. Alexandra D’Arcy (2010) makes a similar point in an entry on the Oxford University Press blog:

³¹ This research is part of an ongoing collaboration between the English Departments at the University of Texas at Austin and the University of Freiburg, Germany (Project directors: Lars Hinrichs, Benedikt Szmrecsanyi).

I am not only a linguist but a *sociolinguist* (of all things!). I describe language as actually used and I revel in the differences and variations of language in practice [;...] there is no place for prescription in my world. The notion of *should* does not apply.

The website *Language Log*, where several eminent linguists blog about linguistic matters, has its own category of “Prescriptivist Poppycock” where the writers disprove prescriptivist the validity of prescriptivist precepts (Lieberman 2011), collect examples of usage guide authors breaking their own rules (Lieberman 2006) and accuse prescriptivists of “grammatical egocentrism” and “cluelessness” (Zwicky 2009). Besides regarding the prescriptivist endeavor as misguided, it is often seen as pointless because language supposedly changes regardless of human intervention. Even prescriptive grammarians such as Henry Fowler are doubtful of their influence on actual language use:

What grammarians say should be has perhaps even less influence on what shall be than even the more modest of them realize, usage itself evolves little disturbed by their likes and dislikes. (Fowler 1965, p.622)

On the other hand, prescriptivist usage guides such as the *Dictionary of Modern English Usage* or *The Elements of Style* have sold millions of copies. They are assigned readings for college students and are used as writing guidelines by journalists, scholars and novelists. Common sense suggests that they could quite possibly have had an influence on written English. A comprehensive, empirical study of the influence of prescriptivist rules on language change, however, has yet to be done.

2. Previous Research

A few studies have investigated the influence of prescriptivism on Standard English in isolated areas of grammar. Auer (2006) and Auer & Gonzalez-Diaz (2005) study the impact of 18th century grammars on the development of the subjunctive in Britain. They find that prescriptivist grammarians had a slight, temporary influence on the use of the subjunctive, but were ultimately unable to stop or reverse the general trend towards use of the indicative in its place. They conclude that their findings

can be interpreted as a warning against the danger of overestimating the explanatory potential of prescriptivism and a call for a more careful reanalysis of its impact in processes of language change in the history of English. (Auer & Gonzalez-Diaz 2005, p.336)

Similarly, Busse & Schroeder (2010) find that prescriptivists were unable to stop the drift of *hopefully* from being used in its adverbial meaning toward usage as a sentence adverbial in British English. And while Facchinetti (2000) sees usage following prescriptivist guidelines on *shall* and *will* from the 17th century onwards, it remains unclear whether this change was initiated by prescriptivist grammarians or whether grammarians were just reinforcing a change in progress. Other studies, however, claim to detect influences of prescriptivism on English usage: Busse & Schroeder (2010) for example, suggest that the increased usage of *different from* (instead of *different to* or *than*) might be an effect of prescriptivist influence. Leech et al (2009, p.230) find that writers increasingly use the relative pronoun *that* instead of *which*, a development they attribute to usage guides and teaching of prescriptivist rules as well.

3. The Variable

In this paper, we will investigate the influence of prescriptivism on the use of the relativizers *that*, *which* and relativizer omission (zero) in written English. Unlike Leech et al (cited above), we exclude all non-restrictive clauses, possessive relative clauses and pied-piping constructions and focus on restrictive relative clauses where *which*, *that* and zero are interchangeable, i.e. the variable discussed in the prescriptivist literature cited below. Most major usage guides address the issue of relativizer choice in restrictive relative clauses. From a purely descriptive standpoint, speakers are free to choose between the variants *who*, *that*, *which* and zero (e.g. *The house [that/which/Ø] Jack built*). Writers of usage guides, however, introduce a much stricter rule: There is a consensus in the prescriptive literature (Peters & Young 1997) that *that* is the correct variant in this context. Thus, *The house that Jack built* is the preferable, ‘most correct’ version of the example above. William Strunk (1999, p.59), for example, advises his readers that

that is the defining, or restrictive, pronoun, *which* the nondefining, or nonrestrictive. [...] it would be a convenience to all if these two pronouns were used with precision. Careful writers, watchful for small conveniences, go *which*-hunting, remove the defining *whiches*, and by so doing improve their work.

30 years before Strunk, Henry Fowler (1965, p.635) wrote a little more cautiously that

the two kinds of relative clause, to one of which *that* & to the other of which *which* is appropriate, are the defining & the non-defining; & if writers would agree to regard *that* as the defining relative pronoun, & *which* as the non-defining, there would be much gain both in lucidity & in ease.”

More recently, Eric Partridge (1957, p.364) writes on the restrictive relative clause that “it is ushered in by *that*” or no relative pronoun at all (zero).

The historical development of English relativizers between the 16th and 20th century has been studied by Catherine Ball (1996) She summarizes her findings with regard to the variable under investigation here in the following graph:

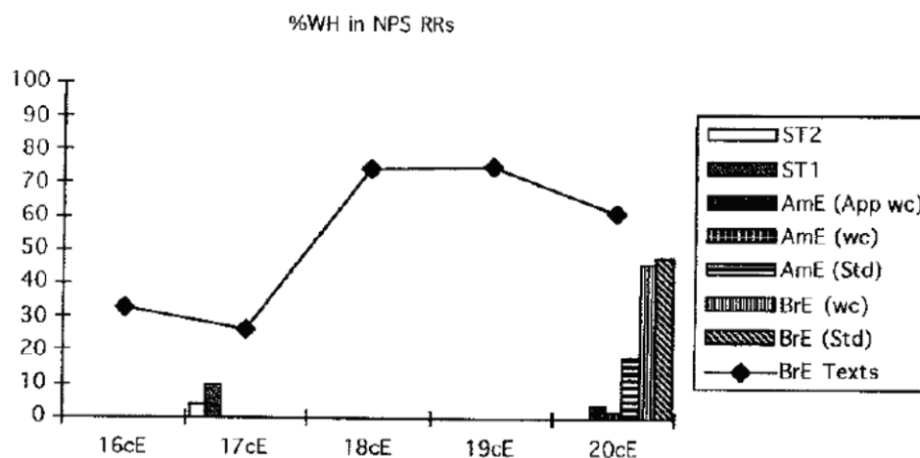


Figure 1: WH in non-personal subject relatives, spoken and written (Ball 1996, p.250).

There are issues of representativeness and interpretability of the graph here. The data this graph was built from comprises only 583 relevant constructions for four centuries. Moreover, these have been summarized in bins by the century, which it would make more sense to represent in a bar plot rather than a continuous line. Nonetheless, if we accept her evidence, the development between the 17th and the 18th century appears drastic. During this time, a shift in the paradigm of relativization seems to have taken place, favoring *which* as the dominant option, at least in written British English, which is represented by the line.

The apparent, although slight, reversal of the trend that we see towards the 20th century in Figure 1 finds some confirmation in Leech et al. (2009, pp.226-33), even if the data are not completely compatible. The authors run counts of all cases of *wh-* and *that* as relativizers (regardless of syntactic context) in written British and American English around 1961 and 1991. They rely on the four central corpora of the Brown family, which are also the database for the present study and are presented in more detail below. Two observations are of interest here. Firstly, in both British and American English the frequency of *that* increases over time, while that of *wh-* relativizers goes down. Secondly, this trend is by far more pronounced in American English than in British English. The problem with these counts is that they are made independent of syntactic context and hence include non-restrictive relative clauses, which do not allow *that*, pied-piping constructions, which require *which* and the relative pronouns *who* and *whom*. Their numbers are consequently not very sensitive to the constraints that operate on relativization in English and should be regarded as a first estimate on the basis of which more fine-grained analyses can be conducted.

Relating the figures in Leech et al. to Ball's historical findings, it seems as if the development of the 18th century in the paradigm of relativization is in the process of being reversed. Two reasons suggest the possibility of attributing at least part of this to the influence of prescriptivism. The phenomenon as defined for the present purposes appears to have taken form during the beginning of the 20th century, Fowler (1965) – whose first edition was published in 1926 – being the first noteworthy example. There is thus a temporal correspondence, even if a very loose one so far. The diverging turns American and British English appear to be taking may be another pointer to the influence of prescriptivism, since

[p]rescriptivism maintains it [sic] hold over written AmE through channels which are absent from the UK, such as handbooks for obligatory freshmen English courses, and the pronouncement of 'language mavens' in the press. (Leech et al. 2009, p.264)

Based on the above findings, the hypothesis for this paper is as follows: ongoing change in the frequency of relativizers in restrictive, non-animate, non-pied-piping relative clauses can at least in part be attributed to the influence of prescriptivism in the form of usage manuals and style guides. While the findings in Ball (1996) and Leech et al. (2009) suggest that this hypothesis might be true, they do not provide strong evidence for it, nor do they make any claims beyond the speculative. To investigate the hypothesis further, the present study attempts to gauge prescriptive influences on choice of relativization strategy against other current trends in (written Standard) English morpho-syntax and tease these apart through multinomial logistic regression modeling.

4. The Dataset

The hypothesis will be tested against data from the four core collections of the Brown family of corpora: the Brown corpus (written Standard American English, 1961), the Lancaster-Oslo-Bergen corpus (LOB, written Standard British English, 1961), the Frown corpus (written Standard American English, 1991) and the F-LOB corpus (written Standard British English, 1991). Each of these contains approximately 1 million words of published writing, divided into individual text samples of ca. 2,000 words each. These samples fall into 4 broad genre categories: Press, General Prose, Learned and Fiction. Figure 2 gives a summary of the basic setup of the four corpora, including the dimensions of variation they have been collected to cover specifically (for detailed descriptions of the corpora see Leech et al. (2009, chapter 2).

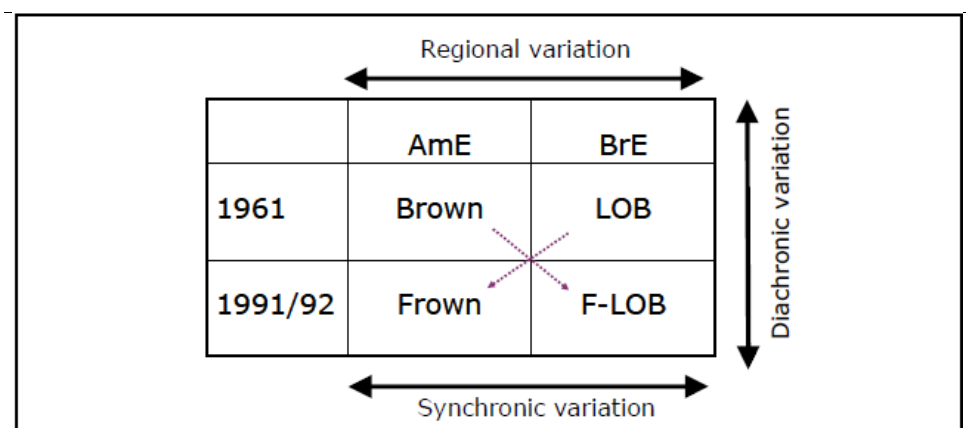


Figure 2: Setup of the Brown corpora (Busse & Schroeder 2010, p.89).

The present study makes use of POS-tagged versions of the Brown family corpora (for a detailed description see (Hinrichs et al. 2010) which have been additionally tagged for the variable under investigation. Tags for “zero” relativization were inserted automatically in the appropriate places and manually post-edited for correctness. This version also already incorporates a differentiation between the cases of interest here and other cases of relativization. Non-defining “which” and “which” in combination with pied-piping, for instance, have a different POS-tag (“<DDLX>”) than defining “which” without pied-piping (“<DDL>” or “<DDLO>”, depending on whether the relativizer occurs in a subject or an object gap). Here is a sample from the Learned section of the Brown corpus which illustrates the tagging employed in this study:

```
<s>
<AT>The <NN1>effect <IO>of <$T> <NN2>drugs <WPRS>that
<VV0>act <II>on <AT>the <JJ>iodide-concentrating <NN1>mechanism
<%T> <VM>can <VABI>be <VVN>counteracted <II>by <NN1>addition
<IO>of <RR>relatively <JJ>large <NN2>amounts <IO>of <NN1>iodine
<II>to <AT>the <NN1>diet <.>.
</s>
```

(Brown J14)

As can be seen from the highlighted parts, every defining relative clause has received a tag indicating the beginning of the antecedent (“<\$T>”, “<\$W>” and “<\$Z>” for “that”, “which” and “zero” respectively), one immediately before the relativizer (in this case “<WPRS>”; they differ according to form and syntactic function of the relativization device) and one at the end of the relative clause (“<%T>”, “<%W>” or “<%Z>”). This extended tagging allows for operations on the texts which go beyond a mere count of *that*,

which, and *zero*, making it possible to extract various additional syntactic and discourse-related variables. A note on corpus size is in place. With 4 million words altogether, the four subcorpora of our dataset are rather small in terms of present-day corpus linguistics, which often works with figures well above 100 million words. Yet Marianne Hundt and Christian Mair (1999, p.224) maintain that for “phenomena [...] comprising the core-grammar of Standard English [...], enough data can usually be extracted from one-million-word-samples at least for initial orientation”, and this describes quite accurately what the present study attempts. For the outcome variable itself, 17257 tokens were recorded altogether, which is enough to yield robust statistical results. Depending on the nature of the predictors that are to be incorporated into a model corpus size can become increasingly problematic with the Brown corpora, but we are confident that for our present purposes such issues do not present themselves.

5. The Predictors

The aim of this paper is to gauge the extent to which prescriptivism is affecting the paradigm of restrictive relativization in written Standard English, in the context outlined above. An important part of this consists of identifying the unique influence of prescriptivism as opposed other broad trends that have been found in recent developments of English grammar, most notably colloquialization (Leech et al. 2009, pp.239-49) and information densification (Leech et al. 2009, pp.249-52); (Hinrichs & Szmrecsanyi 2007). Both these drifts can be assumed to operate on the system of relativization. For instance, the “zero” option may be favored as a means to densify a text by omitting an optional word. Similarly, “which” has been found to be very infrequent in spoken language (Hinrichs 2006). Hence, if there is a decline in percentages of *which* this may be better explained as a result of the general colloquialization of written English than as adherence to certain style advice. What the proposed study attempts, then, is to investigate how choice of relativization correlates with these trends, and to what extent indicators of adherence to prescriptivism help predict which relativizer is used in restrictive contexts. In order to achieve this goal, the following predictors were extracted:

Time: A binary variable comparing the samples from 1961 (Brown, LOB) to those from 1991 (Frown, F-LOB). The influence of this variable will be related to the findings of Hinrichs (2006) and Leech et al. (2009) Since we are interested in language change, variation along this axis will be of special interest.

Variety: This is also a binary variable, comprising American (AmE) and British English (BrE) in their written Standard forms. The findings of Leech et al. (2009) and Hinrichs (2006) cited above suggest that the two varieties have developed quite differently with regard to relativization in the time period under investigation. Including variety as a predictor will enable a more fine-grained analysis of these differences.

Genre: This category includes the four text categories of the Brown corpora: Press, General Prose, Learned, and Fiction. Each 2000 word sample in the corpora comes with a file ID, from which the according genre can be inferred. Since different genres have been found to develop different stylistic preferences (Biber 2003); (Hundt & Mair 1999); (Biber & Finegan 1989), this factor can be assumed to play a role in the choice of relativization strategy. Including it also makes it possible to relate the findings to what other authors have already found out about the general drifts various genres follow.

Type-Token Ratio (TTR): Since one force that may be influential besides prescriptivism is information densification (Biber 2003); (Hinrichs & Szmrecsanyi 2007); (Leech et al. 2009), there needs to be some measure of how dense the text sample in question is. TTR, i.e. the ratio of different words versus the number of words in total in the text sample, is taken here to be an indicator of information density. For the calculation,

only surface forms are considered (no lemmatization). The value of TTR is not stable across text sizes but decreases with larger texts. However, this problem is mitigated by the fact that the size from which the value is calculated, namely the individual 2,000 word sample, is roughly the same.

Frequency of nouns in the text sample: The “nouniness” of a text has been found to be a reliable indicator of its degree of information density (Leech et al. 2009, p.211). We therefore measure the relative frequency of nouns in the 2,000 word sample a given relativization device is extracted from. For easier interpretability, the relative frequencies here – as well as for all other predictors of this kind – are normalized to a value per 10,000 words.

Frequency of subordinating conjunctions in the text sample: As an indicator of the syntactic complexity of the text as a whole, we include this measure in our list of predictors. More complex texts indicate a higher degree of formality, which may favor certain relativization strategies over others.

Personal Pronoun Frequency in the text sample: We choose this variable as a measure for the “colloquialness” of a text sample at hand. There are plenty of indicators of colloquialization, but the problem is that many of these are also related to other phenomena. Declining use of the passive, for instance, could also be a reaction to prescriptivist pressure as the passive is a construction under constant attack by style advisors. Similarly, increasing use of contractions (Hundt & Mair 1999) may arguably be a result of information densification. The use of personal pronouns appears to be relatively unrelated to these other trends.

Frequency of Stranded Prepositions in the text sample: This variable is included since an aversion against stranding is the subject of another prescriptivist mantra: “Never end a sentence with a preposition.” If the hypothesis holds true, and prescriptivism in general has an influence on choice of relativization, high predictive power is expected for this variable. Instances of stranding are here defined as a preposition followed by an item of punctuation like a comma, full stop or question mark.

Ratio between Passive and Active Constructions: For the same reasons as preposition stranding, a measure for the use of the passive voice will be included as well. Prescriptivists tend to abhor this construction (e.g. Strunk & White 1999, p.18). Hence, fewer passive constructions may indicate stricter adherence to style manuals’ ‘rules.’ Admittedly, it may also indicate greater colloquialization. Here, it will be worthwhile to look at how the weights of individual predictors relate to each other in the final mode. Two means of estimating the frequency of passives were initially included: passive voice constructions per 10,000 words and the ratio between passive and active verbs in the text sample. We decided to use the latter as it is estimated to be more accurate, describing only cases of true variation rather than raw counts. Also, in a test for the condition number of all continuous predictors, passive-active ratio turned out to be less correlated with other predictors than passives per 10,000 words. In order to count passive constructions, cases of auxiliary *be* + past participle of any full verb were recorded, as well as cases with adverbs in between. For active verbs, all indicative forms of full verbs, auxiliary *have* + past participle of a full verb (present past), and auxiliary *be* + present participle of any full verb (progressives) were considered.

Syntactic Function (or “gap”) of relativizer: A relative clause can have an independent subject, in which case the relativization device appears in the object position of the clause, or it can have the relativization device itself in subject position. An exception to this is *zero* relativization, which can only be employed with an independent subject in the relative clause. To test with regard to *which* and *that* whether one syntactic function favors either relativization device over the other, we extract this binary factor from all constructions in question.

Length of relative clause: Depending on the complexity of the clause with which a noun-phrase is post-modified, different devices for relativization may be privileged. If, for instance, the relative clause is very long, there may be a desire to mark it overtly and in this case *which* “can give a stronger, more unambiguous signal for the beginning of a relative clause” than *that* or zero (Leech et al. 2009, p.227). The length of the clause is measured in individual words.

Distance between head of the antecedent and beginning of the relative clause: This variable is included as a first estimate to assess the adjacency between noun-phrase and relative clause. Most relative clauses directly follow the noun-phrase they modify, but there is also the possibility of additional syntactic material being inserted in between, as in “No house was ever built that could not have been built better ...” (Brown E35). At the time of writing, a reliable method for extracting adjacency was not available. The distance in number of words between the head of the antecedent noun-phrase and the beginning of the relative clause stands in as an ersatz variable. If the value here is 0 the two are adjacent; with increasing values the chance of non-adjacency increases as well. It has to be said, though, that the correlation is far from perfect. Better methods for extracting adjacency are in the process of being developed.

Definiteness and number of the antecedent head: Since relative clauses provide additional information about a noun-phrase, the degree of definiteness as well as the number of this noun-phrase may play a role in how the relative clause is introduced. Hence we extracted these two binary factors for every construction in question to include as a predictor in our model.

6. Results

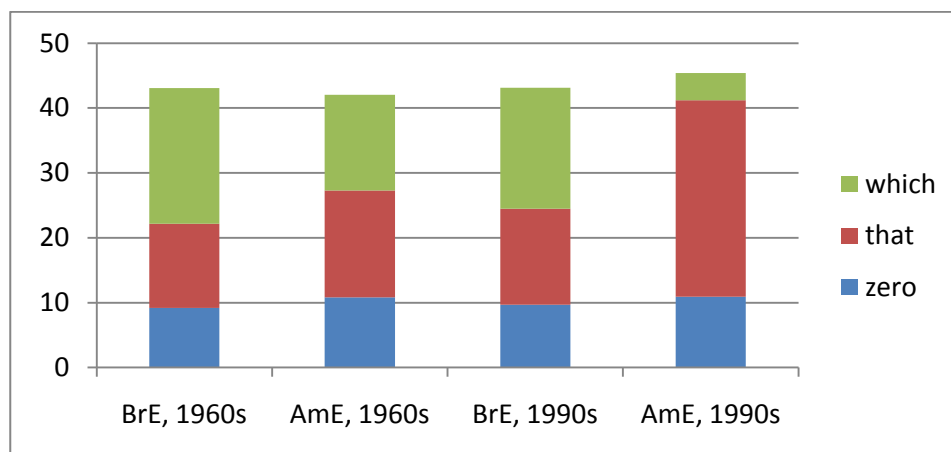


Figure 3. Frequencies of relativizers (N/10,000) in the four corpora

Figure 3 shows the distribution of absolute counts for *which*, *that*, and *zero* in Brown, LOB, Frown and F-LOB. The numbers are not normalized, hence the slight difference in height between the four plots. The big picture of both diachronic development as well as synchronic variation can be gleaned from this figure. With regard to *zero*, not much seems to be happening along either of the axes time and variety. In accordance with Bell's findings, variation and change in the use of *that* and *which* appear to be best explained in

terms of each other, with zero remaining an infrequent but stable alternative diachronically and synchronically.³²

With regard to the situation of competition between *which* and *that*, both major observations in Leech et al. (2009, pp.226-33) are corroborated by our findings. There is an increase in frequency of *that* at the expense of *which* between 1961 and 1991 in both varieties, but this development is progressing at very different speeds in British and American English. Whereas in the United States there is a drastic shift in preferences, the UK is taking a much more gradual course in the same direction. The more specific context of non-pied-piping restrictive relative clauses that our counts reflect even exacerbates the force of the change in American English, where by 1991 *which* has become the least frequent option by far.

These findings lend support to the hypothesis that prescriptivism is playing a role in the choice of relativization device. The diachronic increase of *that* can be interpreted as a style rule's gaining support over time, whereas the differences between British and American English point to the diverging institutional contexts for the enforcement of such rules in the two countries. The development is particularly interesting if read against Ball's (1996) findings, since the prevalent trend in her data up until the beginning of the 20th century was one towards *which* as a device for introducing restrictive relative clauses in written English. Our findings indicate a reversal of this trend. In order to gain a more in-depth understanding of the factors that influence choice of relativization device, the predictors in (5) were entered into logistic regression models for *which*, *that*, and *zero* separately. Since *zero* appears to be fairly stable across time and space the focus of this paper is the alternation between *which* and *that*, only odds ratios for factors influencing these two are given. The odds ratios always reflect how a predictor influences the probability of seeing the relativization device in question versus the probability of seeing either of the other two alternatives. This relationship explains the fact that both figures are not exactly each other's inverse. A relativizer in subject position of the relative clause, for instance, favors both *which* and *that*, but this is because *zero* is categorically restricted to an object gap in the relative clause.

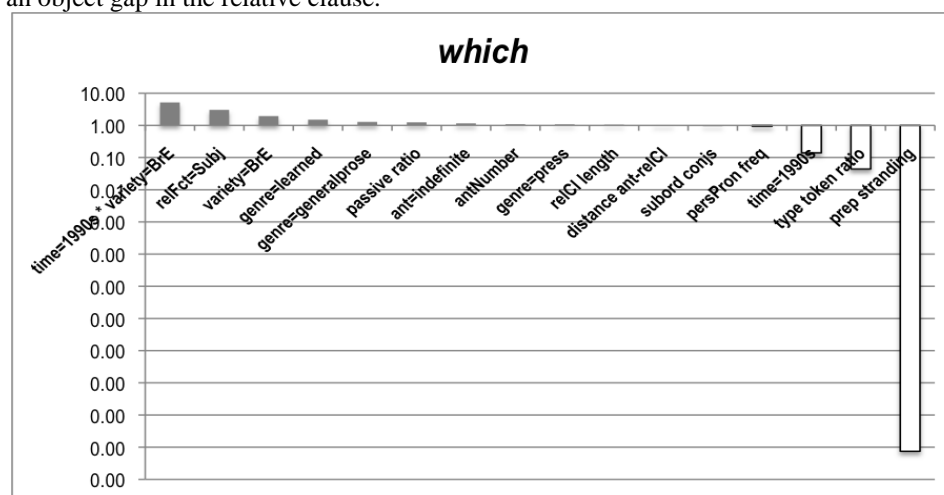


Figure 4. Factors favoring/disfavoring the choice of *which* in restrictive relative clauses (odds ratios from logistic regression). Only significant factors are included in the model.

³² We are looking at written English only here. In spoken English, zero is, of course, a much more frequent choice of relativization (Biber et al. 1999: 609-612).

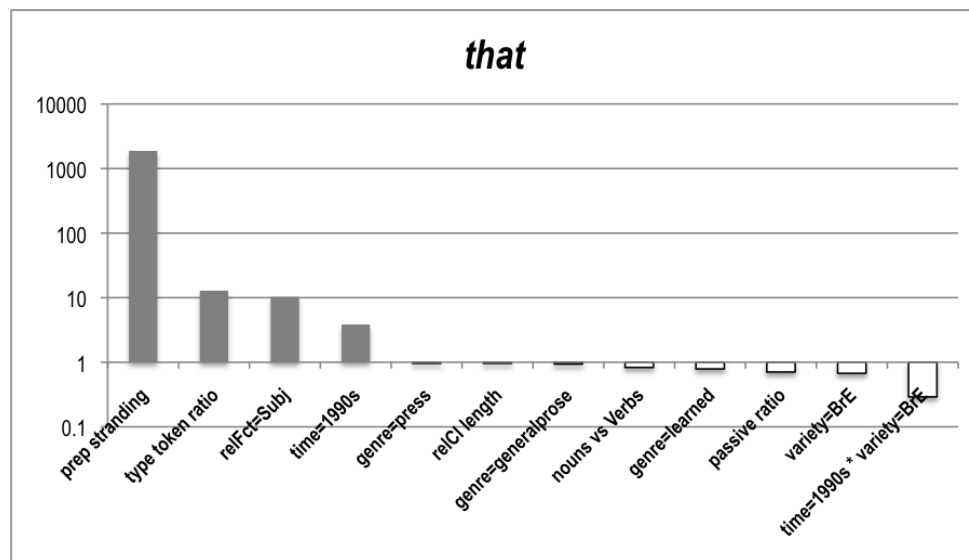


Figure 5. Factors favoring/disfavoring the choice of *that* in restrictive relative clauses.

Figures 4 and 5 include all the predictors that came out as statistically significant ($p < 0.05$) in the respective models. Of all these, we will restrict our discussion in this paper to the ones which can be interpreted as reflecting a prescriptivist style rule. The observations from the raw counts above are reflected in the positive weight British English has as a predictor for *which* and the negative one for *that*. Also, the weight of the interaction term between variety and time in both models underlines the degree to which the developments in the two varieties differ. The ratio of passive versus active constructions lends further support to the hypothesis that prescriptivism is a potential source of these trends. Texts with more passive constructions tend to favor the use of *which* in restrictive relative clauses, whereas texts with fewer passive constructions tend to favor *that*. In other words, if authors adhere to one style rule (“avoid the passive voice”) they are also more likely to follow another one (“that is the defining, or restrictive, pronoun...”).

Another prescriptivist mantra, never to end a sentence with a preposition, is reflected in the predictor frequency of stranded prepositions. This one as well comes out as significant in both the model for *which* and the one for *that*. In fact, it is the predictor with the highest factor weight. Contrary to our hypothesis, though, increasing frequencies of stranded prepositions, i.e. violations of the above rule, favor adherence to the prescriptivist call for using *that* for introducing a restrictive relative clause. This result is surprising indeed, at least under our initial hypothesis, and requires further investigation. Understanding prescriptivism as a unified set of rules with equally strong bearing obviously does not suffice to explain the variation observed in our data. In the present case, the proscription of preposition stranding may be a rule that is on its way out. Rodney Huddleston and Geoffrey Pullum lend support to this speculation when they state with regard to usage manuals that only “some of the more old-fashioned ones still state that ending a sentence with a preposition is incorrect or at least inelegant” (Huddleston & Pullum 2002, p.138).

Neither will it be useful to discuss the realization of these rules in a given text in isolation from other factors. Developments like the one under scrutiny here are always “embedded in particular text types and discourse contexts” (Leech et al. 2009, p.12) which

exert a certain influence on the grammatical and lexical choices of an author and a variety of different broad trends are at work beside, and potentially in competition with, style rules of the prescriptive kind. It is therefore necessary to turn to some of the other predictors in our models and ask how the text features and trends they represent interact with prescriptivism in influencing the choice of relativization strategy.

If we look at the distribution across different genres in our data we see that *which* is favored first and foremost by academic writing, whereas *that* is most prevalent in fiction (the baseline genre, which is why, in the odds ratios table for *that*, all other three genres come out as negative). These two represent poles on a continuum of formality, with academic prose at the formal, “uptight” (Hundt & Mair 1999) end and fiction, which often has a large proportion of direct speech, at the other extreme. *Which* is arguably the most formal relativization device: it is required in elaborate syntactic contexts like pied-piping constructions and largely absent in spoken English (Biber et al. 1999, pp.609-18). Hence it seems plausible to include this axis alongside prescriptivism as an explanatory factor in our model for choice of relativizer. Additional support for doing so is given by the fact that increasing frequencies of personal pronouns, another indicator of less formal and more “involved production” (Biber 1988), disfavor *which* in our data.

7. Discussion

Our investigation of relativization devices in restrictive, non-pied-piping relative clauses in the four core corpora of the Brown family confirms Leech et al.'s (2009) findings for relative clauses on the whole. In fact the rise of *that* at the expense of *which* in American English is put into even starker relief in the context of true variation accounted for in this paper. It seems as if, at least in the United States, advice from usage guides and style manuals has a real bearing on writers' choices, to the extent that *which* has become a very rare option for introducing restrictive relative clauses in American English. In addition, the variables choice of relativizer and frequency of passive constructions are correlated in a way that allows us to interpret adherence to prescriptive rules as a common cause for outcomes in both.

The discussion of our individual predictors, on the other hand, has shown that prescriptivism is far from being a unified and isolated influence, and writers are not slaves to a set of style rules. Rather, they are sensitive to a number of general, genre-specific and discourse-related stylistic conventions and make their choices accordingly. One ongoing change in such conventions is “colloquialization” (Leech et al. 2009, pp.239-49), the move away from formal and towards more oral language in written English. Our analysis shows that this trend is at least partly responsible for the development of relativization strategies in restrictive contexts. It is difficult at times to tease colloquialization and prescriptivism apart exactly since many of the contemporary style guides are written in a spirit of ‘plain English.’

Finally, a note on the ideological dimensions of the discussion around prescriptivism is in place. Studying contributions to the website LanguageLog, one cannot fail to notice the harsh tone with which descriptive linguists often denounce the “rules” of style guide authors and their followers. Arnold Zwicky, for instance, speaks of “a pack of hypocrites or fools” and makes specific reference to usage advisor Bryan Garner as “an idiot” (Zwicky 2006). This kind of rhetoric is even more surprising if one considers that scholars in other disciplines hold rather different opinions of the people criticized by Zwicky. Peter Elbow (2011), for instance, recently recommended Garner's *Dictionary of Modern American Usage* specifically at the 2011 conference on college composition and communication. The point here is not to take sides on the debated issue and prove the

other side as being in the wrong. Rather, it seems that authors from different backgrounds have very different understandings of the role and purpose of usage advice.

We maintain that descriptive linguists including ourselves are well-advised to remember that our perspective is one of several and not necessarily privileged over others. The grounding of our discipline in empirical research can often suggest a greater claim to factuality and disinterested scientific inquiry. This does not mean, however, that we are free from ideology and able to write from a 'positionless position.' In his classic discussion of feminist language prescriptions, Michael Silverstein (1985, p.222) reminds us that "[t]he total linguistic fact is [...] irreducibly dialectic in nature". Speakers as well as, we maintain, linguists are themselves part of this dialectic situation and consequently unable to describe it entirely 'objectively' from the outside. The meaning of any utterance is in large part dependent on the indexicalities connected to it in a given situation. These indexicalities are not universally valid semantic meanings, but context-specific and ideologically mediated ones, and they work "all the way down" (Silverstein 1998, p.138); cf. (Blommaert 2005). That is, we can never strip an utterance of all its indexicalities until only the True, universal meaning remains.

What this means for the present context concretely is that a re-evaluation of our object of study as well as a serious engagement with the goals and purposes of usage manuals may be necessary. As for our own research, the question is warranted whether the habitual claim that we describe "the English language" is entirely justified. Even with much larger corpora than the ones used in the present paper, we only have a minute subset of the entirety of language produced by speakers of English at hand, and one which is usually stripped of much of its social context. There are, of course, procedures to ensure the representativeness of such a subset, but they remain imperfect by definition. Usage advisors, on the other hand, seem less interested in "the language." They frequently emphasize the non-categorical nature of their guidelines as well as their validity in a well-defined institutional space. In other words, they seem to be interested first and foremost in maintaining discourse conventions for specific genres. As much corpus linguistic work (Leech et al. 2009; Mair 2006; Biber 2003; Hundt & Mair 1999; Biber & Finegan 1989) has shown, focus on the development of different discourse conventions can be a very rewarding line of inquiry for linguists interested in ongoing changes in the English language. At the very least, then, we should stop denouncing usage guides as a priori distorting, stupid and downright wrong and instead take seriously the influence prescriptivism has on changing norms of discourse. The present paper constitutes an attempt to make a step in this direction.

Undoubtedly, much remains to be done. As we continue to investigate individual usage rules the question to what extent they can be meaningfully subsumed under an umbrella term "prescriptivism" remains to be answered. Similarly, procedures for isolating prescriptivism from other factors, as for instance colloquialization in the present paper, need to be developed. In addition to refining our statistical models and increasing the size of our database, a crucial step will consist in conducting meta-studies in order to determine the institutional sites at which the rules in question are most actively disseminated and enforced. Perception experiments in universities in the UK and the USA are one project we envisage. Similarly, an analysis of the changing style sheets and policies of individual editorial boards may provide rewarding results for contextualizing the probabilistic developments extracted from corpus linguistic analyses.

References

- Auer, A., 2006. Precept and practice: The influence of prescriptivism on the English Subjunctive. In D. Kastovsky, ed. *Syntax, style and grammatical norms: English from 1500-2000*. p. 33.
- Auer, A. & Gonzalez-Diaz, V., 2005. Eighteenth-century prescriptivism in English: A re-evaluation of its effects on actual language usage. *Multilingua*, 24(4), pp.317-341.
- Ball, C.N., 1996. A diachronic study of relative markers in spoken and written English. *Language Variation and Change*, 8(02), p.227–258.
- Biber, D., 2003. Compressed noun-phrase structures in newspaper discourse: The competing demands of popularization vs. economy. In J. Aitchison & D. Lewis, eds. *New Media Language*. London/New York: Routledge, pp. 169-181.
- Biber, D., 1988. *Variation across speech and writing*, Cambridge: Cambridge University Press.
- Biber, D. & Finegan, E., 1989. Drift and the evolution of English style: A history of three genres. *Language*, 65(3), pp.487-517.
- Biber, D. et al., 1999. *The Longman Grammar of Spoken and Written English*, London: Longman.
- Blommaert, J., 2005. *Discourse: A Critical Introduction*, Cambridge, UK: Cambridge University Press.
- Busse, U. & Schroeder, A., 2010. Problem areas of English grammar between usage, norm, and variation. In A. Lenz & A. Plewnia, eds. *Grammar between norma and variation*. VarioLingua. Frankfurt: Peter Lang, pp. 87-102.
- Crystal, D., 2008. *Dictionary of linguistics and phonetics* 6th ed., Malden, MA: Wiley-Blackwell.
- D’Arcy, A., 2010. Ode to a Prescriptivist. *OUPblog*. Available at: <http://blog.oup.com/2010/02/prescriptivist/> [Accessed March 23, 2011].
- Elbow, P., 2011. Casual, everyday unplanned spoken language: What can it offer careful writing? - Even academic writing?
- Facchinetti, R., 2000. The modal verb shall between grammar and usage. In D. Kastovsky, ed. *The History of English in a Social Context: A Contribution to Historical Sociolinguistics*. Berlin, New York: Mouton de Gruyter, pp. 15-1.
- Fowler, H.W., 1965. *A Dictionary of Modern English Usage* Reprint. E. Gowers, ed., Clarendon Press.
- Hinrichs, L., 2006. Prescriptivism creeping in through the back door: The “sacred that-rule” and relative pronoun variation in British and American English.
- Hinrichs, L., Smith, N. & Waibel, B., 2010. Manual of information for the part-of-speech-tagged, post-edited “Brown” corpora. *ICAME Journal*, 34, pp.189-231.
- Hinrichs, L. & Szmrecsanyi, B., 2007. Recent changes in the function and frequency of Standard English genitive constructions: a multivariate analysis of tagged corpora. *English Language and Linguistics*, 11(3), pp.437-474.
- Huddleston, R. & Pullum, G.K., 2002. *The Cambridge Grammar of the English Language*, Cambridge: Cambridge University Press.
- Hundt, M. & Mair, C., 1999. “Agile” and “uptight” genres: The corpus-based approach to language change in progress. *International Journal of Corpus Linguistics*, 4, pp.221-242.
- Leech, G. et al., 2009. *Change in Contemporary English: A Grammatical Study*, Cambridge: Cambridge University Press.
- Lieberman, M., 2006. The direct and vigorous hyptic voice. *Language Log*. Available at: <http://itre.cis.upenn.edu/~myl/language-log/archives/003417.html> [Accessed March 23, 2011].

- Liberman, M., 2011. The linguistic narcissism of Christopher Hitchens. *Language Log*. Available at: <http://languagelog.ldc.upenn.edu/nll/?p=3014> [Accessed March 23, 2011].
- MacKay, D.G., 1980. On the goals, principles, and procedures for prescriptive grammar: Singular they. *Language in Society*, 9(03), pp.349-367.
- Oxford English Dictionary, prescriptivism. *Oxford English Dictionary*. Available at: <http://www.oed.com.ezproxy.lib.utexas.edu/Entry/150654>.
- Partridge, E., 1957. *Usage and Abusage*, Baltimore: Penguin.
- Peters, P. & Young, W., 1997. English grammar and the lexicography of usage. *Journal of English Linguistics*, 25(4), p.315.
- Pullum, G.K., 2009. 50 Years of Stupid Grammar Advice. *The Chronicle of Higher Education*, p.B15.
- Silverstein, M., 1985. Language and the culture of gender: At the intersection of structure, usage and ideology. In E. Mertz & R. Parmentier, eds. *Semiotic Mediation*. Orlando: Academic Press, pp. 219-259.
- Silverstein, M., 1998. The use and utility of ideology: A commentary. In B. Schieffelin, K. Woolard, & P. Kroskrity, eds. *Language Ideologies: Practice and Theory*. New York: Oxford University Press, pp. 123-145.
- Strunk, W. & White, E.B., 1999. *The Elements of Style* 4th ed., Longman.
- Zwicky, A., 2006. Language Log: Five more thoughts on the That Rule. *Language Log*. Available at: <http://itre.cis.upenn.edu/~myl/languagelog/archives/002189.html> [Accessed May 31, 2011].
- Zwicky, A., 2009. Prejudices, egocentrism, impositions, and intransigence. *Language Log*. Available at: <http://languagelog.ldc.upenn.edu/nll/?p=1354#more-1354> [Accessed March 23, 2011].

axel@bohmann.de

ADDRESSES