Language prescriptivism : attitudes to usage vs. actual language use in American English

Kostadinova, V.

Author: Kostadinova, V.
Title: Language prescriptivism : attitudes to usage vs. actual language use in American English
Issue Date: 2018-12-18

Language Prescriptivism

Attitudes to usage *vs.* actual language use in American English

Printed by

Ridderprint BV
Pottenbakkerstraat 15
2984 AX Ridderkerk
The Netherlands

+31 (0)180 463 962
www.ridderprint.nl

Cover design: Olga Kepinska and Viktorija Kostadinova.
Cover illustration: A keyword analysis of the dissertation text produced with WMatrix.

Language Prescriptivism

Attitudes to usage *vs.* actual language use in American English

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op 18 december 2018
klokke 12:30 uur

door

Viktorija Kostadinova

geboren 3 november 1986
te Kochani, Macedonië

Promotores:	prof. dr. Ingrid Tieken-Boon van Ostade
	prof. dr. Anne Curzan (University of Michigan)
Promotiecommissie:	prof. dr. Mike Hannay (Vrije Universiteit Amsterdam)
	prof. dr. Raymond Hickey (University of Duisburg and Essen)
	prof. dr. Maarten Mous
	dr. Dick Smakman
	prof. dr. Marina Terkourafi

За тато

Contents

Acknowledgements

I would first like to thank the people who have helped most directly with the content of this dissertation. I am grateful to my promotor, Ingrid Tieken-Boon van Ostade, for the opportunity to work on this project, for her help, and for valuable discussions throughout the years. I thank the members of the Bridging the Unbridgeable project for the stimulating discussions during our project meetings. I would also like to acknowledge the NWO funding provided for the project. I thank Anne Curzan for her inspirational course at the LSA Linguistic Institute in 2013, as well as her comments on the final manuscript. I am grateful to the members of the committee, Mike Hannay, Raymond Hickey, Maarten Mous, Dick Smakman, and Marina Terkourafi for their evaluation and useful feedback.

I thank all my respondents and friends in Los Angeles. I especially want to thank those who helped me find respondents in the first place: Elizabeth Conn, her boyfriend, Aaron, and her parents, Cathi and Richard, for their interest in my research, as well as their kindness, generosity, and invaluable support. I *literally* would not have been able to do this without them.

I would like to thank the many friends and colleagues who I've met at conferences, and with whom I've discussed my project. Those discussions proved to be a panacea for the pain of not knowing where the project was going. I particularly want to thank Folgert Karsdorp for introducing me to Python, and for thus providing the trigger for my data analysis; Marten van der Meulen for being both an enthusiastic and critical debater on issues of shared interest; and Mark de Rooij for statistical help. I thank Elly Dutton for (proof)reading the n-th final draft of my thesis and convincing me that the end was in sight.

I thank everyone at LUCL for all the help, support, drinks, talks, and the rare but important chats at the coffee machine. I am especially appreciative of my time with the other members of the PhD Council, particularly Andreea, Bora, Elly, and Marieke. I would like to thank Niels Schiller, Gea Hakker, and all other former and current members of the LUCL management team and secretariat for always being helpful and friendly.

My colleagues at the University of Amsterdam have been continuously encouraging in the final year of the process. I thank them all for their interest in my work and for helping in various ways, which has meant a lot to me. In addition to all of this, Rina de Vries kindly offered to do me the huge favour of taking care of the Dutch summary of my thesis.

I am beyond grateful to my friends and paranymphs, Elly and Olga, for everything; from our early day shenanigans, through all the 'hangage' and 'cuaffage', to their support and encouragement in the last stage of this process. Words cannot describe (it's a cliche cos it's true) how much they have meant and done for me. Bobby has been a great friend and colleague, and I thank him as much for the laughs as for his help with LaTeX and statistics. Jamie's conversations were distracting, enriching and always fun. My Skopje friends Sashe Nikolovski, Sashe Tasev, Slavcho, Kata, и компанија, have given me unwavering support and encouragement over the years, and have enriched my life with wonderful visits, trips, parties, and discoveries.

Мајче, ете со што се мачев последните шест години. Ти благодарам за поддршката, љубовта и бескрајната верба во мене.

CHAPTER 1

Language prescriptivism

1.1 Introduction

The view of language prescriptivism as a phenomenon which originated in the eighteenth century and is, at best, only marginally related to processes of linguistic variation and change can still be considered particularly prevalent in descriptive linguistics. For instance, compared to the prescriptive and normative grammars of the eighteenth and nineteenth centuries, twentieth-century grammar writing is strikingly different. While this paradigm shift in English grammar writing was completed long ago, and the most influential reference grammars of English, such as Biber et al. (1999) or Huddleston and Pullum (2002), are descriptive grammars of English, written by linguists and based on linguistic scholarship, the normative tradition of prescribing correct language use has survived to this day. In other words, prescriptivism is "alive and well" (cf. Beal 2009: 47). Perhaps the most persistent and entrenched twentieth-century manifestation of this tradition is the genre of 'usage guides' (Tieken-Boon van Ostade 2010). These books of language advice – variously referred to as usage guides, handbooks of usage, usage manuals, and usage books – first appeared in England at the end of the eighteenth century (Leonard 1929; Tieken-Boon van Ostade 2010). A similarly popular practice of publishing books

on language advice arose in America during the nineteenth century, when the first American usage guide was published (Connors 1983; Tieken-Boon van Ostade 2015), resulting in an equally productive and long-lasting tradition. Since then, the number of titles published and copies sold has grown in both countries. William Strunk Jr. and E.B. White's *The Elements of Style*, undoubtedly the most popular American usage guide, is considered a classic, and has become part of popular culture. Since its publication in 1959 – the version revised by E.B. White – over ten million copies have been sold (Roberts 2009; Pullum 2010a).

On the other side of the Atlantic, an often cited example of the success such publications can achieve is Lynne Truss's guide to punctuation, *Eats, Shoots & Leaves: The Zero Tolerance Approach to Punctuation* (2003), which has sold in excess of two million copies (Curzan 2014: 41). The majority of the advice which these books provide has in general been met with strong criticism from linguists (cf. Burchfield 1991). Pullum (2010a: 34), for instance, argues that "the success of [Strunk and White's] *Elements* [is] one of the worst things to have happened to English language education in America in the past century", and that "the people who rely on it have no idea how badly off-beam its grammatical claims are". *Eats, Shoots & Leaves* provoked a similar reaction, this time in the form of a book-length historical account of attitudes to English usage by Crystal (2006). In his treatment of the subject, Crystal contextualises the historical development of attitudes to usage, highlighting the futile attempt of usage commentators to restrain, control, or "fix" the English language. Intertwined with the historical account of the development of attitudes to usage through the centuries are arguments about topics such as the nature of language variation and change, or the crucial role of situational context in discussions of language standards and norms (Crystal 2006: 152). These are arguments which linguists often use to point out the failures and shortcomings of language advice found in usage guides (e.g. Pullum 2010a,b). Linguists often perceive these prescriptive approaches to usage to be acutely misleading with respect to the nature of language use and linguistic change. This is often pointed out with reference to the strikingly different meanings that certain basic linguistic terms, such as "grammar", "language", and "rule" (cf. Curzan 2014), have in general language use. Lamenting the confusion surrounding the understanding of what grammar is, Leech et al. (2009: 1), note that "[r]ather than see grammar as the vast and complex system of rules which helps us organize words into constituents, clauses and sentences, [among usage commentators] the term is restricted to refer to a collection of variable and disputed usages which have been selected arbitrarily in the course of almost 300 years of prescriptive thinking

about good grammar and proper English".

However, this does not tell the entire story about usage guides. Despite the problems in their approach to language use and the falseness of many of these claims, there are three other important – and perhaps less controversial – characteristics of the usage guide tradition which are worth pointing out. The first one is its stability and entrenchment in contemporary society. This stability is both historically conditioned, by the long tradition of the genre, and synchronically relevant, as is evident from the continuing popularity of usage guides. Thus, despite well-documented historical accounts of the linguistically arbitrary nature of prescriptive language ideologies and their guaranteed failure in the face of inevitable processes of language change, "[m]anuals of English have sold well for generations" (Crystal 2006: viii), and new titles continue to appear. There have been at least thirty new usage guides published in America since the turn of the century. The American usage guide genre now numbers more than two hundred titles (see Section 4.3), and, as Creswell (1975: 1–2) noted almost forty years ago, "[u]sage guides remain the go-to source of authority on matters of language usage". Judging by the number of guides which have appeared since, his observation remains equally applicable nowadays.

The second characteristic of the usage guide tradition, which is perhaps related to the relatively large number of publications, is that not all usage guides are the same, or indeed equally prescriptive. In fact, it has been recognised that the dichotomy between prescriptivism and descriptivism is never neatly manifested in practice. Cameron (1995: 3–5), for instance, argues convincingly that prescriptivism and descriptivism can be seen as mirroring similar ideological positions, and are often intertwined in actual language use (see also Pinker 2014: 188–189). This is supported by scholarship on eighteenth-century normative grammars which shows that, contrary to the stereotypical notion, some normative grammars are not exclusively prescriptive, and that they often feature prescriptive and descriptive approaches side by side (e.g. Hodson 2006; Tieken-Boon van Ostade 2006; Straaijer 2009). Much the same applies to usage guides. Although criticism has been directed at some usage guides, others have received positive reviews from linguists. This is perhaps more readily noticeable in relation to empirically based usage guides, such as *Webster's Dictionary of English Usage* (cf. Algeo 1991a).

The third relevant characteristic of usage guides is their value. Linguists have often recognised the indispensable social function of good usage guides (cf. Weiner 1988: 182–183), as well as the value of usage guides in general, despite the level of misguidedness of some of their language advice. Crystal, for instance, recognises that

despite the fact that usage guides fail to "fix" the language, or teach speakers how to use language effectively,

> they do have a value. They help to alert us to the issues of change that worry the more conservatively minded members of society. They also perform a valuable service in drawing attention to those features of language where it is all too easy to be lazy or careless, and where sense or intelligibility suffers as a result. (2006: 157)

Thus, beyond the obvious social usefulness of usage guides, they have a rather subtle and linguistically relevant value, in that they inevitably contain information about linguistic variants, even when some of these variants are the subject of criticism. It is these aspects of the usage guide tradition which serve as the starting point for my investigation of the American usage guide tradition and its role in perpetuating prescriptive attitudes to language, its influence on language users and language use, and its relationship with linguistics. The question of the influence of usage guides is thus a question of the influence of prescriptive ideology in general. In the introductory chapter, I will discuss the importance of studying this question, and will elaborate on the motivations and assumptions of the present study.

I am here concerned specifically with American English, and I will at times make references to comparisons between this variety and British English. The reason for this is that the American usage guide tradition is distinct from the British tradition, and also from the growth of English as a global language. It is of course important to keep in mind that this focus on one specific variety is not meant to imply that similar observations would necessarily apply to other contexts, such as World Englishes. For instance, Inner Circle varieties, in terms of the model of Kachru (1992), are different from Outer Circle and Expanding Circle varieties, in terms of norms, ideologies, and socio-historical context. For instance, one potential problem that might immediately be identified is the fact that the intended readership of usage guides is *native* speakers of the variety in question, as the spread of English beyond the Inner Circle has shown that the concept of the native speaker may no longer account for the types of speakers of English that exist in these 'new' contexts (cf. e.g. Singh 1996). However, addressing the question of prescriptivism, usage guides, and language advice literature in relation to Outer and Expanding circle contexts in any reasonable depth is beyond the topic of this dissertation (however, see Cameron 2012 for an interesting discussion of language prescription in the context of Global English).

1.2 Defining prescriptivism

Prescriptivism is related to "an ideology (or set of beliefs) concerning language which requires that in language, as in other matters, things shall be done in the 'right' way" (Milroy and Milroy 2012: 1). Fundamentally, prescriptivism is an approach, or a response, to variation and change in language, in that it sets out to counteract that change. As such, it is part and parcel of the process of standardisation, which involves, in the words of Milroy and Milroy (2012: 6), "the suppression of optional variability in language". In other words, from the point of view of prescriptive ideology, linguistic variants are strongly, almost viscerally, associated with values. At the most fundamental level, those values are 'good' and 'bad'; extended to language, additional values refer to such things as standardness, correctness, appropriateness, clarity, and legitimacy. The use of *ain't* is an example of this; it is an incorrect form, and is not recognised as a legitimate English word, because it is not a variant which is acceptable in standard English (see Chapter 5). The instinct to regulate which expressions legitimately belong to a language and which do not means that prescriptivism "implies above all, authority; it also implies order, stability, predictability and reason" (Drake 1977b: 1). The attempt to regulate the value of linguistic variants on a micro-level corresponds to an attempt to regulate language variation and change on a macro-level, by enforcing a doctrine of correct language use. "Adherents to the doctrine of correctness", Finegan (1980: 10) writes, "strive to mold linguistic practice according to selected patterns of grammar; they attempt to retard the pace of language change or halt it altogether". In the history of English, the ideology of language prescriptivism is traditionally associated with English normative grammars written during the eighteenth and nineteenth centuries in the United Kingdom and the United States (e.g. Leonard 1929; Tieken-Boon van Ostade 2000, 2006), although it has also been recognised that prescriptivism survives in present-day English as a characteristic of usage guides (e.g. Crystal 2006; Albakry 2007; Tieken-Boon van Ostade 2010).

The term "prescription" is also often used alongside the term "prescriptivism", and there seems to be some confusion as to what is usually meant by these two terms. While "prescriptivism" seems to be used to refer to prescriptive ideology in general, regardless of the time period in which this ideology is found to be manifested, the term "prescription" is used by Milroy and Milroy (2012) to refer to the latest stage in the English standardisation process, which follows the stage of codification (although the

boundaries between stages are not clearly delineated). For instance, in the Milroys' discussion of the stages of standardisation, they use the term "prescription" (see also Tieken-Boon van Ostade 2008b, 2016). While Tieken-Boon van Ostade (2016) argues that prescription and prescriptivism should be distinguished as important stages in the standardisation process of English, the difference between the two remains unclear, mostly coming down to arguments related to the negative connotations associated with the term "prescriptivism". For the present study, I distinguish between prescription and prescriptivism; by "prescription" I understand the stage in the standardisation process which follows codification (in line with Milroy and Milroy's model), while by "prescriptivism" I understand the ideology according to which language use should adhere to specific and clearly defined norms and standards of correctness and appropriateness, while language change should be resisted. In this view, prescriptivism can be seen as characterising an approach to language which can be found both in the codification and in the prescription stages. These stages are associated with different functions of prescriptivism. Perhaps the most fundamental difference in the function of prescriptivism is that during the codification stage prescriptivism is part of the process of codifying the language standard and establishing the norms of the language, while during the prescription stage prescriptivism has the function of maintaining the language standards and norms. The codification and prescription stages are also associated with different kinds of metalinguistic works, i.e. normative grammars and usage guides respectively.

In addition to determining its ideological characteristics, a useful and often applied approach to defining prescriptivism is to compare it with descriptivism. Descriptivism is an approach which characterises linguistics, i.e. the objective study of language as it is, not as it should be (cf. Huddleston and Pullum 2002: 2–3; Baker and Hengeveld 2012: 19–20). Descriptivism is usually associated with twentieth-century linguistics, a scientific discipline diametrically opposed in outlook and approach to the eighteenth-century grammarians' prescriptive conception of language. Consequently, it is normally assumed in mainstream linguistics that prescriptivism and descriptivism are entirely at odds with each other. Compared to prescriptivism, "[d]escriptivism emphasizes change over stability, diversity over uniformity, usage over authority, and the spoken language over the written language" (Drake 1977b: 1). If prescriptivism is inextricably linked with what Finegan (1980) refers to as the doctrine of correctness, a term first adopted by Leonard (1929), descriptivism is associated with the doctrine of usage. Proponents of the descriptivist view on language variation and change "make no explicit value judgements about the logic, utility, or aesthetics – i.e. the

'correctness' of particular lexical or grammatical items, but report the known facts about the ways in which a given form, meaning, or pronunciation is actually used and in what circumstances" (Finegan 1980: 11).

While this neatly established dichotomy is still the predominant way of thinking about prescriptivism, the way that linguists refer to or account for prescriptivism in descriptive studies is not always straightforward. It is perhaps partly due to the entrenched opposition between prescriptivism and descriptivism, or to the strong commitment to descriptivism in linguistics, that linguistic accounts of prescriptivism come in various kinds and at various levels of specificity. Concretely, in relation to the question of the influence of prescriptivism, statements tend to differ; sometimes they are contradictory, and often they are vague. For simplicity, I group existing accounts of the relationship of prescriptivism to language into three kinds of positions with respect to the ways in which prescriptivism is conceptualised and its influence accounted for in descriptive linguistic studies. While this does not represent the entirety of scholarly positions or assumptions about prescriptivism, it will serve as a useful generalisation of a diverse set of statements in the context of the present discussion. It is crucial to identify these three positions at the outset of this study, because they show how prescriptivism has been conceptualised as a phenomenon in relation to language, and, specifically, what that means for the question of the influence of prescriptivism.

The first position assumes that prescriptivism is diametrically opposed to the objective study of language, and that, as such, prescriptive ideology is fundamentally mistaken in its conception of language. Consequently, prescriptive ideology is seen as having no role to play in language development. This stance on prescriptive ideology is representative of the stance on language ideology in general. Kroskrity (2004: 499) traces the "marginalization or proscription of linguistic ideology" back to the beginning of the twentieth century, and the development of the new science of linguistics. "Speakers, through their linguistic ideologies, were neither part of language nor capable of being agents of linguistic change", Kroskrity (2004: 499) writes, and continues by saying that "[r]ather than being viewed as partially aware or as potentially agentive, speakers – in Chomskyan models – were merely hosts for language". This meant that linguistic ideology in general, and consequently prescriptive ideology, was not taken into account at all, or, when it was, as, for instance by Bloomfield (1927, 1944), it was ultimately concluded that "speakers' linguistic ideologies – even those cast as prescriptive norms – had a negligible effect on their actual speech" (Kroskrity 2004: 499). Cameron (1995: 3) makes a similar observation in relation to prescriptive ideology, noting that in linguistics "the evaluative concerns

of speakers (embodied in their 'prescriptivism') are by implication seen as both alien and perverse". This dissociation from prescriptive ideology is very often stated in descriptive grammars (e.g. Huddleston and Pullum 2002: 2–5) and linguistics textbooks (e.g. Fromkin 2000: 20; Baker and Hengeveld 2012: 19–20).

The second position differs from the first in that it maintains that prescriptivism is an important sociolinguistic factor. This position is formulated and discussed in a number of recent critical treatments of language prescription and prescriptivism (Milroy and Milroy 1985; Cameron 1995; Curzan 2014), which argue convincingly that prescriptivism should not be written off as a narrow view of language use which is common among a small group of people who distinguish themselves by a fondness for rules and an ignorance of how language works. Milroy and Milroy (1985), as noted above, argue that prescription is a stage in the standardisation process of English, and that it follows the codification stage. Cameron (1995) approaches prescriptivism as the natural tendency for societies to regulate and establish standards and norms of language use; she calls this broadly defined phenomenon "verbal hygiene". More recently, Curzan (2014) significantly furthers the discussion of prescriptivism and its influence by reformulating the main question at the centre of most discussions of prescriptivism, which usually focus on the failure of prescriptivism to achieve what it purports to do, i.e. stop language from changing. As Curzan points out, the important question when studying prescriptivism is not its success or failure, but its essential nature, its importance in the context of metalinguistic discourses, and the effects it has on language users and language itself. "[M]aking value judgments on language", Cameron (1995: 3) argues, "is an integral part of using it and not an alien practice 'perversely grafted on'". This position assumes that "[a] prescriptive attitude has nevertheless played a noticeable role in shaping the English language" (Chafe 1984: 96), and, furthermore, that this assumption needs to be studied in practice. The majority of empirical investigations of this question have been conducted by historical sociolinguistic scholars focusing on prescriptivism in the eighteenth and nineteenth centuries (e.g. Dekeyser 1975; Tieken-Boon van Ostade 1982, 1994, 2006, 2000, 2002; Chafe 1984; Tottie 1997; Auer and González-Díaz 2005; Auer 2006; Anderwald 2012, 2014; Yáñez-Bouza 2015). These works include those which study both the establishment and rise of normative and prescriptive language rules (for a discussion of the difference between these two, see Vorlat 1979) during the eighteenth century, and their influence during the eighteenth and the nineteenth centuries. This reasoning has been extended to present-day English; in this context, prescriptivism is understood to have an influence on both language practice and language speakers, and

it is argued that this influence should be studied empirically.

The third position is somewhere between these two, and is the least well defined. It is usually not strongly or explicitly expressed, and is found in cases where either the importance of prescriptivism is recognised, but not discussed in further detail, or where prescriptivism is used as an explanatory tool for processes of language change, although this is also often not explained precisely. A good illustration of this position can be found in Calle-Martín and Miranda-García's (2009) diachronic study of the split infinitive, spanning the period from 1640 to 1920. In it, they observe that the split infinitive is a construction against which there is a strong prescriptive bias, "a fact which makes [split infinitives] practically disappear from corpus data, particularly in the case of consciously edited texts" (2009: 349). In making this observation, the authors implicitly recognise the influence of prescriptive ideology on edited written English. However, this raises additional questions about what exactly constitutes strong prescriptive bias – as opposed to weak prescriptive bias – and how its influence could be identified or measured rather than being merely assumed. Another example of the use of prescriptive influence as an explanatory tool is found in a study of grammatical change in twentieth-century written English, where, on the basis of corpus data, Leech and Smith (2009: 196) identify "a steeper decline of the passive in [American English] possibly due to prescriptivism". More specifically, they note that "[a]n additional reason for a passive decline, probably increasing through the century, has been the hostility (especially in the US) of prescriptive forces – including usage gurus, house style manuals, crusaders in favour of 'plain English', and latterly, grammar checking software – all either overtly or covertly disparaging the use of the passive" (Leech and Smith 2009: 183). They offer a similar explanation for the decrease of the frequency of *which*, as opposed to *that*, in American English (Leech and Smith 2009: 181). A third example of an implicit statement about the potential influence of prescriptivism on language use can be found in a study of the de-grammaticalisation of the infinitive marker *to* (Fitzmaurice 2000b). Fitzmaurice selects infinitives split by *not* as a case study for the investigation of the process of de-grammaticalisation, rather than infinitives split by adverbs; in relation to the latter type of split infinitives, she observes:

> This construction is a traditional bugbear of traditional grammarians, and therefore its high profile militates against an objective assessment of the progress of the de-grammaticalisation of the infinitive marker. By contrast, the negative split infinitive – the construction split by a negative operator – has received rather less attention from prescriptivists and has

> thus remained less obtrusive in speakers' conscious linguistic behaviour. This type of split infinitive therefore seems worth further investigation. (Fitzmaurice 2000b: 172)

This reasoning seems to imply that there is some kind of influence of traditional prescriptivism on the use of the split infinitive in speakers' conscious linguistic behaviour, which in turn makes it a problematic candidate for an investigation of linguistic processes such as grammaticalisation, which operate below the level of consciousness. While, of course, this is not explicitly stated, such a reading is not impossible on the basis of Fitzmaurice's observations. A more plausible interpretation would be that the author does not believe that the process of de-grammaticalisation is influenced by prescriptivism, but that she is avoiding the issue altogether, in case there might be some influence. It remains unclear what exactly is meant by the argument that "its high profile militates against an objective assessment of the progress of the de-grammaticalisation of the infinitive marker". In other words, it is unclear whether "its high profile" works against an objective assessment because traditional prescriptive grammar may have somehow 'contaminated' this process, or whether it is merely a safer option for the analyst to deal with features which speakers may not be as aware or conscious of as they would be expected to be in the case of prescriptively targeted features. Furthermore, Fitzmaurice argues that, because negative split infinitives have not been the subject of prescriptive usage commentators, it can be assumed that speakers will not be affected by prescriptive ideas about their incorrectness. Establishing this relationship between the two implies that prescriptive usage commentators are influential in disseminating prescriptive ideas about language use. In a discussion of the contributing factors for the disappearance of zero adverb forms, Tagliamonte (2012: 227–228) entertains a potential prescriptive influence which could explain "why zero adverbs in the United States, and perhaps elsewhere in North America, have endured longer than in the United Kingdom", even though she is more tentative in accepting this as the only influence. In all of these cases, potential prescriptive influence is recognised in the context of patterns of language variation and change.

Observations about potential prescriptive influence are also found in studies of speakers' language judgements and intuitions. For instance, in a review of methodological practices adopted in the collection of grammaticality judgement data, Buchstaller and Corrigan (2011) note the potential influence of normative prescriptive ideology when speakers are asked to 'translate' a standard language sentence into their vernacular during data collection. The authors note that this method "presupposes a

situation of relatively low prescriptive pressure in which informants are comfortable providing the dialectal equivalent of the 'standard'" (Buchstaller and Corrigan 2011: 32). This statement is an implicit recognition of the potential influence of prescriptive ideologies among speakers; in this case, such an influence is seen as a distortion of speakers' natural, or real, language use. These examples are of course not exhaustive, and are meant to illustrate some of the issues raised in descriptions or accounts of prescriptive influence which belong to the third of the three positions identified here. At the same time they raise a number of questions about the nature of prescriptive influence, the importance of accounting for this influence systematically, and the problem with relying on prescriptivism as an explanation in processes of variation and change.

The difficulty in formulating straightforward answers to these questions is one of the motivations for the present study. The study is part of a broader research project, conducted by Ingrid Tieken-Boon van Ostade, which also deals with the stages of prescription and the prescriptivism associated with it in the context of British English (Ebner 2017) and in the context of media discourse (Lukač 2018). It is also a continuation of the work done on the emergence and rise of the normative grammar tradition (cf. Tieken-Boon van Ostade 1982, 1994, 2000, 2002, 2012c), in the sense that it attempts to evaluate the possible influence of the norms which started to become established in that tradition. In other words, this study attempts to investigate how prescriptivism is manifested after the codification stage, that is, during the prescription stage. This study will thus explore the nature of twentieth-century prescriptivism and its influence on language variation and change, and language speakers in the context of American English. The present study is based on three important principles, formulated by Curzan (2014), which form the basis for many of its main assumptions. The principles are the following:

- The history of the English language encompasses metalinguistic discussions about language, which potentially have real effects on language use.
- The history of the English language encompasses the development of both the written and the spoken language, as well as their relationship to each other.
- The history of the English language encompasses linguistic developments occurring both below the level of speakers' conscious awareness – what is sometimes called "naturally" – and above the level of speakers' conscious awareness. (Curzan 2014: 48)

On the basis of these three principles, I will address the question of prescriptivism, specifically with respect to the twentieth century, from three perspectives. The first perspective involves the study of usage guides as a source of metalinguistic discourses. The second perspective deals with the question of empirically investigating and ascertaining the potential influence of prescriptivism in spoken and written American English. This will be investigated by means of a corpus-based analysis of the patterns of actual use of a select number of language features (described in detail in Chapter 3). Finally, data on speakers' attitudes will be brought to bear on the question of the influence of prescriptive ideology on speakers' conscious language behaviour and language practices. In what follows, I discuss in detail the importance of studying these three perspectives, as well as the initial assumptions adopted in the study.

1.3 Usage guides

As mentioned above, some scholars have noted the great significance of usage guides, and the potentially considerable influence that such guides may have on language use (Landau 2001; Crystal 2006). The first principle formulated by Curzan (2014: 48; cf. Section 1.2 above) stresses the importance of metalinguistic discussions, and their potential influence on both language use and language users. Commenting on the attitudes expressed in these books, which are often dismissed by descriptive linguists as irrelevant, Landau observes:

> While linguists may deplore the attitudes expressed in usage guides, there is no doubt that such books are popular, and with good reason. The attitudes of others towards one's own language must be considered seriously by anyone who hopes to achieve practical goals. To deplore such attitudes, to argue that such attitudes ought not to exist, is to indulge in fancy and usually means that one is fortunate enough not to need ambition, but wants to show one's sympathy with those who do. Since those who are ambitious and insecure are the great believers in prescriptive attitudes and buy the books that perpetuate them, scholars who are scornful of such attitudes must realize sooner or later that they are addressing only each other. (2001: 262–263)

So, while the influence of usage guides on language use and language users can be assumed to exist, the nature, degree, and mechanisms of this influence have not often

been empirically assessed. If we take these books as a site for the perpetuation and strengthening of metalinguistic discourses and prescriptive attitudes to usage, it is important to provide a more detailed account of the genre and its treatment of usage features.

The first reason for investigating usage guides in the context of this study stems from the fact that the question of the influence of usage guides on language use is a question of the influence of prescriptivism in general. However, before I attempt to investigate the influence of prescriptivism, I turn to a consideration of prescriptive ideology itself. Usage guides present a reliable source for the investigation of metalinguistic discourses, because they are a stable and long-lasting genre, which has played a consistent social function throughout the twentieth century. This allows for an investigation of attitudes to language use, prescriptive or otherwise, as well as, crucially, how these attitudes have changed over time. It is important to take this dimension into account, because existing accounts of prescriptive influence often assume that prescriptive attitudes are unchanging and monolithic. For example, in assessments of prescriptive influence on the split infinitive, scholars usually start from the assumption that split infinitives are proscribed, and then use that as a basis for making judgements about the success or lack thereof of this proscription on the basis of corpus data (cf. the discussion of Calle-Martín and Miranda-García (2009) and Fitzmaurice (2000b) above). In other words, if corpus data show that split infinitives are increasing, this would be interpreted as evidence that the proscription has not been effective, while a decrease in the frequency of use of split infinitives would be considered to be evidence that the proscription has been effective. However, as I will show in later chapters, this approach may not be sufficiently nuanced, because prescriptive attitudes to split infinitives have changed considerably in the course of the twentieth century, to the point where the majority of usage guides now advocate splitting infinitives, and few speakers are now aware of its status as a usage problem (cf. similar results for British English in Ebner 2017). In light of this, Fitzmaurice's (2000b) assumption that speakers' conscious awareness of the split infinitive somehow makes this construction a problematic candidate for studying language change is questionable. Evidence of how prescriptive attitudes have changed is thus the first step in accounting for their potential influence.

Another reason for drawing on usage guides in the study of attitudes to usage is that they evolve to reflect their social and cultural historicity, and this includes popular or accepted attitudes to usage. One way of studying attitudes to usage, as Algeo (1991b: 5) notes, is by examining and summarising what is said about usage in

usage guides. More concretely, usage guides are crucial in perpetuating prescriptive language attitudes. They are a unique product of the English language prescription stage, which is part of the standardisation process (Tieken-Boon van Ostade 2010; Peters 2006, 2012), and as such are central to language prescriptivism. Usage guides can be distinguished from descriptive grammars and dictionaries in that their primary function is to offer language advice on disputed points of usage to native speakers of the language; perhaps in part as a consequence of this function, they feature a great deal of personal opinions or subjective judgements on language usage (Busse and Schröder 2009: 82; Peters 2012). They are also important for the stabilisation of attitudes to correct usage, because "[w]hereas table manners are codified in handbooks of etiquette, 'correct' use of language is codified in handbooks of usage" (Milroy and Milroy 2012: 1). As such, they provide evidence for what kinds of attitudes to language have become institutionalised through publication and are, consequently, widespread.

Thirdly, usage guides are nowadays arguably the most influential type of metalinguistic works in which popular or prescriptive attitudes to language are expressed. As discussed above, the publication of usage guides has remained steady since their first appearance at the end of the eighteenth century. Furthermore, they also represent a fairly stable genre, albeit not an entirely homogeneous one (Straaijer 2018). However, despite variations within the genre, usage guides, compared to other forms of metalinguistic discourse, such as newspaper columns, letters to the editor, or oral complaints by speakers, are on the whole less ephemeral, less varied, and more tractable than other forms of metalinguistic discourse.

Finally, objective empirical descriptive studies of the genre of usage guides are rather rare (e.g. Creswell 1975; Peters and Young 1997; Albakry 2007; Busse and Schröder 2009, 2010; Busse 2015; Tieken-Boon van Ostade 2015). Alongside these important contributions to the empirical study of the usage guide genre, discussions of particular usage guides have previously appeared in the form of positive or negative reviews. Examples of the latter are Pullum's discussion of *The Elements of Style* (2010a), and his review of Simon Heffer's *Strictly English* (2010b), and Algeo (1991a) provides a positive review of *Webster's Dictionary of English Usage* (1989). Aside from providing views on the quality of the usage guides in question, what these publications suggest is that usage guides come in various kinds, and they are not all shockingly prescriptive or uninformed. Dekeyser (1975), for instance, concludes, on the basis of an empirical analysis of the nature of language pronouncements in nineteenth-century grammars, that in the course of that century, normative grammars

became less normative and less prescriptive. Similarly, work on eighteenth-century grammars has shown that many of these grammars are not exclusively prescriptive or descriptive, but can actually contain elements of both approaches to language (Straaijer 2009; Tieken-Boon van Ostade 2011). Dekeyser (1975) also posits a potential influence of the development of the science of linguistics on normative and prescriptive works on language. This is an interesting proposition, and one worth investigating in the context of present-day usage guides. Steven Pinker makes a related observation in a lecture on writing style,[1] noting that the most tangible influence of the science of linguistics on the treatment of points of usage is the scientific mindset, which requires empirical evidence in the resolution of usage contentions (cf. Landau 2001: 268). A comprehensive diachronic study of the usage guide genre can provide insights into the development of prescriptive ideologies over time (see also Straaijer 2018; Tieken-Boon van Ostade forthcoming).

Motivated by these considerations, the present study aims to provide an empirical exploration of the American twentieth-century tradition of usage guides. The questions at the centre of this investigation relate to the prescriptive nature of the usage guide tradition in the United States, and the nature of its influence on language practice and on language speakers in this variety of English (for British English, see Ebner 2017). I approach usage guides as a popular metalinguistic genre, strongly associated with normative or prescriptive approaches to language use. The aim is to investigate and assess the attitudes to language use presented in these books, and evaluate it in the context of popular attitudes to language use. I will also focus on testing the stereotypical association of these books with prescriptivism. These books are investigated in order to gain an understanding of how popular ideas about language use since the middle of the nineteenth were presented to a general audience, and to track changes in these ideas in the course of the twentieth century, and up to the present day.

1.4 Language variation and change

What makes prescriptivism an important factor to consider in processes of language variation and change is the fact that language variation and change represent concerns which are shared by prescriptive and descriptive approaches to language (cf. Peters and Young 1997: 315). In descriptive, linguistics change is seen as a natural part of

[1] Available online at: https://www.youtube.com/watch?v=9GubdYZPYPg&t=297s.

language itself, and as something to be considered in investigating the workings of the language system. Prescriptivism, on the other hand, considers change as something detrimental to the language, and its main goal is to stop language from changing and to 'fix' it in the state in which it is – or in its allegedly correct state – for ever. However, as mentioned above, prescriptive and descriptive factors are rarely neatly distinguished in language use. Even though the methods and goals of the two approaches are strikingly different, it is important that we gain a better understanding of the mechanisms by which what is seen as a natural process of language variation and change is influenced by prescriptive ideology. The possibility of hypothesising the existence of prescriptive influence on language variation and change, as well as of determining the importance of studying this influence, rests on a number of assumptions about the nature of language variation and change, which I will address in this section.

The first set of assumptions relates to the nature of language change. In Section 1.2, I mentioned the three principles formulated by Curzan (2014) which are important for a study of prescriptivism. One of those principles recognises the importance of language variation and change in both spoken and written language. In other words, evidence from both spoken and written language needs to be considered in assessing the potential influence of prescriptivism, based on the assumption that written language developments are as much part of the English language as spoken ones. This is particularly relevant in the context of prescriptivism, because prescriptive influence is generally more likely to be manifested in written language than in spoken; this, however, does not trivialise that influence. An additional reason for the importance of accounting for the effects of prescriptivism on standard written English is the fact that "standard English, while being one variety among many from a purely descriptive-linguistic point of view, has nevertheless been the most studied and best documented one because of its social and cultural prominence" (Leech et al. 2009: 1).

The second assumption related to the nature of language change is that a change in one word or one linguistic feature is as important as a change in the language system. Curzan (2014: 61), for instance, argues that a study of prescriptivism as a factor in change is important, despite the usual assumption that prescriptivism targets only a small set of linguistic features, because language change often happens word by word. Specifically for the empirical study of prescriptivism, this means that the effect of prescriptivism on the use of one feature is already indicative of prescriptive influence. While prescriptivism is unlikely to change a fundamental aspect of the language system, this does not make its influence less important or trivial. And in some cases, language ideologies "[have] noticeably changed the grammar of English within

my generation's lifetime, resulted from the feminist challenge to the once standard 'generic *he*'" (Kroskrity 2004: 496–497).

The final important assumption about the nature of language change is that changes above the level of consciousness can be non-trivial (Curzan 2014: 61; see also Tottie 1997: 84). While the majority of sociolinguistic studies of language change deal with changes below the level of consciousness, deliberate language change has been shown to be important in its own right. Relevant work in this respect has been done by Thomason (2007, 2011), who shows that contrary to "the general assumption [...] that such changes are relatively trivial, confined mainly to the invention or borrowing of new words, changes in lexical semantics, and the adoption of a few structural features from a prestige dialect [...] adult speakers can and do make deliberate choices that bring about nontrivial lexical and structural linguistic change" (Thomason 2007: 41). The importance of the model of deliberate language change developed by Thomason is based on the argument that while "speakers' choices can indeed lead to drastic linguistic changes [...] these changes only rarely have a permanent effect on the speech of an entire community; and where they do have a permanent effect, it is because of particular social circumstances" Thomason (2007: 58). One of the examples given of what may constitute particular social circumstances is "the deliberate actions of language standardizers" (Thomason 2007: 58).

The second set of assumptions distinguished here relates to language variants, since "[a]t any given moment during a linguistic change, speakers typically experience the change as variation, with some speakers using one variant and other speakers using other variants or with the same speakers using multiple variants, perhaps in different registers" (Curzan 2014: 46). Prescriptivism is specifically concerned with a subset of language variants which have, for one reason or another, become socially salient. Thus, what are usually described as usage problems in the context of prescriptive approaches to language use, are actually language variants which are characterised by language variation, and, potentially, language change. Such variants can then be defined as part of linguistic variables in the sociolinguistic variationist vein. Furthermore, changes in variants in a particular period are identified and analysed on the basis of the assumption that language change is "statistical in nature, with a given construction occurring throughout the period and either becoming more or less common generally or in particular registers" (Denison 1998: 93; see also Leech et al. 2009: 8). The other assumption about linguistic variants relates to the co-occurrence patterns of linguistic variants on the basis of registers and communicative functions. Building on the work of Biber (1988) on the co-occurrence of linguistic features

alongside stylistic or communicative dimensions, and a study of the influence of prescriptivism on the basis of language features as prescriptivism-related factors (Hinrichs et al. 2015), the present study will explore prescriptive influence on the basis of multiple features.[2] The importance of considering multiple features is that prescriptivism is not a monolithic entity, and that while it is true that each prescriptively targeted feature presents a separate case in itself, it can be expected that if the use of language in a particular situation, spoken or written, is influenced by prescriptive concerns, this influence will in turn affect a number of prescriptively targeted features. Thus, the assumption is that in texts where prescriptive influence can be hypothesised to affect the frequency of use of a particular prescriptively targeted variant, it will do so for other such variants as well.

Measuring the potential effects of prescriptivism on language variation and change will be done through a corpus-based study of the language variation and change of a number of selected language features, which is the focus of Chapter 6. The approach taken follows historical sociolinguistic principles, in that linguistic variants are, where possible, identified and defined in a context of a linguistic variable. Both text frequencies and proportions of the occurrence of particular variants are then used as metrics to investigate patterns of occurrence of these variants, and to arrive at an understanding of the patterns of variation and change.

1.5 Speakers and attitudes

I have so far used the term attitudes to usage to refer to observations and judgements on language use found in usage guides or similar metalinguistic works. In this section, I turn to another type of attitudes: speakers' attitude (cf. final paragraph of Section 2.7 for studies which look at attitudes to usage in metalinguistic works and Section 4.3 for more details on the distinction between attitudes to usage and speakers' attitudes). Since prescriptivism is a set of beliefs or attitudes about language, the other area in which the influence or effects of prescriptivism can be expected to be found is in speakers' attitudes towards language use, and towards usage features

[2]The terms 'linguistic variant/variable' and 'linguistic feature' will be used throughout the study, but they have a somewhat different reference. I use '(linguistic) feature' in the sense of Biber (1988) to refer to any language construction, word, or phrase which is targeted by prescriptive ideology. The terms 'linguistic variant' and 'linguistic variable' will be used specifically in cases where a particular feature can be identified as a linguistic variant, as not all features can be defined in terms of linguistic variants as part of a linguistic variable.

in particular. This relates to Curzan's third principle, which considers conscious awareness of language variants to be equally important in the development of the language. When we talk about prescriptive attitudes, we are referring specifically to attitudes to language usage, or, even more concretely, attitudes to specific language features. Garrett (2010: 7–10), for instance, distinguishes a number of different levels at which attitudes to language can take form in a language community. He treats in some detail the question of attitudes to usage features in relation to standard languages and the notion of correct or proper language use. These attitudes are related to explicitly expressed ideas about language, and are highly prominent in society.

However, attitudes towards language use in relation to prescriptivism have rarely been studied empirically, even though the research area of language attitudes is immensely varied, both theoretically and methodologically (cf. Ebner 2017). The term 'language attitudes' is in itself fairly general (see Chapter 2 for a definition), which means that language attitude studies can include anything from stereotypes associated with accents or dialects to attitudes towards second languages and their effects on second-language acquisition. Attitudes to language have been studied extensively from numerous perspectives, most notably in the social psychological tradition (e.g. McKenzie 2010; Giles and Rakic 2014), as well as in the tradition of perceptual dialectology (e.g. Preston 1999a). One reason for the relative absence of prescriptive attitudes and attitudes to usage in studies of language attitudes may be that such studies have traditionally been concerned with unconscious language attitudes, or with language attitudes in multilingual societies. Prescriptive attitudes are considered to operate above the level of consciousness, and as such are seen as attitudes imposed on speakers from outside. Perhaps the prevalence of such prescriptive attitudes makes them a rather predictable subject for research. However, this paradoxically points to two contradictory interpretations of prescriptive attitudes.

The first interpretation is that prescriptive attitudes held by speakers are a reality, but that they are not interesting from a research point of view precisely because they *are* predictable. This implies that prescriptivism influences speakers and their attitudes, but that this influence may not be particularly relevant to research. The second interpretation is that prescriptive attitudes are not as strong or consequential as subconscious attitudes, and as such are not capable of affecting language practice. This implies that prescriptive attitudes have no influence on language users and their language use. In any case, the influence of prescriptivism on speakers has also rarely been looked into. Aside from historical sociolinguistic studies relying on social network analysis in the study of the influence of the emerging standard language

ideology on individuals (e.g. Austin 1994), evidence of the extent to which speakers maintain prescriptive attitudes in present-day English is still rather limited. The inclusion of ordinary speakers in studies of the spread and maintenance of prescriptive attitudes to usage has similarly been neglected, with the exception of recent work done in the context of the Bridging the Unbridgeable project at Leiden University, whose aim was to fill this gap. Surveys of attitudes to usage in the tradition of Leonard (1932) and Mittins et al. (1970) did not engage with the attitudes of ordinary language users, but limited themselves to language professionals (see also Ebner 2017).

This study is a step in the direction of providing insights into how prescriptive attitudes operate among language users. Regardless of what is usually seen by descriptive linguists as the inadequate conception of language in prescriptive ideology, the notions of correct and standard language usage are entrenched in speakers' linguistic lives. "[T]he ongoing prescription that is part of standardization", Curzan (2014: 52) notes, "is part of many speakers' daily experience with the language, both written and spoken". This process of standardisation "has left a strong mark on modern-day attitudes, amongst some at least" (Garrett 2010: 8–9). Looking at the influence of prescriptivism on speakers' attitudes will thus allow for a much more nuanced and grounded investigation of the manifestations of prescriptivism in practice. Prescriptive language attitudes are also important in the discussion of prescriptive ideology and its influence on speakers, because they are potential factors in deliberate language change. Speakers have been shown to have agency in deliberate language change, i.e. language change above the level of consciousness (Thomason 2007). It is important to note that whether such agency will have a significant effect on the population as a whole, and consequently on the language as a whole, is dependent on many factors. In the context of prescriptive language attitudes or language attitudes which are instrumental in keeping the standard *vs.* non-standard functional distinction in language varieties, these attitudes can be expected to play a significant role.

Finally, understanding speakers' attitudes to usage is a crucial component in evaluating the influence of prescriptivism, not only on speakers, but also on language variation and change. Often, the relationship between prescriptive ideologies on the one hand and patterns of language variation on the other may differ depending on the language features investigated. For some features, for instance the passive in American English (Leech and Smith 2009; Anderwald 2014), prescriptivism may have been a stronger influence than for other, such as the split infinitive (see Chapter 6). Understanding speakers' attitudes to those variants may throw light on these kinds of differences. In relation to eighteenth-century normative grammars, for instance,

Tieken-Boon van Ostade (2006: 553) argues that by studying normative grammars in detail, "we will end up with [...] a clearer picture of how and why actual usage continued to differ from the norm which was imposed upon the language by the prescriptive grammarians of the eighteenth century and beyond". The same applies to present-day usage guides, and this thesis is an attempt to improve our understanding of the relationship between the usage guide tradition and speakers' attitudes. The question I will be concerned with here is what other sociolinguistic mechanisms are at play in the maintenance of prescriptively targeted forms, such as the non-literal use of *literally* (see Chapter 7).

1.6 Research questions

On the basis of the motivations, problems, and assumptions discussed above, this study offers an empirical investigation of the phenomenon of twentieth-century prescriptivism in American English. I focus on a set of six linguistic features, and will study these from the point of view of the three perspectives elaborated above: the American usage guide tradition, the patterns of language variation of these features, and speakers' attitudes towards these features. The concrete research questions, based on the empirical investigation of six linguistic features, which will be introduced in Chapter 3, are given below. The first one is the general research question, while the other three are the subquestions addressed separately in Chapters 5 through 7.

(1) What is the influence of prescriptivism on language use and on speakers' attitudes to language use?
(2) What is the treatment of the six linguistic features in American usage guides across time?
(3) What are the patterns of actual use/usage of these linguistic features?
(4) What are American native speakers' attitudes towards language use?

By exploring these questions empirically, this study aims to contribute to our understanding of prescriptivism as a factor in language variation and change. These findings will hopefully provide new insights for both historical and variationist sociolinguistics alike.

1.7 Outline

In Chapter 2, I will review the literature relevant to the study of prescriptivism and its relationship to language variation and change, as well as to speakers' attitudes. Drawing on previous studies, I will also establish the relevant theoretical background for the interpretation and discussion of the results of my various analyses in Chapters 5, 6, and 7. Chapter 3 discusses six selected language features, or usage problems, which I will focus on throughout the study, while Chapter 4 presents the methodology used, including the various types of data analysed. Chapter 5 presents the results from the analysis of usage guides, Chapter 6 covers the actual use perspective on the basis of a corpus-based analysis of the selected language features, and Chapter 7 discusses the results and findings from the language attitudes study. The study ends with a conclusion.

CHAPTER 2

Studying prescriptivism

2.1 Introduction

In the previous chapter, I pointed out that prescriptivism is understood as an ideology of correctness about language use. More generally, it can be understood as a type of language ideology which is manifested through speakers' beliefs, ideas, or rationalisations about language use (see Section 2.7). As such, prescriptive ideology is above the level of awareness of speakers, and it is perhaps the most widely discussed ideology in a standard English context such as the United States. This is also clear from the long tradition of usage guide publications in the United States, which started in the middle of the nineteenth century, and seems to be on the increase today. While linguists have generally been sceptical about the effects of prescriptivism on language structure, language change, and language users, a considerable number of studies have looked at prescriptivism and related phenomena in more detail. In this chapter, I outline the major research strands which provide the theoretical background for the present study.

In Section 1.2 above, I made a distinction between prescriptivism and prescription, drawing on Tieken-Boon van Ostade's work. It is important to keep this distinction in mind in the context of the theoretical background which follows, in that prescriptivism

can be studied in the context of both the codification stage and the prescription stage. In other words, I refer to "prescriptivism" as an ideology and an approach to language which is characteristic of the "prescription" stage in the model for language standardisation proposed by Milroy and Milroy (1985), but is not restricted to this particular stage. The prescriptive approach to language has been found to be characteristic of some eighteenth-century normative grammars as well. In this context, previous studies which have evaluated the influence of eighteenth-century prescriptivism on language use (e.g. Auer and González-Díaz 2005) have also provided useful points of departure for the present study.

The chapter is organised around these major research strands. Section 2.2 gives an overview of usage guide studies and their findings. Apart from studies focusing on usage guides, prescriptivism has also been studied in the context of normative grammars; this research is discussed in Section 2.3. Sections 2.4 and 2.5 review literature on the effects of prescriptivism, and on how such effects are reflected in language variation and change. In relation to the influence of prescriptivism on language variation and change, the notions of linguistic variants and usage problems are of particular relevance; these are discussed in Section 2.6. Finally, Section 2.7 outlines the relevant research on language ideology and language attitudes with reference to American English.

2.2 Studies of usage guides

English usage guides are books of language advice, generally intended for native speakers of the language. Busse and Schröder (2009: 72) define the usage guide as "an integrative all-in-one reference work written for educated lay people that bridges the traditional divide between a grammar and a dictionary". A usage guide is "an integrative all-in-one" work in that it often covers advice on multiple and various aspects of the language, such as grammatical and lexical points, punctuation and spelling conventions, and often pronunciation as well. The extent to which any of these language dimensions is included in a particular guide probably depends on decisions made by its author. A usage guide is a "reference work" insofar as it is intended to be consulted by users who are unsure about a certain linguistic choice. In other words, a distinguishing feature of usage guides is their "external function", which is "to enable the user to make choices between linguistic variants that can be functionally equivalent in a given context" (Weiner 1988: 173). Such users, as noted above, tend

to be "educated lay people" (Busse and Schröder 2009: 72), who are less interested in how language works than they are in obtaining practical usage advice when confronted with a usage choice they are uncertain about. In that respect, usage guides bridge "the traditional divide between a grammar and a dictionary" (Busse and Schröder 2009: 72, 84). The British and the American usage guide traditions have been the subject of a number of studies, focusing on various aspects of the genre, from its origins and social functions to its form and content. The following section outlines these studies.

However, usage guides are also methodologically and qualitatively different from both grammars and dictionaries. The methodological difference consists in the selection of language features that are covered in usage guides. Unlike grammars, which treat the grammar of a language, and unlike dictionaries, whose aim is to describe the entirety of a language's lexicon, usage guides treat usage problems (Weiner 1988: 173–174). Usage problems are items of common usage which are considered problematic for a variety of reasons, ranging from supposed grammatical incorrectness, as in the case of *between you and I*, to social controversy, as, for instance, the use of *ain't* (Ilson 1985; Weiner 1988; Algeo 1991b; Albakry 2007; Tieken-Boon van Ostade 2015). Often, these usage problems are so-called old chestnuts, or shibboleths, whose correctness or incorrectness is based on prescriptive notions about language use.

This dimension of usage guides has been decisive in establishing the genre as separate from grammars and dictionaries, as evidenced by the identification of the first usage guide (Tieken-Boon van Ostade 2010: 16), *Reflections on the English Language*, written by Robert Baker (1770). Leonard (1929: 35) notes that Baker was "[t]he first writer to codify his preferences into a book, the ancestor of those handbooks of abuses and corrections which were so freely produced in the nineteenth century". This in turn relates to the qualitative difference between usage guides on the one hand and grammars and dictionaries on the other. While present-day grammars and dictionaries strive to objectively describe and record the entirety of the language as it is used, usage guides are characterised by "their discursive and sometimes very personal treatment of subject matter" (Peters 2012: 248; see also Peters 2006). Thus, the consensus is that by their very nature usage guides are predominantly "a specific form of prescriptivist discourse" (Straaijer 2018: 12). While this may be the case with the majority of the usage guides, it is important to note that not all of these books are prescriptive. Usage guides such as Gilman's *Webster's Dictionary of English Usage* (1989) are recognised as offering a balanced perspective of usage based on evidence of language use (cf. Algeo 1991a). Usage guides thus represent a specific form of metalinguistic discourse,

which is usually marked by a high degree of language prescriptivism.

The first American usage guides were published in the middle of the nineteenth century; some of the first such publications were Seth Hurd's *Grammatical Corrector* (1847), Andrew Peabody's *Handbook of Conversation: Its Faults and Graces* (1855) and the anonymous *500 Mistakes Corrected* (1856) (cf. Connors 1983; Tieken-Boon van Ostade 2015). Since then, usage guides, handbooks of usage, or usage manuals, as they are variously called, have appeared regularly (on variations within the genre, and related genres, see Straaijer 2018). The variation within the genre makes it difficult to know how many usage guides have been published, but a realistic estimate would be that at least 200 have appeared in America since the middle of the nineteenth century (see Section 4.3 for more details on the basis for this estimate). Not all of these usage guides have proved to be equally successful or popular among the general public. Those that are, however, tend to have a secured status as classics, are usually well known, and can be assumed to be fairly influential. This means that the American usage guide tradition has been characterised by a small number of individuals whose influence has been significant, and whose names tend to be associated with prescriptivism. For instance, Richard Grant White is probably the most famous nineteenth-century usage guide writer and 'chief prescriptivist' (Drake 1977b; Finegan 1980; see also Busse 2015). The twentieth century saw the publication of a number of popular usage guides, perhaps the most famous being Strunk and White's *The Elements of Style* (1959). More recently, Garner's *Modern American Usage* (1998) seems to have gained a special place in the American usage guide tradition. In the course of its development, the genre has undergone some changes in terms of approach, methodology, and the manner in which language advice is offered. A noteworthy innovation in this respect is the use of panels of experts as the basis for usage pronouncements (Allen 2009: 357–358; Peters 2006) in some usage guides, as an attempt to achieve a more objective representation of usage norms and standards.

Apart from their pragmatic function of offering usage advice, these books have at various times been linked to broader social functions in American society. Connors (1983: 87), for instance, argues that "such small manuals [...] were the fruits of self-improvement fads and a burgeoning system of class distinctions, partially linguistically based, in America". The connection between books of language advice and self-help literature has also been made (cf. Landau 2001; Beal 2009; Yáñez-Bouza 2015: 25). According to Drake (1977b: 18), nineteenth-century America saw a revival of prescriptive impulses in language after 1850, when the rise of the so-called genteel culture produced greater emphasis on and concern with issues of language correctness.

This genteel culture was characterised by "an increased interest in language, especially in 'linguistic etiquette' in genteel publications; in the reaction against innovation; in the application of intellect and logic to language; in the high premium placed by the genteel on books and authority; in the anglophile tendency of the genteel; and in the desire for a responsible, stable community" (Drake 1977b: 18). Later in the nineteenth century, and at the beginning of the twentieth, industrialisation, the emergence and growth of new types of businesses, and migration allowed for unprecedented social mobility; in this context, knowledge and mastery of the standard was considered crucial for the social and professional advancement of the rising middle class (Drake 1977b: 24). This was attested by the popularity of language advice books. According to Drake (1977b: 19), for instance, "[t]he dictionary by the [18]60's had become a big business, due largely to the great waves of immigrants seeking linguistic passport to the society [...] and due to many native born Americans using linguistic conformity as a means to mobility".

Further insights into the social circumstances which stimulated the need for language advice, and consequently opened up the market for language advice publications, can be gained from a consideration of the increasing availability of education, and the emergence of new professions. Thus, in a study of the role of high schools in nineteenth-century America, Ueda (1987: 34) notes that "[a] standard written language was needed to foster the nexus of communication that underlay orderly social and economic relations in an impersonal, urbanizing society". The new businesses that arose around the end of the nineteenth century required a particular set of communication skills, including the use of language. These skills, Ueda (1987: 76) argues, were associated with "standard of living and status in the community"; consequently, parents wanted to make sure their children acquired those skills in high school. This broader social relevance of education, as well as the value assigned to the development of linguistic skills, meant that the prescriptive approach to language advocated in schools and colleges would have had a substantial influence on speakers. The topic of the importance of the standard in the teaching of English in America is also addressed by Marckwardt (1968), who sheds light on the historical process by which prescriptive attitudes to language use became influential in the teaching of English and composition in schools and colleges at the beginning of the twentieth century. Lindley Murray's grammar, first published in 1795, was the model on which grammar instruction was based in those days (see also Schweiger 2010), and his and other similar books "reflected the authoritarian tradition characteristic of the eighteenth-century grammarians" (Marckwardt 1968: 3). Language-related

publications for the general readership "were even more rigid and unyielding in their attitudes than the elementary school grammars" (Marckwardt 1968: 3; cf. also Bailey 1992: 14–15).

At the same time as these books were becoming popular, the student body in educational institutions across the United States was growing significantly, and high schools and colleges were becoming accessible to people from various walks of life. After World War I, the student body included children of native-born parents without any formal education, as well as children of non-native immigrant parents (Creswell 1975: 90). In such a social climate, in which the relevance of linguistic skills was highly valued, and correct and good English were seen as the prerequisites of social advancement, usage guides and other related genres sold 'correct English' as a commodity. A good example of an author and a businesswoman who seems to have been fairly successful at selling 'correct English' was Josephine Turck Baker (1873–1942) (Kostadinova 2018a). This perhaps sheds some light on the historical process through which prescriptive ideas about language became popular and widespread in twentieth-century America. A parallel can thus be made with eighteenth-century England, when language correctness was especially important for the rising middle classes, "as the correct use of language would be not only a key to upward social mobility but also a means of distancing themselves from their social inferiors" (Yáñez-Bouza 2015: 25; see also McArthur 1986: 8). This is a think what Landau (2001: 261) means when he describes usage guides as the "twentieth-century descendants of the eighteenth- and nineteenth-century grammars" and as "an American phenomenon", even though the tradition was first attested in the United Kingdom. He also argues that the popularity of usage guides in American English is a reflection of the linguistic insecurity of Americans, as well as the fluidity of American society. Because this fluidity allows social mobility, "Americans are naturally more highly motivated to acquire the skills that will help satisfy their ambitions" (Landau 2001: 262).

The usage guide tradition, as well as its significant social functions over the last 150 years, has very often been criticised in linguistic scholarship. "Throughout its history", Albakry (2007: 29) argues, "the entrepreneurial academy of usage commentary has been mainly prescriptive, with its judgements based solely on intuitions [... or] the opinions and personal preferences of handbook writers themselves". The fact that usage guides tend to prescribe, rather than describe, language use has defined the genre for a long time; one implication of this attitude of linguists towards books of this sort has been the relative absence of critical studies

of the genre by linguists – with some notable exceptions, which I address below. For instance, a few short studies of college textbooks or handbooks of composition were inspired by the concern with what standard should be taught in English in the course of the twentieth century (Allen 1935; Dobbins 1956; Malmstrom 1964[1958]; Tibbetts 1966; McDavid 1973; Connors 1983). Some of these were concerned specifically with the consensus on points of usage in these books, and found that handbooks were in considerable disagreement about issues of usage (Dobbins 1956: 46), which indicates a lack of consensus on what constitutes correct and good English.

Tibbetts (1966: 310) is probably one of the first "survey[s] of books on composition, rhetoric, and grammar" to include in the analysis a large number of such books, and to discuss the descriptive dimension of nineteenth-century and early twentieth-century textbooks, focusing specifically on the period betwee 1850 and 1925. While the exact number of textbooks surveyed is not explicitly given, mention is made of some eighteen books from which the author draws evidence in support of the argument that "older textbooks were far less prescriptive than we have been led to believe" (Tibbetts 1966: 310). Creswell (1975) compares the treatment of more than 200 usage items on which a usage note is included in the *American Heritage Dictionary* (1971) to the treatment of those items in a selection of dictionaries and usage guides. His selection of the ten books he analyses in the category of usage guides is an example of the difficulty in describing and delimiting the genre of usage guides. While some of these ten books, such as Krapp (1927) and Nicholson (1957), are usage guides, others – e.g. Leonard (1932) and Crisp (1971) – are usage *studies*. The latter two are in fact significantly different from usage guides, and resemble sociolinguistic studies of attitudes to usage rather than usage guides; I address this point in detail in Section 2.7 below. Beyond such limitations, however, Creswell (1975) is one of the few meticulous and quantitatively precise studies of the consensus in the treatment of various types of twentieth-century publications on usage, both dictionaries and usage guides, with the usage notes in *American Heritage Dictionary* serving as a starting point.

Genre distinctions in the context of metalinguistic works such as usage guides, however, remain problematic (Straaijer 2018), and this is certainly evident in many of these previous studies. The majority of these studies do not consider the issue of genre in much detail. One exception can be found in Connors (1983), who distinguishes between handbooks of composition and rhetoric and other types of manuals of usage. The question of defining the usage guide genre is later approached by Weiner (1988) as well, but not in the context of a study of usage guides in the vein of Connors

(1983). Weiner's discussion is important and informative, but it approaches the topic from the point of view of "the practice of usage guide writing: the scope of the genre, the selection of subject-matter, the use of evidence, the principles of guidance, and the organisation of material", i.e. from the perspective of "one who has attempted the genre" (Weiner 1988: 172). Finally, Algeo (1991b) can also be seen as addressing the question of the nature of the genre, albeit indirectly, through his typology of usage guides. He identifies seven types of guides, grouped into two broader categories. The first of these – and the largest one – comprises guides that are subjective and moralising in nature, while the second contains "works that aim at objectivity and reportage" (Algeo 1991b: 6). The prototypical usage guide belonging to the first category is Fowler's *Modern English Usage* (1926), while for the second Algeo provides the example of Leonard's *Current English Usage* (1932) as a case in point. I already mentioned above that Leonard's survey is very different in kind from usage guides, so Algeo's inclusion of this study in the same general group of usage guides is an example of the lack of and difficulty with a clear delineation of genre boundaries.

Irrespective of these issues with respect to genre delineation, a number of previous studies have addressed the question of the nature of usage pronouncements in usage guides and related genres. Berk (1994: 110) investigates language pronouncements in 26 reference books on language and found that while "[t]he reference books surveyed here reflect the entire spectrum, [...] prescriptivism is the dominant theme". She also found that the majority of the books are prescriptive and that "[a] common pattern among them is to acknowledge that language does change but to justify prescriptivism in the service of a distinction between formal and casual use" (Berk 1994: 112). Algeo (1994) compared two British and two American usage guides, in order to investigate their pronouncements on differences between the two varieties. He found that the four usage books do not agree on British *vs.* American usage, but modified his conclusion with a call for a more comprehensive empirical study (Algeo 1994: 107).

Two important larger-scale studies were done by Meyers, one in 1991 and another in 1995. Meyers (1991) studied the usage glossaries found in fifty college handbooks of composition, and compared them to the pronouncements in twelve handbooks of usage from the nineteenth century. In this study, Meyers found that there was significant variation and lack of consensus in the pronouncements in these works. Building on previous work done by Allen (1935), Meyers (1995) looked at the extent to which textbooks or handbooks (he uses the two terms interchangeably) are in keeping with the developments in usage and descriptive accounts of such usage. He concluded that there is a discrepancy between the textbooks' pronouncements

and actual usage, and that consequently these textbooks propagate the notion that such pronouncements are absolute standards and it is thus impossible for them to be changed or influenced by usage. Finally, Meyers concluded that no change or increase in linguistic awareness whatsoever had taken place between the 1930s and the 1990s. What he means here is that the textbooks he studied did not show any signs of revising their pronouncements to reflect actual use or descriptive linguistic knowledge. In a similar study, Peters and Young (1997) survey "a set of forty books from Britain, America and Australia and [examine] their treatment of eleven points of grammar to see how far prescription rather than description prevails" (Peters and Young 1997: 315). They further examined referencing practices of authors as an indicator of *ipse dixit* pronouncements and, consequently, of a prescriptive approach to language. What Peters and Young (1997: 321–322) found was that American usage books tend to use more references and to support their pronouncements with secondary sources.

In a more recent study of usage guides, Busse and Schröder (2009: 82) show that usage guides as metalinguistic reference works are characterised by the highest level of personal opinions expressed by authors compared to grammars and dictionaries. This is in line with the observation that usage guides are typically characterised by the discursive treatment of their subject matter (cf. Peters 2012: 248), referred to above. Busse and Schröder (2009) only deal with three editions of Fowler's *Modern English Usage*, so it may be difficult to generalise their conclusions with respect to the entire body of usage guides. Busse and Schröder (2010) look at the relationship between reference works, such as usage guides and dictionaries, and patterns of actual language use, in order to explore the extent to which the observations on language found in these reference works reflect patterns of actual use. They find that usage guides tend to reflect patterns of actual use to a great extent, and that over time it seems that usage guides have become more descriptive in their treatment of usage, although the extent to which this is true varies for different language features. Tieken-Boon van Ostade (2015) is a study of one particular usage guide, *Five Hundred Mistakes Corrected* (1856), and how it reflects the context in which it was written. An important conclusion drawn by the study is that the usage guide in question contains "a wealth of linguistic data" from the period in question, which can be particularly valuable for historical sociolinguistic research (2015).

What this last study has shown is that usage guides can sometimes provide interesting clues to sociolinguistic aspects of language variation and use. In addition to usage guides, other types of metalinguistic texts, such as normative grammars, have proved to be valuable sources of information for historical sociolinguists on

sociolinguistic variation in the past. An interesting example is provided by Arnovick (1997) in her discussion of the influence of normative grammars on the use of *shall* and *will*, where she notes that the rules about these forms appeared when the forms were undergoing a process of levelling, thus making the rules additionally significant. In other words, the rules were in a sense related to actual processes of change in the language, as "eighteenth-century grammarians maintain that basic illocutionary and semantic distinctions are worth making" (Arnovick 1997: 146–147). This dimension of metalinguistic texts has sometimes found its place in historical sociolinguistic studies. Nevalainen (1997: 158), for instance, refers to historical or early grammars for evidence of treatment of variation in *-ly vs.* zero adverbs. Another good example of the inclusion of evidence from grammars and usage manuals in the study of historical sociolinguistic variation can be found in Poplack (2006: 457), where Poplack describes how, alongside the study of historical language data, she and her colleagues used the descriptions, prescriptions, pronouncements, or value judgements found "throughout the prescriptive history of English", or, more specifically, in a "collection of nearly 100 English grammars and usage manuals" published between 1577 and 1898, to supplement the analysis with respect to any potential social values that could be discovered in these types of books.

2.3 Studies of prescriptivism

With some notable exceptions, the majority of the work dealing with the study of prescriptivism has been carried out in the last four decades. Despite the considerable variation that can be observed in this work, a number of different strands of research can be identified. This classification into groups of studies of prescriptivism is done merely for practical reasons, and reflects the kinds of background information each group of studies contributed.

The first type of studies are (usually book-length) discussions of the phenomenon of prescriptivism and its various facets, ranging from its nature and historical origins to its relationship with linguistics and its manifestations in everyday life. These works were crucial in formulating many of the theoretical assumptions of the present study, and provided useful perspectives on prescriptivism as a sociolinguistic phenomenon. One of the foundational texts in the study of prescriptivism, James and Leslie Milroy's *Authority in Language: Investigating Language Prescription and Standardisation*, was published in 1985, setting the stage for subsequent investigations of prescriptivism.

Two other important books in this category are Cameron (1995) and Curzan (2014). An important aspect of these studies is that they argue convincingly that linguists need to take prescriptivism seriously in sociolinguistic studies of language variation and change, as well as in studies of language attitudes. Milroy and Milroy (1985) and Cameron (1995) unearth various ideological aspects of prescriptivism, and foreground the importance of a critical investigation of prescriptive phenomena and the ways in which they bear on language use. Curzan (2014: 24) redefines prescriptivism by distinguishing four different strands of prescriptivism: standardising, stylistic, restorative, and politically responsive prescriptivism. In the context of this redefinition of prescriptivism, Curzan (2014) elaborates on how these different strands affect actual language use, and points out the need to take prescriptivism into account in telling language history. Beal (2009) discusses the value of good English in its historical context, and focuses on its present-day manifestations, thus addressing the important issue of what she calls "new prescriptivism", especially with reference to accent.

Another aspect which characterises these studies is that they predominantly deal with prescriptivism in the twentieth and twenty-first centuries. For example, Milroy and Milroy (1985) bring contemporary sociolinguistic research to bear on deconstructing prescriptive attitudes, and they discuss how these kinds of attitudes can have a negative influence in education, and specifically in language assessment. Cameron (1995) offers in-depth analyses of prescriptivism and verbal hygiene in various contexts such as the role of publishing, and copy-editing in the perpetuation of particular (prescriptivist) language ideologies, or the issue of prescriptivism in the context of changes in the curriculum in the United Kingdom. In the most recent of these studies, Curzan (2014) presents an informed discussion of the potentially far-reaching effects of prescriptivism mediated by grammar checkers in text-processing software.

Another strand comprises historical studies of the development of attitudes to usage and prescriptive ideas over time, exemplified by works such as Leonard (1929), Drake (1977b), Finegan (1980), Baron (1982), and Bailey (1992). While these studies address the topic of the historical development of attitudes to language, and in particular prescriptive or popular attitudes associated with notions of correctness, they all differ somewhat in their approach, and in the historical and cultural contexts they deal with. Leonard (1929) analyses the attitudes to language usage in eighteenth-century normative grammars, without focusing specifically on British or American English. While the study of Leonard (1929) may not be directly relevant for the analysis of usage guides, as most of the publications discussed are grammars,

with the exception of Baker's *Reflections* (1770), it is a classic study which provides a useful historical background for the origin and rise of prescriptive attitudes. Drake (1977b) charts the historical processes which shaped the public understanding of correct language usage in America in the nineteenth and twentieth centuries (see also Drake 1977a). Finegan (1980) explores the historical development of attitudes to language usage, with a particular focus on the American context. He specifically addresses the juxtaposition between the "doctrine of correctness" and the "doctrine of usage", in order to show the legitimacy of both positions in the so-called war on authority in language matters. Baron (1982) documents the development of language reforms mainly in eighteenth- and nineteenth-century America, by focusing on the work of the most prominent language reformers. The work contains a great deal of historical detail about concerted efforts to plan and reform the English language at different times in history. What is also important about this group of studies is that they look at language attitudes to usage, not, strictly speaking, speakers' attitudes. I referred to the importance of keeping these two separate in the context of this study in Section 1.5, and I discuss the terminological decision to employ this word in Section 4.3.

The third group of studies of prescriptivism are historical sociolinguistic studies dealing with the Late Modern English period, and are concerned more specifically with the origin of prescriptive ideology in the processes of standardisation and codification. Prominent studies in this strand are those describing the rise of prescriptivism, the social conditions that contributed to the emergence of prescriptive grammar and usage guides, and the origin and establishment of eighteenth-century prescriptive rules. Quite possibly the earliest study on the relationship between prescriptivism in normative grammars and patterns of actual language use is Fries (1925); see also Fries (1940). In this early corpus linguistic study, Fries examines the attitudes to the use of *shall* and *will* in about 60 normative grammars from the sixteenth century onward, and traces the changes in attitudes to *shall* and *will*, as well as grammarians' rules about the temporal reference uses of these two forms. He found that the general rules dictated that *shall* is used with future temporal reference in the first person but with the meaning of obligation in the second and third persons, while *will* follows the opposite pattern. Fries (1925: 1016) then tested these rules against the frequency of occurrence of *shall* and *will* in self-compiled corpora of British and American English plays. The results showed conclusively that normative rules with respect to the uses of *shall* and *will* were not supported by the observed patterns of use in the corpus data: *will* was found to be more frequent with first person in declarative

clauses, while *shall* was more common with first person in questions. With second person subjects, *will* was found to be more common in questions and in declarative clauses. In addition to these findings, a clearly decreasing pattern was identified in the use of *shall* in second and third persons, while in terms of regional differences, American usage data yielded lower frequencies of *shall* in almost all contexts of use.

Further studies dealing with various aspects of eighteenth-century prescriptivism, in terms of either individual grammarians or specific features, can be found in Tieken-Boon van Ostade's work. Tieken-Boon van Ostade (1994) looks at the relationship between normative grammar rules on pronoun usage and actual use in the course of the eighteenth century, and concludes that there is a continuum of variation from actual use to the standard norm imposed in the grammars. Tieken-Boon van Ostade (2006) shows how Lowth's social network may have affected the language norm he prescribed in his grammar, while Tieken-Boon van Ostade (2008a: 205) revisits the issue of multiple negation in eighteenth-century normative grammars, and concludes that even though multiple negation seems to have been ousted from the standard variety before the eighteenth century, it persisted as a vernacular language feature, used by the lower social classes (see also Nevalainen 2000). This, according to Tieken-Boon van Ostade (2008a), might explain the inclusion of this feature in the grammars. Other studies on the rise of prescriptivism are Percy (2009), who traces the origin of prescriptivism back to periodical reviews in the eighteenth century, and Straaijer (2009), who analyses the level of prescriptivism in normative grammars on the basis of a quantitative analysis of the types of modals used by different grammarians when making language judgements.

These studies are related to the origin and source of prescriptive rules, or investigate the rise of prescriptivism and the nature of eighteenth-century prescriptivist attitudes and discourse. The second type of historical sociolinguistic studies, as I mentioned above, are concerned with what comes after the stage of codification, i.e. the stage of prescription (Tieken-Boon van Ostade 2012a,b). The effects of prescriptivism are thus a crucial object of research in these studies, and as such they are of particular interest to the present study; these will be discussed in more detail in the next section.

2.4 The effects of prescriptivism

In this section I turn my attention to an examination of studies of the effects of prescriptivism on variation and change. These are larger-scale studies which analyse patterns of precept in normative literature (i.e. prescriptive rules) and compare them to patterns of language variation and change established on the basis of corpus data. In the context of the distinction between codification and prescription, as well as the difference between normative grammars and usage guides, it is important to make a clear distinction between studies which investigate the influence of prescriptive pronouncements in normative grammars on language use, and studies which evaluate the influence of usage guides. While it is important to note that the second group of studies is more relevant for my own analysis, the work done on evaluating the potential effects of normative grammars on language use has also provided many useful perspectives, especially with respect to methodology. In what follows, I address these two bodies of work separately.

Auer and González-Díaz (2005: 318) make a useful distinction between studies of the influence of prescriptivism on a micro level and on a macro level. Micro-level studies, they observe, are "based on social network theory and the influence of prescriptivism on the idiolect of selected people" (cf. Auer and González-Díaz 2005 for further studies). Macro-level studies look at general patterns of change in prescriptive attitudes on the one hand, and frequency patterns in large-scale language use data on the other, through the application of an approach sometimes referred to as "precept *vs.* practice" (see Auer 2009: 4–11 on the origin of this approach). An early macro-level study whose goal was to investigate the success of nineteenth-century British normative grammarians' attempts to regulate language use, specifically in the context of number and case relations, is Dekeyser (1975), a study already mentioned in Section 1.3 above. Number and case relations represent cases of government and concord, which were, according to Dekeyser (1975: 2), part of the focus of the syntax of normative grammars. His analysis of prescriptions is based on a self-compiled corpus of 60 grammars and error books, while his analysis of actual language use is based on a self-compiled three-million-word corpus of texts comprising novels, non-fiction, and letters or essays. The results of the analysis showed that while prescriptivism had no effect on diachronic developments in language use in the course of the nineteenth century (Dekeyser 1975: 276), some effect could be hypothesised with respect to genre differences, as more colloquial texts in Dekeyser's corpus

displayed higher proportions of 'incorrect' constructions – which he calls "licentious forms" – than the more formal texts. Another relatively early study whose aim was "to find out how great an effect eighteenth-century grammarians had upon actual usage as far as the double negative is concerned" is Tieken-Boon van Ostade (1982), which found that the constructions of negative concord which normative grammarians used as examples of incorrect usage were actually not found in eighteenth-century texts, while the negative concord feature which did occur in the texts was not commented on by grammarians. Tieken-Boon van Ostade thus concluded that normative grammarians had no effect on the subsequent decrease in the use of negative concord, as this is likely to have happened before the eighteenth century, and, furthermore, that their prescriptive pronouncements were probably based on the usage of earlier periods.

Chafe (1984) addresses the issue of the role of prescriptivism in the development of differences between speech and writing. Through the examination of data on features such as *shall/will* and dangling participles, he proposes that prescriptivism has had an influence, but that this influence may be manifested in different ways in the language. These different manifestations are presented through three models for the potential influence of prescriptivism on similarities and differences between speech and writing. These models will be discussed further in the next section, as they are more relevant to the question of how prescriptive influence is conceptualised in relation to language variation. Chafe also notes that, when looking at prescriptive influence on multiple language features, "[t]here is a sense in which each feature affected by prescriptivism has had its own history" (1984: 102). Using newspaper language data from the nineteenth century to investigate the use of *shall* and *will*, Facchinetti (2000) finds that distinctions between these forms made by normative grammarians – i.e. *shall* expressing futurity with first person subjects, and obligation with second and third person subjects – are generally adhered to in the data analysed. On the other hand, she also finds that grammarians were mistaken in associating perceived misuses of *shall* with Irish speakers, as her data show that "at least in the first part of the nineteenth century, the Irish employed *shall* with first person subject more frequently than the English" (2000: 130).

In the area of the subjunctive, it has been shown that normative grammars may have exerted limited short-term influence on the decreasing trends in the use of the construction in the course of the eighteenth century (Auer and González-Díaz 2005; Auer 2006), while in the context of double comparatives, it has been argued that the role of prescriptivism was that of a reinforcing influence of an already strong decreasing trend (Auer and González-Díaz 2005), similar to that observed for

negative concord (Tieken-Boon van Ostade 2008a: 205). For the nineteenth century, Anderwald (2014) shows that American normative grammars reacted vehemently against the progressive passive, a construction which was a typical nineteenth-century development. Anderwald's data showed that, despite strong negative criticism, the construction rose in frequency. However, corpus data for the twentieth century indicated that the progressive passive construction declined sharply after its peak around the 1940s, especially in newspaper language. Anderwald (2014: 14) links this decrease to the publication and unparalleled popularity of Strunk and White's *Elements of Style* (1959), and concludes:

> Surprisingly, then, while for the purportedly over-prescriptivist nineteenth century a prescriptive influence on actual language change could not be convincingly demonstrated, the middle of the twentieth century, the descriptive century per se, showed the most convincing correlation of the publication of a notorious style guide (Strunk & White) and the actual striking reversal in the fortunes of the progressive passive in written American English.

This is an important finding in the context of the present study, because it shows that the effects of prescriptivism need to be investigated for the period of the twentieth century as well. In this context, as I established at the beginning of the previous chapter, usage guides are the central source of data on prescriptive ideology and prescriptive attitudes to usage. Consequently, studies examining the influence of usage guide prescriptions on actual language use also provide an important context for investigating prescriptivism and attitudes to usage in the twentieth century. Albakry (2007) presents a quantitative analysis of the effects of style and usage guide prescriptions on actual language practice in the context of written media registers, by looking at the extent to which style guides and their judgements have influenced newspaper language in American English. He does so by focusing on five language features: clause-initial coordinators, stranded prepositions, split infinitives, functional shift, and modified absolute adjectives. Albakry's analysis is particularly relevant, in that he establishes a difference between strongly and weakly dispreferred features, and hypothesises, on the basis of his results, that strongly dispreferred language features are less frequent in newspaper data than weakly dispreferred ones. However, one limitation of this study may be considered the lack of distinction between usage guides and style guides. This limitation relates to the general problem of clearly delineating different types of genres of metalinguistic works, which I discuss in Section 2.2. One way in which style guides and usage guides can be distinguished is that style guides

tend to be associated with a particular publication (i.e. a newspaper or a magazine), and as such contain language usage rules and guidelines relevant for that publication. Usage guides, on the other hand, are oriented towards the general public (for more on this distinction, cf. Ebner 2016; Straaijer 2018) Similar in some respects to Albakry (2007) is a more recent study evaluating the effects of usage pronouncements on the development of norms in British and Australian English (Peters 2014). Peters (2014: 596) shows, through the example of hyperstandardised language features such as *-ise/ize*, *alright*, and singular *data*, that in both British and Australian English "the standardization of language norms does not depend on a framework of continually reinforced prescription. Rather, the norms develop their own momentum in common usage".

The studies discussed so far have shown that prescriptivism seems to have a temporary effect on language change. Evidence to the contrary comes from two diachronic studies: one on the development of perfect infinitives, and the other on the stigmatisation of *you was* forms. Exploring the development of perfect infinitives, Molencki (2003: 175) argues that with respect to counterfactual infinitival constructions, "certain natural language processes [in this development] were either retarded or prevented owing to the prescriptivists' activities" in the eighteenth and nineteenth centuries. Specifically, he shows that before the explosion of normative grammars in the second half of the eighteenth century, perfect infinitives were used to express counterfactual meanings, i.e. expressing unreal, hypothetical situations, but this was stopped by grammarians who considered the expression of temporal anteriority to be the only possible function of the perfect infinitive. Laitinen (2009: 200) looks at the role of eighteenth-century normative grammars in stigmatising singular *you was*, as opposed to singular *you were*, and suggests that "the role of normative grammars in the diachronic development of this particular variable was substantial". According to his analysis, *you was* started spreading as a typical change from below towards the end of the seventeenth century. Normative grammarians then picked up on the variable *you was/were* and started proscribing *you was*, which resulted in its effective stigmatisation by the beginning of the nineteenth century.

Two studies addressing directly the question of the influence of prescriptivism on language change are Tottie (1997) and Hinrichs et al. (2015). Tottie (1997) examines the influence of literacy and prescriptivism on the variation between *that* and *which*, and between *that* and *who*, in both British and American English. With respect to American English specifically, she found that the patterns of use of *that* as opposed to *wh*-forms seem to be affected by the opposing influences of literacy

and prescriptivism. According to Tottie (1997), spoken and general written data from American English show that *that* is the more frequent option, and she associates this with the influence of prescriptivism. On the other hand, among highly educated American speakers, *wh*-forms are more frequent than in general or spoken usage. The author explains this as the result of literacy. In other words, Tottie concludes, literacy and prescriptivism are opposing forces in American English: the former affects the linguistic behaviour of speakers, especially in formal contexts, while the latter affects editorial practices, and consequently, written usage. Hinrichs et al. (2015) is a more recent empirical investigation of the effects of prescriptivism on twentieth-century American English, which has shown that prescriptivism may have had an influence on language change. However, this influence is mediated by other language change and social processes. Specifically, Hinrichs et al. (2015) have shown that certain language features are more sensitive to prescriptive influence than others, and that additional processes such as colloquialisation, or strong language authority, might play a crucial role in eventually determining which language features will be influenced by prescriptivism, and which will not. The research by Hinrichs et al. (2015) is methodologically ground-breaking, in that it applies a novel approach to the empirical testing and measuring of prescriptive influence, and has significantly influenced the present methodological approach. This aspect of the study will be discussed in more detail in Chapter 4.

A final study I consider relevant to mention in the context of analysing the effects of prescriptive language ideology is Kroch and Small (1978). This is an early study of the effects of grammatical ideology on speech on the basis of data from a small group of speakers. Furthermore, the study is a multifactorial quantitative study, and it is a rare attempt to account for both internal and external factors in the study of the effects of, in their terms, "grammatical ideology" on speech. By comparing the use of standard and non-standard forms by radio hosts and call-in listeners, Kroch and Small (1978) found that radio hosts were more likely to use the standard forms. This led the authors to the conclusion that grammatical ideology does have a measurable effect on speech.

All of these studies have provided the point of departure for my conceptualisation of prescriptivism, as well as for the methodological approach, which will be discussed in Chapter 4. Drawing on these studies, in the remainder of this chapter I will address the important question of how prescriptive influence has been conceptualised, and how it can be operationalised and measured. These issues will be addressed with respect to language variation and change on the one hand, and speakers' attitudes and ideologies

on the other.

2.5 Prescriptivism and language variation and change

An important issue related to the difficulty in ascertaining the influence of prescriptivism on language is how this influence is conceptualised. In other words, what do we in fact mean when we ask: does prescriptivism affect language variation and change? In this section, I review previous studies of the effects of prescriptivism on language variation and change by focusing specifically on the ways in which these studies explore those effects, and how they ascertain patterns of variation and change. Since I am concerned primarily with the period from the beginning of the twentieth century until the present day, I will also draw on a number of important studies of grammatical changes in twentieth-century English.

In discussing the nature of the potential influence of prescriptivism on language variation and change, it is crucial to distinguish the effects of prescriptivism on language change from those on language variation. The reason this distinction is crucial is that the majority of the large-scale studies of prescriptive influence have shown that prescriptivism rarely has a lasting influence on long-term processes of language change (e.g. Dekeyser 1975; Chafe 1984; Auer and González-Díaz 2005; Auer 2006; Anderwald 2012, 2014; Yáñez-Bouza 2015). If any effects were identified in these studies, they were temporary (Chafe 1984; Auer and González-Díaz 2005; but see Molencki 2003). On the other hand, on the basis of register variation patterns of these variants, written language can reasonably be hypothesised to be influenced by prescriptive pressures (Anderwald 2012: 267). For instance, Dekeyser, Chafe, Auer and González-Díaz, and Anderwald found that while over time proscribed features do not seem to be affected by prescriptivism, at particular points in time proscribed features are less frequent in edited or formal registers, and are most frequent in spoken data or in fiction.

Taking this register effect further, Chafe (1984) makes interesting and relevant points about the difference between speech and writing, and how the influence of prescriptivism can be conceptualised and explained. He identifies three possible scenarios in which prescriptivism can be considered to account for differences between writing and speech. In the first case, a feature that has been established in the spoken language is not adopted in writing, partly due to the inertia associated with written language norms, and partly due to prescriptive attitudes. Chafe (1984: 96–97)

cites the use of *this* to introduce new information, as in *Then this guy appeared*, as an example of such a case. The second situation Chafe (1984: 97) describes is "[a] development typically associated with prescriptivism"; it refers to a feature that starts out being used in both speech and writing. Once its frequency of use is noticed and commented upon by prescriptivists, the frequency of use of that feature in writing decreases, perhaps to the point of disappearing. However, spoken language still retains the feature. An example of this is the use of *I shall* as opposed to *I will* to express future temporal reference. Initially, both were used to express future, but under the influence of normative grammar prescriptions, Chafe argues, future *I will* became infrequent in writing, with the distinction between the two forms being observed in writing. However, after the initial influence of prescriptivism faded, *I will* increased in frequency, with *I shall* disappearing from written registers. The third pattern of influence is found in the case of features which are not part of the spoken language to begin with, but are rather more typical of written texts. In such cases, of which the split infinitive is given as an example, prescriptivists observe the pattern in written language and criticise it, and, under the influence of this criticism, the pattern disappears. This kind of development, Chafe (1984: 99) observes, "contradicts the notion that prescriptivism always increases the distance between writing and speaking".

These findings, as well as the patterns observed, serve as important conceptual tools in the formulation of different types of prescriptive influence. One of these patterns, for instance, has been found in studies of the subjunctive, which show that the strong reactions of normative grammarians in favour of the subjunctive, which were triggered by the noticed decrease in use of the feature in the course of the eighteenth century, resulted in a slight temporary increase in the frequency of use of this feature (Auer and González-Díaz 2005). To sum up, the way in which the effects of prescriptivism on language variation and change have been tested is by looking at changing frequencies of proscribed or prescribed variants, and proposing prescriptive influence as an explanation for observed variation and change patterns, in cases where, as noted in one study, all other possible factors seem unlikely to be significant (Auer and González-Díaz 2005). With respect to the influence of prescriptivism on actual language use, then, we need to take into account various levels of language use. It would be one thing to speak of the influence of prescriptivism on the language system itself, and another to speak of the influence of prescriptivism on the frequency of usage in particular genres. As Anderwald (2014: 14) shows with respect to the development of the progressive passive in nineteenth-century British and American English data, "text-type sensitivity of the progressive passive was shown

to be extremely pronounced in both national varieties". This raises the question of identifying patterns of variation and change, as well as the constraints on language variation and change that have been established in previous studies. In what follows, then, I draw on studies of grammatical variation and change in twentieth-century English, inasmuch as this research bears on the present study.

In the context of recent grammatical change, Denison (1998) provides the most comprehensive overview of changes that have taken place in English since 1776, and this account serves as the point of departure for data on many of the features analysed for this study (for more on this, see Chapter 4). According to Denison (1998: 93):

> Since relatively few categorical losses or innovations have occurred in the last two centuries, syntactic change has more often been statistical in nature, with a given construction occurring throughout the period and either becoming more or less common generally or in particular registers. The overall, rather elusive effect can seem more a matter of stylistic than of syntactic change, so it is useful to be able to track frequencies of occurrence from eModE through to the present day.

Mair and Leech (2006) also discuss a number of recent changes in English, as well as empirical data on the frequency patterns of particular variants identified as undergoing change. Mair (2006) identifies a number of important processes in twentieth-century English, such as colloquialisation, which add to our understanding of stylistic or extralinguistic factors in shifting frequencies of variants. In other words, processes such as colloquialisation may be seen as opposing tendencies to the influence of prescriptivism, and this may in turn have implications for the operationalisation of these constraints on language variation and change (a question I discuss in more detail in Chapter 4). Finally, Leech et al. (2009) also provide the background for the present study, as they cover a number of important contemporary changes, such as the use of the passive or of *that* and *which* in restrictive relative clauses with inanimate referents. In addition to the analysis of changes in specific linguistic variants, Leech et al. (2009) also identify a number of linguistic or other determinants of language change, which are important to consider in relation to prescriptivism, similar to the process of colloquialisation, identified by Mair (2006). One of the most important conclusions of this research, as mentioned above, is that the twentieth century is rarely marked by profound changes in grammatical structure; rather, observed changes are more visible as changes in statistical tendencies and variation.

This brings me to another point which is important in ascertaining potential

prescriptive influence, the difference between structural change and stylistic change. Szmrecsanyi (2016), for instance, distinguishes environmental change from grammatical change proper. He discusses the problem of relying solely on text frequencies when studying grammatical change and carefully teases out possible confounding variables in a corpus study of variation and change. Drawing on his distinction between environmental change and grammatical change proper, for the purposes of the present study I take prescriptivism to belong to the environment of what Szmrecsanyi calls "textual habitat". He argues that disentangling environmental change from grammatical change proper is possible. This is an important distinction for the present study, because prescriptivism can more readily be considered as a factor in environmental change than as a factor in grammar change proper. This is also related to the fact that grammatical change takes longer to complete and is slower and more imperceptible than environmental change. Environmental change in frequency could depend on many things, including, I argue, prescriptive influence. For instance, in a hypothetical study of variation in the use of a particular linguistic feature, we might find that certain heavily edited texts contain no instances of that feature, which may in turn lead to the conclusion that the feature has disappeared from the language. However, this would be an instance of environmental change: the hypothetical feature would not be found in these texts simply because the editing process has influenced the use of the feature. The distinction between grammatical change and environmental change may also offer an explanation for the widespread assumption that prescriptivism has no influence on language because it rarely has an influence on grammatical change proper. However, if we take this distinction into account, then we can more meaningfully evaluate prescriptive influence in terms of influence on environmental change. In other words, if we are to understand how prescriptivism affects language, we need to be able to distinguish between its influence on structural diachronic changes, which has been shown to be minimal, and its influence on stylistic, or probabilistic synchronic variation patterns, which, as shown, is more likely to occur.

Alongside the effects of prescriptivism on language variation and change, a separate question is the extent to which prescriptivism has an effect on speakers, both in terms of attitudes and in terms of the language use of individual speakers. This is a question that has not been empirically investigated as often as the first one, in part due to the difficulty of obtaining reliable evidence for a satisfactory answer. One way of going about it is to analyse changes in language use of individuals over time, as in the study of Austin (1994), or to analyse spoken data on the basis of situations in

which speakers may be expected to be under pressure to produce what they consider to be grammatically correct speech; the comparison of language use of radio hosts as opposed to that of call-in listeners in Kroch and Small (1978) may be considered an example. A question relevant to the influence of prescriptivism on speakers is about whether specific kinds of speakers tend to be affected by prescriptivism, if such kinds could reasonably be identified. It may, for instance, be argued that socially mobile speakers are more likely to be influenced by prescriptivism. According to Fitzmaurice (1998), the purpose of eighteenth-century normative grammars was to provide language advice to the lower or middle classes, while Tieken-Boon van Ostade (2008a: 208) notes that these speakers were "the class of people which formed the target audience of grammars like the ones by Lowth and Murray". In discussing the well-known rule on the use of *who* and *whom*, Aarts (1994: 74) notes that "...it is true that *whom* is now virtually dead in informal English, it is also true that most educated speakers of English are still aware of the rule which says *whom* is the correct form to be used when the relative pronoun is not the subject." This is why speakers' attitudes are crucial in ascertaining prescriptive influence.

As a final point with regard to the question of the effects of prescriptivism, it should be pointed out that it is also possible for prescriptivism to be influenced by language variation and change. In other words, over time, prescriptivism and its manifestations (i.e. rules and prescriptions in popular usage guides, or prescriptive speakers' attitudes) may, and in many cases are bound to, align with patterns of language variation and change. In some of the cases I discussed above, normative literature was found to be influenced by actual language use. In the context of more recent manifestations of prescriptivism, Albakry (2007: 25–26) reasons:

> If the practices of newspaper writers coincide with the usage practices allowed for by those preparing usage manuals, can we be so sure that newspaper writers are "heeding" the pronouncements of those handbooks? Not quite, since it is also possible that some of the authors of usage books have been observing what educated writers, including perhaps newspaper writers, are doing in the language. Furthermore, both groups themselves could be observing normative patterns among educated writers. In this scenario, newspaper writers may not necessarily be paying attention to what usage books say but are rather adhering to larger cultural and linguistic norms.

While I will not investigate such claims explicitly, I will consider them in the interpretation of the results of the analyses I present here. In certain cases, as will be argued in Chapter 7, speakers' attitudes that differ from prescriptive attitudes may be

crucial in determining which proscribed language features will end up being accepted over time.

In conclusion, the distinction between language variation and language change is crucial in assessing the potential effects of prescriptivism in an accountable way. Furthermore, a failure to make this distinction is the reason that people who are sceptical of the influence of prescriptivism on language change usually also dismiss the influence of prescriptivism on language variation. Those people often cite cases of language change as evidence that prescriptivism has no influence. However, looking at synchronic variation and register variation, prescriptivism may prove more likely to be influential than is generally assumed.

Most of the studies of prescriptive influence have shown that normative or prescriptive texts alone are not always a reliable source of prescriptive influence in society, which is another crucial aspect of prescriptivism. In many cases in which prescriptivism has been found to have effects on language variation and change, it has also been shown that such influence was crucially supported or conditioned by broader social or cultural processes. For instance, in the context of the decrease of preposition stranding, Yáñez-Bouza (2015: 125) found that "late eighteenth-century prescriptivism cannot be held principally to account [for the decrease]; in other words, late eighteenth-century precepts did not trigger change, but rather reinforced an existing trend". The same has been shown in the context of the disappearance of negative concord (Tieken-Boon van Ostade 2008a). Furthermore, Yáñez-Bouza (2015: 126) suggests that there must have been "latent awareness" of the stigmatised status of preposition stranding, because "something must have happened at this early stage, or even before it, to bring the steady rise in usage to a halt". She then argues that what must have happened in the course of the seventeenth century was the development of a latent awareness of the incorrectness of preposition stranding, which accounted both for its decrease, and for the stigmatisation of the feature in eighteenth-century normative grammars. That this is also the case in twentieth-century English has been suggested by studies dealing with recent data. Hinrichs et al. (2015), for instance, show that the decline in restrictive relative *which*, in favour of *that*, has been crucially conditioned by processes such as institutionalisation and colloquialisation, while in the area of spelling, the influence of usage guides is sooner or later superseded by the influence of common usage (Peters 2014). In this respect, we could generalise that even though prescriptive literature may not have permanent long-term effects on language change, at certain points in time, prescriptivism – understood more broadly – may be an important factor in variation.

Furthermore, prescriptive and descriptive attitudes identified in metalinguistic works such as normative grammars and usage guides are not isolated from their social context, and may in time start to reflect the broader attitudes associated with language.

2.6 Usage problems and linguistic variables

As pointed out above, the main concern of prescriptivist literature on language is the establishment and maintenance of clear distinctions between right and wrong usage. The term 'usage' itself, as has been noted previously (e.g. Peters 2006; Albakry 2007; Allen 2009; Busse and Schröder 2009), can be used descriptively, to refer to patterns of usage in a language, or it can be used to refer to the ideological category 'correct usage' or 'good usage'. In the latter meaning, the term 'usage' is used to "[refer] to a finite set of stigmatized linguistic features" (Albakry 2007: 29). This boils down to a set of "features of divided usage", such as "perceived errors of grammar like *you was* or *less* for *fewer*" (Tieken-Boon van Ostade 2015: 57), known as 'usage problems'. Algeo (1991b: 2), similarly, observes that "[u]sage is a choice among alternatives to which users attribute social value". It is thus useful to consider how the concept of the usage problem, as well as the study thereof, can be usefully connected to the concept of the linguistic variable. In what follows, I will argue that most of the grammatical usage problems are linguistic variables, which is important in approaching an investigation of how they are used. While this is applicable to many usage problems, there are those which cannot be conceptualised as linguistic variables; this is a point I will address in more detail below.

Linguistically, many of the grammatical usage problems can be seen as variants of language structures that are usually assumed to have the same denotational or referential meaning (for a discussion of the assumption of referential equality, see Lavandera 1978), but are considered by speakers to be 'unequal', based on a number of conventionally established norms of correctness. This means that *I have not seen anybody*, and *I have not seen nobody* are, from a linguistic point of view, formally different, but referentially the same, because they are realised differently while having the same propositional truth-value or refering to the same reality (cf. Milroy and Milroy 2012: 14–15). Sociolinguistically, however, these two sentences differ in the way they are perceived by speakers, due to the fact that their social and stylistic meanings differ. Milroy and Milroy (2012: 14–15), for instance, argue that the reason that one of these forms would be considered acceptable or correct "was probably

socially motivated, and the general compulsion to select one form out of a set of equivalents was a consequence of the trend towards standardisation". The first item, *I have not seen anybody*, is the neutral standard form, while the second one, *I have not seen nobody*, is the marked form that is perceived as non-standard, and, consequently, as incorrect or faulty. As such, it is associated with dialectal or uneducated speech, and would be avoided by educated speakers.

A broader perspective on usage problems thus foregrounds the most essential aspect of this phenomenon: usage problems are linguistic variants which have come to be particularly salient for speakers in the way that they encode stylistic or social meaning, and subsequently, a set of other properties related to prescriptive ideology. As Lippi-Green (1997: 30) puts it, "[w]e exploit linguistic variation available to us in order to send a complex series of messages about ourselves and the way we position ourselves in the world we live in". In this context, the study of attitudes towards usage problems among ordinary speakers of English acquires an additional and important dimension. Usage problems become ways in which people can perceive or express social identities. This realisation leads to another important distinction between usage problems that are more socially salient for ordinary speakers and those that are not socially salient. Thus, the working definition of the term 'usage problem' in the present study is that usage problems are two (or sometimes more) structural or lexical variants that are both used by at least some members of a speech community, but are seen as different in terms of correctness, acceptability, and style from the perspective of the standard language norm.

It is of course difficult to pinpoint what the standard language norm is, in the sense that the norm is always an ideal, or something that speakers strive towards. This is the reason that Milroy and Milroy (2012) argue that it makes more sense to speak of the standard as an ideology, and note that the ideal of the standard is almost impossible to achieve in practice, especially in spoken language. Accepting that the standard language norm is an ideal, or an abstraction, however, does not mean that such a norm does not exist or that it does not exert influence on speakers. In a particular language community at any one time, it is possible to come up with a set of rules about language use that demarcate standard from non-standard usage. The distinction between single and multiple negation is a case in point. The standard language norm dictates that single negation is the default unmarked variant, while multiple negation is the 'deviant' form which violates the standard norm. What I think Milroy and Milroy are referring to when they say that that the standard is an ideal, or an abstraction, is that despite the existence of single negation as the standard

form, speakers may still use multiple negation depending on many different factors. In other words, if we look at standardisation as the removal of optional variability from the language, then standardisation, to some extent fails in reality, and it is in that respect indeed an ideology. However, there are undoubtedly contexts where optional variability is minimised and where standard language norms are maintained. Academic texts, journalistic prose, and educational and institutional settings, for instance, are all contexts where the standard language norm is used and manifests itself. What I mean by the standard language norm, therefore, is the generally received or accepted norm with respect to what is perceived to be standard language use in a community. This is also the norm that is prescribed in usage guides. It should be borne in mind, though, that these norms can change over time, and there may sometimes be a mismatch between what is found in standard language use and what is prescribed in some usage guides.

Ilson (1985) discusses several criteria for establishing what constitutes a usage problem. The first criterion for a feature to be a usage problem "is that it should be a *problem*; that is, something that people actually say, rather than something they'd never dream of saying" (Ilson 1985: 166). Interestingly, this criterion is already expressed by Gould (1867), a usage guide writer from the nineteenth century, who notes that "[p]ossible, or imaginary, errors do not seem to be worth the trouble of exposure or refutation" (1867: iv). The second criterion for a usage problem identified by Ilson is that usage problems are usually features that are not restricted to a particular regional dialect, but are used across wider geographical space. This may relate in an interesting way to the notion of "vernacular universals" (cf. Chambers 2004; Nevalainen 2006b; Trudgill 2009). The final criterion, according to Ilson, is that the discussion of a particular feature should not reflect a social taboo – the reason that most slang expressions, for instance, are not usage problems. As far as the reasons for the existence of usage problems are concerned, Ilson (1985: 167–168) refers to various phenomena which might be considered to be contributory factors in the rise of usage problems. The first reason he cites is the idea that considerations of logic, aesthetics, or style dictate that certain language features are better than others. The second reason is that certain linguistic features become associated with the language of the lower social classes, and consequently become stigmatised through the association with qualities such as education, social class, or social standing. The third reason is based on Nunberg's unpublished work, and is related to the rise in the English speech community of a specific genre of language use, or discourse, which becomes central and especially valued; for the English speaking world this genre is the non-fictional

essay (Ilson 1985: 167). The final reason Ilson (1985: 167) cites relates to the concept of diglossia. He argues that "when English-language usage books recommend some items and condemn others, however arbitrarily, they may be responding to a need to maintain a distinction within Standard English corresponding to the High variety and the Low variety of some other languages". Building on this definition, Ebner (2017: 7) defines usage problems as "social constructs" which have a "divisive function in society", and whose use is conditioned by social, historical, and situational constraints. An important point mentioned by Ebner (2017) is that different usage features are associated with different levels of awareness; I return to this in the discussion of the results of the present study.

Another set of criteria for what makes a usage problem comes from Algeo (1991b: 2), who notes that "for something to be a question of usage, three factors must be present: there must exist alternatives of use; language users must be able to choose among them; and those same or other users must think that the choice means something". These three factors, alternatives, choice, and value, are "implicationally related" (Algeo 1991b: 3). They also correspond to some extent to the first of Ilson's criteria, which is that in order for something to be a usage problem it needs to be a problem, i.e. it needs to be a variant in the language. This account fits into the linguistic nature of usage problems, and brings us to the relationship between usage problems and linguistic variables, which is a central principle for the empirical study presented here.

The notion of the linguistic variable goes back to the sociolinguistic work of Labov (1972b: 8), who postulated a number of criteria for defining the linguistic variable; these criteria refer specifically to the selection of linguistic variables for the study of the social stratification of language. According to these criteria, a linguistic variable should be well integrated into the language system, it should be highly frequent in language, so that enough tokens can be collected from relatively short stretches of naturally occurring conversation, and it should be socially stratified. This involves ascertaining the possible environments of the variable, as well as the total number of environments in which it occurs (Labov 1972b: 71), thus allowing for an analysis of socially constrained variation following the "principle of accountability" (Labov 1972b: 72). It is also well known that Labov applied this approach mostly, though not exclusively, to phonetic variables. Subsequent sociolinguistic work has applied this notion to syntactic or grammatical variables as well, defining syntactic variables on the basis of two variant forms which have the same referential meaning (e.g. Sankoff 1972).

This approach is not without problems. Lavandera (1978), for instance, has pointed out the difficulty of assuming equal referential meaning of two syntactic variants, and thus, the difficulty in extending the notion of the linguistic variable from phonological to syntactic variables. This was followed by the development of a series of different approaches to conceptualising the linguistic variable (for an overview of this discussion, see Campbell-Kibler 2010: 424–425), from defining variable rules, to using the notion of the linguistic variable rather loosely, to refer to a set of variants used for "saying the same thing" (Chambers and Trudgill 1980: 80, quoted in Campbell-Kibler 2010: 425). The latter position is the one taken in the present study, specifically in the instances in which such an approach is relatively straightforward. While theoretical and methodological assumptions about linguistic variables are important, these issues are beyond the scope of the present study.

As mentioned above, the concept of linguistic variables is difficult to apply to certain usage problems. In the present study such difficulty is encountered in the analysis of the discourse particle *like* and non-literal *literally*. A variationist analysis proper of these two variants would involve ascertaining all possible environments in which these variants could occur, as well as identifying other linguistic variants which have the same meaning or function as the variants in question. Thus, for the discourse particle *like* this would involve identifying all variants which have the same meaning or function, and ascertaining all possible environments in which all of these variants could occur. In a similar vein, non-literal *literally* cannot be seen as a variant of the word *literally*, because what we have in this case is not the same denotational meaning expressed by different forms, but rather the opposite: the same form expressing different meanings. Approaching the use of non-literal *literally* in a variationist way would involve establishing a lexical variable such that *literally* would be one of at least two variants expressing the same meaning of intensification and emphasis. This kind of analysis was not undertaken here, because it requires extensive theoretical work in ascertaining the linguistic variables, which in itself entails problems which are beyond the scope of this study. A more straightforward approach was taken for the analysis of these two features, by relying solely on establishing the text-frequencies of both the discourse particle *like* and *literally* and by distinguishing three different uses of *literally*; this approach is described in detail in Section 4.4.

Approaching usage problems as linguistic variants allows us to make some observations which are important in establishing the approach to the question of the influence of prescriptive ideology on the variation in usage problems. The first relates to Labov's notion of sociolinguistic stereotypes (Labov 1972b: 139). Usage

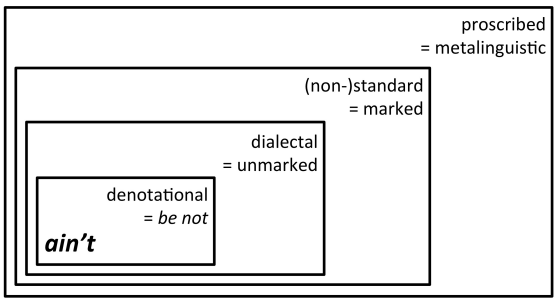

Figure 2.1: A model of the indexical order of meanings of *ain't*

problems are sociolinguistic stereotypes, because they are above the level of conscious awareness of speakers, and they are overtly commented on. This connection between proscribed features and the notion of sociolinguistic stereotypes has previously been observed in the context of *you was* (Laitinen 2009: 201). The second observation is that usage features are related to what Biber (1988: 36) defines as "the aesthetic function" of variation, which is stylistic in nature and refers to the attitudes and ideologies of language users with respect to correctness, acceptability or preference (Biber 1988; Albakry 2007).

A final important aspect of usage problems is that they are not all the same in terms of 'problematicity'. This point stems from the criteria or reasons for their emergence. Different usage problems are tied to different kinds of social values, and this means that their use may have different implications in different social contexts. This, I propose, could be usefully related to the notion of indexicality, which is understood as that aspect of the meaning of a language feature which comes from its contextual or pragmatic association with specific contexts of use (Silverstein 1976, 2003). Taking the notion of indexicality further, Silverstein (2003) develops the idea of indexical order, which, he argues, is necessary for the understanding of sociolinguistic meaning phenomena. On the basis of Silverstein's indexical order, I propose that a model can be developed for discussing the meaning of a particular prescriptively targeted feature. For instance, Figure 2.1 presents an example of this model, for the indexical order of meanings of *ain't*.[1]

On the very basic level, *ain't* is a negative form of *be*, and referentially expresses non-existence of something. On this level, *ain't* is referentially equal to any other

[1] I use the example of *ain't* for *be not*, but the model could be applied to the use of *ain't* for *have not*, and indeed, to any other prescriptively targeted language feature.

present tense form of *be not*, and it is on the basis of this referential equality that *ain't* is here considered a variant in the linguistic variable 'present tense *be not*'. The next level in the indexical order is acquired by the use of *ain't* in particular regional or social dialects. This association between such dialects and the use of *ain't* accounts for the development of an indexical relationship between *ain't* and its 'dialect' meaning, which is neutral at this point. One level up, the fact that *ain't* is part of a language variety which is not considered standard adds another order of indexical meaning which associates *ain't* indexically with non-standard language varieties. This association with non-standardness means that *ain't* would not be used in formal educated contexts, which would in turn develop the indexicality of uneducated, incorrect speech; this level thus differs from the previous one in that the use of *ain't* tends to be negatively evaluated. This brings us to the final level of indexical meaning in relation to *ain't*. At this level of the indexical order, *ain't* indexes the ideology of prescriptivism itself, as evidenced in cases where speakers use *ain't* on purpose, either to criticise or make fun of people who use the word (i.e. aligning themselves with prescriptivism) or to criticise or make fun of people who are sticklers about it (i.e. distancing themselves from prescriptivism).

Elsewhere, Silverstein (1996) addresses in detail the question of the semiotic processes by which standard language forms become indexical of particular qualities associated with the use of these language forms. In doing so, he identifies two semiotic processes – folk-extensionalisation and folk-intensionalisation – through which "social differentiations can be displaced onto linguistic differences in usage [...] and these latter can be perceived as a guide to and natural basis for the social differentiation that they index" (1996: 295). He also makes the point that the economic dimension of knowing the standard is an important aspect of the "culture of standardization", which is particularly relevant for certain social groups, such as yuppies, or yumpies. Silverstein's work is also important for his observations about the potential influence of the indexical order of meaning in language itself. Silverstein (2003), for instance, talks about how the meaning associated with the use of *he vs. they* is an example of how higher-order indexicality can influence language use and language structure. Given that these kinds of processes are a kind of prescriptivism – on the basis of Cameron's (1995) concept of "verbal hygiene" and Curzan's (2014) concept of "politically responsive prescriptivism" – this serves as a good example of how prescriptivism can be an influence on processes of language change. This thus provides an additional theoretical framework, and/or justification, for the hypothesis that prescriptivism influences language variation and change, as well as speakers'

attitudes, in important ways which need to be accounted for, and it also demonstrates the crucial place of speakers' attitudes and ideologies in the study and discussion of prescriptive influence, discussed in the next section of this chapter.

2.7 Ideologies and attitudes

The broad field of ideologies and attitudes bears directly on the question of prescriptive influence, because it is through the operation of attitudes and ideologies that speakers may or may not use a prescriptively targeted language variant. Furthermore, the importance of research on language ideology is evident in the fact that prescriptive attitudes to usage are usually associated with the existence of a standard language ideology (Milroy 2001: 530). As discussed in relation to the question of prescriptive influence in Section 2.4 above, large-scale studies of prescriptive influence on language variation and change patterns have shown that in general such influence is contextually or temporally limited at best. However, this is not to say that speakers are not influenced by prescriptivism; in fact, at any given point in time, particular speakers, or groups of speakers, may be more influenced by prescriptive ideas than others. Thus, an investigation of prescriptive influence also needs to be concerned with how such influence is manifested in the context of individual speakers. I already mentioned that this would be investigated on the basis of a study of the actual language use by speakers who are perhaps especially prone to prescriptive influence, but such a study is practically almost impossible on a larger scale. This is why I will be concerned here only with how prescriptivism affects speakers' attitudes and ideologies. More specifically, I draw on work done on language ideologies, language attitudes, and attitudes to usage. In what follows, I address each of these in turn.

Language ideology research provides important assumptions and observations regarding the influence of language ideology on speakers, and consequently on language itself. The term 'language ideology' has been used to refer to many different phenomena related to the relationship between ideology and language, understood in broad terms. For the present study, research into the so-called "ideologies *of* language" (Woolard and Schieffelin 1994: 55) is of particular relevance for the study of prescriptivism, as prescriptive ideology is one of the ideologies of language, and, we might argue, one of the more dominant ones (cf. Garrett 2001: 628). Many definitions have been offered of language ideology. Language ideologies have been defined as

"sets of beliefs about language articulated by users as a rationalization or justification of perceived language structure and use" (Silverstein 1979: 193, cited in Kroskrity 2004: 497), or as "the cultural (or subcultural) system of ideas about social and linguistic relationships, together with their loading of moral and political interests" (Irvine 1989: 255). It is clear on the basis of these definitions that prescriptivism is an ideological category.

This connection is further strengthened if we take into account the phenomenon of language prescription, on which prescriptivism is based (for a distinction between prescription and prescriptivism, see Section 1.2 above and Tieken-Boon van Ostade 2016). Prescription is a stage in the standardisation process defined by Milroy and Milroy (2012). Language ideology has been found to significantly affect language structure, albeit in a limited number of cases. Drawing on Silverstein's work, Kroskrity (2004: 496–497), for instance, cites the decrease, or near disappearance, of generic *he* as "[a] graphic example of the importance of multiplicity and contention in language-ideological processes, one that has noticeably changed the grammar of English within my generation's lifetime" (see also Bodine 1975; MacKay 1980; Pateman 1982). Although the example is indeed an instance of how speakers can change usage, and recent work on the different kinds of prescriptivism has dealt with this case as an instance of politically responsive prescriptivism (Curzan 2014), it may also be argued that it is debatable whether this has "changed the grammar of English", as Kroskrity argues, or merely the usage of pronouns or particular words; whether or not one considers this a change in the grammar of English would depend largely on one's definition of 'grammar'. In this respect, this instance also illustrates how, when talking about the influence of prescriptivism, one needs to carefully distinguish between changes in the structure of the language and changes in the rate of usage of a specific feature. Nevertheless, the example shows the power of language ideology to affect language use significantly.

Closely related to prescriptivism is the standard language ideology (e.g. Silverstein 1996; Lippi-Green 1997; Milroy 2001). Silverstein (1996) discusses the processes by which one variety, which has become identified as the standard, becomes indexically associated with specific personal qualities, as well as the function of the standard language ideology in establishing the linguistic economy in the United States in the twentieth century. Lippi-Green (1997) deals specifically with the issue of standard language ideology and its influence in twentieth-century America. She also makes a connection between the work of Foucault and the idea that language ideology of any kind has to do with organising, controlling, and directing language, or having

power over what language variety is going to be used in specific social contexts and functions. It thus becomes clear how language-prescriptive ideology, as an extension of the standard language ideology, can be explained from a theoretical point of view in this way. Lippi-Green's definition of ideology is more critical, and is concerned with unearthing power differentials and the role of ideology in creating and perpetuating those power differentials. The language ideology that relates specifically to the standard is also addressed by Woolard and Schieffelin (1994), in their comprehensive review of language ideology research. They note that "codified, superimposed standard languages are tied not only to writing and its associated hegemonic institutions, but to specifically European forms of these institutions" (Woolard and Schieffelin 1994: 64). Furthermore, they also stress that the ideological nature of standard languages means that ideas about language standards are naturalised and considered to be fundamental or essential to language, rather than forming a linguistically arbitrary, but socially or culturally conditioned category. This observation also applies to prescriptive ideas about language, which owe their persistence across centuries to their naturalised state. These kinds of standard or prescriptive language ideologies have an important role to play in power differentials in a society; this issue has been looked into perhaps most extensively by Lippi-Green (1997) in the American context.

Apart from the influence of language ideology on social relations and power, a crucial question with respect to prescriptivism is whether language ideology affects language structure and language change. As Woolard and Schieffelin (1994: 69) point out, "modern linguistics has generally held that linguistic ideology and prescriptive norms have little significance – or, paradoxically, only pernicious – effect on speech forms (although they might have some less negligible effect on writing)". However, work on language ideology has shown that in certain cases such as "gender in English, T/V pronoun shift, and Javanese speech levels, Silverstein shows that rationalization not only explains but actually affects linguistic structure, or rationalizes it by making it more regular. To understand one's own linguistic usage is to potentially change it" (Woolard and Schieffelin 1994: 70). This work provides important evidence that prescriptivism can indeed affect language change. However, when we talk about prescriptive ideology, it is important to point out that this ideology may be manifested broadly in two different ways. The first way is through institutionalised discourses on language correctness and rules of language use. This kind of ideology will be explored in the present study through the analysis of the usage guide genre (see Section 2.2). The second important site of prescriptive ideology can be found in

the attitudes of ordinary speakers. This perspective is crucial for understanding the effects of prescriptivism, because widely held prescriptive beliefs about language are sometimes in line with institutionalised prescriptions, while other times depart from such prescriptions.

Language ideology research is directly related to research on language attitudes. In current language attitude research, "[l]anguage ideologies provide the organizational schema through which linguistic diversity is viewed, interpreted and evaluated. In this sense, language ideologies represent broad, socio-cultural schemas that shape the development of intrapersonal attitudes towards particular language varieties and their speakers" (Dragojevic et al. 2013: 11). In relation to prescriptive ideology, a particular set of linguistic features or variants becomes associated with the standard language, as well as with the notion of correct language, and becomes indexically associated with certain positive values or characteristics (Dragojevic et al. 2013: 9–10).

Research on language attitudes has been carried out extensively in different subfields of sociolinguistics and social psychology. It is important at this point to distinguish between language attitude studies and studies of attitudes to usage, as these terms appear to be used in research arenas that have different histories, preoccupations, and research questions. Language attitude studies have been the focus of different subfields of sociolinguistics, the sociology of language, and social psychology. These studies deal predominantly with attitudes to phonological or suprasegmental variation in language, and rarely with attitudes to syntactic, semantic, or lexical variation (Finegan 1985; Giles and Rakic 2014). By contrast, studies of attitudes to usage deal specifically with prescriptive usage norms. Such studies have been very rare in English, but the ones that have been conducted provide important starting points for the present research. In what follows I will briefly outline the most important aspects of language attitude research relevant here, and I will then focus specifically on studies of attitudes to usage.

Language attitudes, understood as attitudes which are very closely associated with the language of others, as a topic of modern sociolinguistic research goes back to sociology of language research on attitudes in bilingual settings (e.g. Agheyisi and Fishman 1970). Methodologically, Tucker and Lambert (1972) pioneered the matched-guise technique (MGT) for the purpose of studying attitudes to language varieties in an indirect way (an overview of MGT methods can be found in Campbell-Kibler 2006: Chapter 3). This research was predominantly concerned with uncovering unconscious or implicit attitudes to language, and it was usually conducted in bilingual contexts. Another important development in language attitude research

is perceptual dialectology (e.g. Preston 1999a,b; Long and Preston 2002). In this tradition, perceptions about dialects and varieties are investigated empirically, using a variety of different experimental methods to assess language attitudes. In all of these branches of research on language attitudes, the underlying assumption is that attitudes are implicit, that speakers are usually not aware of them, and that attitudes require a special set of methods to be elicited. Some recent examples of this include cognitive sociolinguistic work done on language attitudes of Dutch speakers by Speelman et al. (2013), whose main goal was to investigate automatically activated language attitudes, which are impossible to access through direct approaches.

This kind of research is also characterised by a marked lack of concern with prescriptive attitudes. Prescriptivism is usually not discussed; when it is, this is often in introductions to language attitudes, as a way of introducing the notion. An illustrative example can be found in Garrett (2010: 6–10), where he notes that sometimes language use "evoke[s] attitudes with a somewhat (though not entirely) different focus, relating to public controversies over language usage". Here he mentions usage problems such as the use of *hopefully* to mean 'I hope that', rather than 'in a hopeful manner', and double negatives, and discusses the influence of the standard language ideology on speakers. One study which may be considered an exception to this observation is Albanyan and Preston (1998). The authors of this study investigate the attitudes to standard language norms among a group of American university students. To conclude, while language attitude research has provided an important methodological background for the present study, discussed in detail in Chapter 4, most of this research does not provide much information on the particular prescriptivism-related features I am concerned with here. In addition, as I have mentioned elsewhere, work done by Ebner (2017) in the context of the Bridging the Unbridgeable project has explored prescriptive attitudes to usage as well, specifically focusing on speakers' attitudes in British English. Ebner's study is thus in some respects a counterpart for British English to the present study, which is devoted to American English.

For specific information on speakers' attitudes to usage, particularly in relation to usage problems, a number of studies on attitudes to usage provide important evidence for those attitudes among certain groups of people, as well as their changes over time. The most notable studies of attitudes towards usage in English include Leonard (1932), Marckwardt and Walcott (1938), Mittins et al. (1970), and Crisp (1971). The first of these usage surveys, conducted by Leonard, investigated the attitudes to English punctuation and grammar usage of a group of what is described in the

study as "cultivated speakers". The goal of the study was to obtain insights into the contemporary norms of usage on the basis of an investigation of the attitudes to usage of educated speakers, and subsequently to use these insights to provide up-to-date guidance and relevant advice on usage. Dictionaries were seen as limited with respect to providing advice on contemporary usage, because of the time lag of a few years before established usage is recorded in them; grammars were seen as having the additional disadvantage of being "based on traditional pronouncements of dubious value" (Leonard 1932: 95). Leonard thus undertook the study of the then current standard of usage, with regard to both punctuation and grammar, by using questionnaires with various problematic constructions to collect rankings of these problematic construction by so-called judges, who included linguists, authors, editors, businessmen, and teachers, who were believed to "constitute a significant sampling of cultivated usage" (Leonard 1932: 96). The study consisted of two questionnaires, one with 102 problematic expressions, and the other with 130. The rating scale used by the judges distinguished four levels of usage: "Literary English", "standard, cultivated colloquial English", "trade or technical English", and "naif, popular, or uncultivated English" (Leonard 1932: 97). On the basis of those ratings, items were ordered from 1 to 230, and divided into three levels of usage: "established", "disputed", and "illiterate" (Leonard 1932: 99). The rated sentences contained a variety of items, from cases like *had better* or *point of view*, which were found to be established usages, to traditional usage problems, such as *ain't*, the split infinitive, flat adverbs, and singular *they*. The results showed that at the time some of the traditional usage problems were considered acceptable, such as the split infinitive, while others were considered "illiterate", such as *ain't*. One of the conclusions of the study is that "grammar is seen to be not something final or static but merely the organized description or codification of the actual speech habits of educated men" (Leonard 1932: 188). Furthermore, the study showed that on the basis of the ratings by the judges, certain usages considered incorrect in handbooks of usage had actually become part of established usage. Finally, it is also worth noting that aside from the pioneering character of this study in the area of surveys of attitudes to usage, its function and application is very clearly established in the area of teaching. What the study aims to do, first and foremost, is to provide current and reliable advice to teachers of English or composition regarding the accepted usage of the time.

Marckwardt and Walcott (1938: 2–3) supplemented Leonard's study by carrying out his initial intention of comparing attitudes to usage with facts of usage. Even though that had been Leonard's original intention, it was not achieved with his 1932

monograph, which "deals primarily not with usage itself but with opinions about the usage of words and expressions usually questioned or condemned in grammars and handbooks". The aim of the Marckwardt and Walcott study was to add to Leonard's survey of opinions the "facts about current English usage", as the title of their study suggests. Their investigation of the facts of usage was based on the *Oxford English Dictionary* and its supplement; in addition, they also consulted a number of other works, including Hall (1917) and Horwill (1935). By comparing the labels for the established, disputed, and illiterate usage items used in Leonard (1932), Marckwardt and Walcott found that the opinions of the judges were predominantly conservative. With respect to the established usage items, for instance, they found that very few of those were colloquial. For the disputed usage items, it appeared from their analysis that the majority of the usages considered disputable "are, on the basis of the recorded fact, actually in cultivated use today" (Marckwardt and Walcott 1938: 49). Finally, with respect to "illiterate" usage items, they concluded that "illiterate" was too strong a word to designate what would more realistically be described as non-standard, or regional, usage.

Crisp (1971) is a replication of Leonard's approach to assessing opinions on English usage, conducted about forty years later than the original study. His additional purpose was also to investigate the facts of English usage, and compare the gap between fact and opinion to that observed forty years before. As mentioned above, the opinions recorded by Leonard were found to be generally fairly conservative when compared to descriptions of actual usage on the basis of the study by Marckwardt and Walcott. Importantly, Crisp hypothesised that this conservatism in opinion will have decreased somewhat, and expected that attitudes would have been more liberal at the end of the 1960s. An additional variable that Crisp introduced in his study was the grouping of data by geographical region – something not done by Leonard. Crisp's questionnaire included 215 items, based on Leonard (1932) and Marckwardt and Walcott (1938); Crisp also adapted some of the descriptions of levels of usage used by Leonard. The category "illiterate", was, for instance, replaced with "non-standard" (Crisp 1971: 63). Crisp remained fairly faithful to Leonard's approach to selecting informants, collecting the opinions of 1764 informants, consisting of linguists, teachers of English, editors of magazines and news media, dictionary makers or editors, and businessmen (Crisp 1971: 67). On the whole, Crisp also found that what he calls language specialists (a group of informants he compared to the linguists in Leonard's study) were more liberal in their ratings compared to those in Leonard's study. Crisp also found that of the various groups of informants, language specialists

were the most liberal group, while members of the panel of the *American Usage Dictionary* were the most conservative raters. With reference to particular usage items, Crisp identified a number of items which moved in their ratings from "disputed" to "established". On the basis of his findings, it may perhaps reasonably be hypothesised that over time speakers had become more liberal in their attitudes to usage.

Finally, it is worth mentioning that a similar survey of attitudes to usage was conducted for British English by Mittins et al. (1970). In part, at least, this study is similar to Leonard's study in their shared concern with attitudes to usage in education and the teaching of usage standards. Mittins et al. (1970: 5) collected the opinions of 457 informants on 55 items of usage. Their informants consisted of similar groups of speakers to those included in Leonard's and Crisp's studies, but were more heavily skewed towards teachers. Unlike Leonard and Crisp, Mittins et al. included register variation in their questionnaire, by asking informants to judge the acceptability of usage items in formal and informal contexts, as well as in spoken and written contexts. One of the more general findings of this study was that people involved in teaching or teacher training are not necessarily the most conservative on matters of usage. This study is in part replicated by Ebner (2017), in a contemporary sociolinguistic take on the investigation of attitudes to usage in British English. Apart from the general conclusions of these surveys of attitudes to usage, they also contain a multitude of data on specific features, which are too lengthy to cover in detail here, but will be drawn upon in the discussion of individual usage problems in the analysis in Chapter 7, as and when relevant. In addition, despite their methodological limitations, these studies provide an important basis for the present study, as will be discussed in the next chapter.

A final strand of research worth mentioning is the research on attitudes to usage carried out in the context of predominantly historical linguistic and sociolinguistic studies; in this context the term 'attitudes' is used to refer to the normative or prescriptive types of attitudes understood as being overtly expressed and as being spread by institutional, top-down means. Leonard (1929) is often cited as a ground-breaking work of this sort, in which a modern linguist surveys the language attitudes of eighteenth-century normative grammarians in an attempt to show how diametrically opposed their approach to language study was, compared to the descriptive linguistics of the second half of the twentieth century (cf. Tieken-Boon van Ostade 2006). In that sense, we can perhaps trace this strand of research on attitudes to usage back to Leonard's survey. More recent examples include Sundby et al. (1991) and their account of primarily proscriptive views and attitudes to usage in

eighteenth-century normative grammars of English, a study of attitudes to the usage of phrasal verbs in the eighteenth and nineteenth centuries (Wild 2010), and a study of attitudes to preposition stranding from 1500 to 1900 (Yáñez-Bouza 2015). It is important to bear in mind that these studies can be seen as dealing with different types of attitudes. The studies discussed in the previous paragraphs of this section concern the language attitudes of speakers, and are often understood more generally (i.e. not necessarily related to prescriptive language ideology). The studies mentioned in this paragraph, however, use the term "attitudes" to refer to the attitudes to language standards, norms, correctness, or acceptability which are found in metalinguistic works such as normative grammars and usage guides. For instance, the attitudes to usage in Finegan (1980) are those found in books on language by lexicographers or language scholars from the nineteenth and twentieth centuries in the United States. Those that feature prominently as a research topic in historical sociolinguistic studies are the attitudes of grammarians or writers on language, and relate to the correctness or appropriateness of particular language features. This kind of division may suggest that such attitudes are different, but perhaps that is not the case. There are multiple reasons that the latter type of attitudes to usage form part of historical sociolinguistics, and some of these are related to the available sources of evidence. Historical sociolinguistic evidence on language attitudes comes from written texts, and most of the written texts available are predominantly books on language, which were written in a period when modern linguistic science did not exist, and the predominant language ideology was tied to the superiority of the standard. The writers of these books were concerned with distinguishing between correct and incorrect usage from the point of view of the language standard. In essence, normative and prescriptive grammar writing was essentially an attempt to regulate language use. However, despite the fact that those books presented the normative views of a group of people, they may also reveal information about the attitudes to usage found among ordinary speakers. In this context, the relevance of a study of present-day prescriptivism lies in its potential to shed light precisely on the relationship between attitudes to language found in popular metalinguistic works such as usage guides and attitudes of ordinary speakers.

2.8 Conclusion

In this chapter, I have surveyed important work in the study of prescriptivism and attitudes to usage in American English, and to some extent, British English. The

studies cited provide a significant body of knowledge about prescriptivism and its effects in the Anglo-American context. What my account of previous research has also shown is that there are still some open questions as to the influence of prescriptivism and the appropriate ways to study it. As the authors of one recent study note, any kind of study of the effects of prescriptivism on language variation and change "must begin with improved empirical description" (Hinrichs et al. 2015: 807). In the next chapter, I will outline the methods used in the current study, which aims at an improved empirical description of prescriptivism.

CHAPTER 3

The language features: selection and previous studies

3.1 Introduction

The linguistic features I focus on in the present study are: *ain't*, the discourse particle *like*, non-literal *literally*, negative concord, pronouns in coordinated phrases, and the split infinitive; in this chapter I discuss each of these in more detail. These features differ from one another in various respects, and are analysed from the three different perspectives mentioned in Section 1.6 above. In addition to these, the part of the analysis that focuses on the influence of prescriptivism on actual use includes a number of additional language features, also commonly treated as usage problems. In this way, an attempt is made both to zoom in on the influence of prescriptivism on the six features mentioned above, and to zoom out to present a bird's eye view of the influence of prescriptivism, by accounting for multiple language features. Before presenting the methodological approach taken to the three-pronged analysis of these features, I will first discuss the selection of the six linguistic features. The selection process is covered in the first section of this chapter, where I also introduce the Hyper Usage Guide of English (HUGE) database. The subsequent sections are devoted to

each of the six features in turn. In these sections, I will provide brief accounts of each of them on the basis of previous studies.

3.2 Selection of usage problems

A useful tool developed for the analysis of usage problems is the Hyper Usage Guide of English (HUGE) database, compiled by Straaijer (2014) at Leiden University, in the context of the Bridging the Unbridgeable project. The HUGE database contains 77 usage guides, both British and American, and a total of 123 usage problems. In the database compilation process, the 77 usage guides were searched for entries on each of the 123 usage problems; the relevant entries were subsequently entered into the database, and tagged for a range of additional information (for more on the database, see Straaijer 2015, 2018). The database allows users to search for particular usage problems in various ways. It also allows users to explore the frequency with which a particular usage problem is treated in usage guides on the basis of the number of entries and guides that discuss that problem. Table 3.1 was produced using these search options, and shows the number of guides which discuss each of the usage problems in the database. While this ranking applies to all the usage guides in the database (i.e. both British and American), it nevertheless gives a good indication of what the most commonly treated usage problems in American English are. This ranking was the first step in the selection of the usage features investigated here.

Feature	Guides	Feature	Guides	Feature	Guides
shall / will	65	*-ic / -ical*	42	superlative comparison	26
different to / than / from	63	*lend / loan*	42	*-lily* adverbs	26
who / whom	63	*me / myself*	42	*hoi polloi*	25
lay / lie	63	*each other / one another*	41	*contemporary*	24
only	62	*it is I / it is me*	41	*likely*	24
split infinitive	62	*reason is because*	41	*could of*	24
I* for *me	61	*if / whether*	41	*dare*	23
singular *they*	59	*your / you're*	41	*more warmer*	23

less* / *fewer	58	*one of those who*	40	*in / under circumstances*	21
***none* in plural context**	55	*one ... one / he*	39	*'d rather*	21
data is* / *are	54	*them / their* + V-*ing*	39	*there's*	21
disinterested / uninterested	53	***ain't***	39	*corporeal / corporal*	20
neither ... nor ... are / is	53	*compare with*	39	*learn / teach*	20
try and / to	53	***hopefully***	38	*as well (as) ... or better than*	19
like / as	52	*than I / me*	38	*pretty*	19
nouns of multitude	52	*former / latter*	38	*the two first*	19
very unique	52	*equally as*	38	*upon*	16
apostrophe	52	*decimate*	36	double passive	16
a / an	52	*alternative*	36	*thusly*	16
both ... and	52	*flaunt / flout*	35	*like / the way*	15
between / among	51	*off of*	35	*have went*	15
slow / slowly	51	false attraction	35	split auxiliaries	15
who(m) / which / that	51	*on to / onto*	34	*off / from*	14
preposition at end of sentence	50	*either is / are*	34	*quicker / more quickly than*	14
aggravate	50	*most perfect*	34	*omission of relative pronoun*	14
snuck and *dove*	50	*whose / of which*	34	*gay*	14
dangling participle	49	*(not) as / so far as*	33	*thankfully*	14
was / were	49	*may / might*	33	*demonstrative them*	14
me* for *I	49	*from thence*	32	*meet with / meet up with*	13
foreign plurals	49	*like / as if*	31	*all that / so easy*	12
due to / owing to	48	*either of them / each of them*	31		

effect / affect	48	***But / And***	31	*at / in*	11
infer / imply	47	*have to / have got to*	30	*less / least*	9
literally	47	comma splice	30	*when*	8
alright / all right	46	subject-complement	30	*get thither*	6
this / these sort of	46	*averse to / from*	29	*evenings* and *Sundays*	4
compound subject	45	*in / into*	29	*at (the) university*	4
double negatives	44	*spoonsful*	29	*momentarily*	1
that / which	44	*either ... or ...* (*or ...*)	28		
mutual	43	*providing / provided*	28		
can / may	43	*family is / are*	27		
farther/further	43	*very / much amused*	26		

Table 3.1: Usage problems in HUGE, sorted by the number of guides that discuss them; features included in the analysis of prescriptive influence in texts (cf. Section 6.8) are given in bold

The second consideration in the selection process was to include a variety of usage problems. The difference among usage problems can be conceptualised on the basis of a number of criteria, as I have discussed elsewhere (Kostadinova 2018b), and which I will briefly outline here. First, usage problems may differ on the basis of the various strands of prescriptivism distinguished by Curzan (2014: 24–39). Thus, there are usage problems which relate to the standardising function of prescriptivism, and these usually involve distinguishing between standard and non-standard variants, such as *ain't* and negative concord. In addition to standardising usage features, there are usage features characteristic of stylistic prescriptivism; these features relate to different levels of formality and style in language usage. A good example of a standardising prescriptivism-related feature is the split infinitive. Second, a particular type of usage problems can be distinguished in the context of those features which are relatively recent and ongoing changes in the language, and, as a result of the development of new functions or uses, are seen as incorrect or unacceptable. These include the use of *literally* as a modifier of non-literal expressions, and the use of the discourse particle *like*. A third distinction has been made by Albakry (2007), who distinguishes

between "weakly dispreferred" features, of which the split infinitive is an example, and "strongly dispreferred" ones, such as *literally*. One of the aims of the present study is also to ascertain which of the six features are strongly dispreferred and which weakly. The selection of the six features was thus done in such a way that each of these distinctions is exemplified by one of the language features. Beyond these considerations, the choice was to some extent arbitrary, and the selection was made in order to keep research within manageable bounds, as each of the features selected is studied in depth from three different perspectives.

For one part of the analysis of prescriptive influence, specifically for the case of the split infinitive, additional features were used in order to 'measure' or operationalise the influence of prescriptivism on individual texts. This approach is based on two assumptions (cf. Hinrichs et al. 2015). The first assumption is that prescriptive influence can be detected more specifically not by focusing on large sections of large corpora, but by focusing on individual texts. The second assumption is that if individual texts have been influenced by prescriptivism, this influence will affect a number of prescriptively targeted features, not just one. The additional features selected were: sentence-initial *and/but*, singular *data*, *hopefully*, *less* with plural nouns, *these kind/sort of*, *try and* (instead of *try to*), plural *none*, passives, *shall*, and *whom*. The frequency of occurrence of these features was counted for separate texts in the corpora used, in order to see whether the occurrence of a prescriptively targeted feature correlates significantly with that of other such features (this is explained in more detail in Section 4.4).

In the following sections of this chapter, I will discuss each of the six variants separately, in alphabetical order. For each variant, I summarise the most important aspects of its use as reported in previous studies. More specifically, I address the variation in the use of each particular feature, the known factors which influence that variation, and the contexts of use of each particular feature as evidenced in previous accounts. Finally, I also address the extent to which prescriptivism has been discussed in relation to each feature.

3.3 *Ain't*

Ain't is a non-standard feature that is regularly mentioned in descriptive grammars of the twentieth century. Curme (1935: 248), for instance, discusses it as a variant for *am I not?* or *am not I?* in colloquial speech. He further notes that "colloquially *ain't* is

often felt as a useful contraction in *ain't I?*, but it is elsewhere shunned" (Curme 1935: 248), an observation largely in line with the opinion expressed by many usage guide writers (see Section 5.2). Curme (1935: 250) also mentions *ain't* as a less common variant pronunciation of *hain't*, with dropping of *h*. Jespersen (1940: 431) describes the use of *ain't* for *han't* "as a vulgarism (h dropped)", while Quirk et al. (1985: 129) note that "*ain't* is a non-standard contraction commonly used (esp in AmE) in place of *am not*, *is not*, *are not*, *has not*, and *have not*". The account of the use and perception of *ain't* in Biber et al. (1999: 167–168) is informed by corpus evidence of its use, and, consequently, the authors provide some more details regarding the actual use of *ain't*, such as its frequency distribution across different genres; these details are further discussed below. Biber et al. (1999: 167–168) repeat the same characterisation of *ain't* as non-standard, although they also add that it is "relatively widespread in use" and that it "applies to all persons and may correspond to *be* and *have*". In Biber et al. (1999: 1122), *ain't* is also covered as part of the grammar of conversation. Huddleston and Pullum (2002: 1611) note that "in present-day English, *ain't* functions as a negative form for all present tense forms of *be* and *have*". In summary, all of these grammars mention the non-standard use of *ain't*, and Biber et al. (1999) and Huddleston and Pullum (2002) more directly address its status as a proscribed feature. The former, for instance, describe it as a "paradigm case of a frequent though unacceptable form" (Biber et al. 1999: 167), while the latter note that "its effective proscription has been one of the greatest successes of prescriptivists" (Huddleston and Pullum 2002: 1611). Examples of the most common uses of *ain't* are given in (5)–(7).

(5) He thinks he **ain't** a man any more. (1987, fiction, COHA[1])

(6) He **ain't** saying that to my face. (2006, spoken, COCA)

(7) You **ain't** said yes yet. (1932, fiction, COHA)

The correspondence of *ain't* with forms of *be not* and *have not* is likely related to its derivation, which has been discussed in a number of studies (Jespersen 1940: 433–434; McDavid 1941; Stevens 1954; Cheshire 1981: 366–367; Anderwald 2002: 117–121). These accounts of its derivation and historical development sometimes contain important information about the social evaluations of *ain't* across history. As

[1]COCA is the Corpus of Contemporary American English, and COHA the Corpus of Historical American English. See Section 4.4 for details on the corpora.

mentioned above, according to Jespersen (1940: 431) one line of development of *ain't* is as a form of *have not*, through *han't*, where *ain't* can occur as a vulgarism, with the *h* dropped. Jespersen (1940: 434) finds that "in the 19th and 20th c. *an't* and *ain't* are frequent for *is not* as representing vulgar speech". McDavid (1941: 59) notes that "both [aj ejnt] and [aj ant] represent normal phonological developments in different dialect areas"; however, he continues, "the social prestige of the dialects in which the latter was the normal form is responsible for the odium of vulgarity attached to the former, and the influence of normative grammar and pseudologic completed the work of depriving Standard English of one of the inherited forms of the negative paradigm of the verb *to be*". This is similar to Huddleston and Pullum's observation on the success of prescriptivists on stigmatising *ain't*. Denison (1998: 195–197) includes *ain't* among non-standard contractions, also noting that in the eighteenth century *ain't* was found in the colloquial speech of the educated upper classes. This illustrates that *ain't*, like most other usage problems, is another example of the arbitrariness of the attitudes and social stigma associated with it, which is in part revealed by the more nuanced way in which *ain't* could be used in the past (see also Denison 1998).

Ain't has been looked at rather extensively in studies of non-standard varieties of English, such as non-standard British English (Cheshire 1981, 1982; Anderwald 2002), African American English (Labov et al. 1968; Wolfram 1969; Weldon 1994; Howe 1997; Howe and Walker 2000; Walker 2005), Puerto Rican English (Wolfram 1974), Appalachian English (Wolfram and Christian 1976), and Southern White non-standard English (Feagin 1979, cited in Howe and Walker 2000). These studies have shown that *ain't*, generally and predominantly, corresponds to present forms of *be not* and *have not* (Cheshire 1981: 365–366; Weldon 1994; Howe 1997: 270; Kjellmer 1997: 169; Anderwald 2002; Walker 2005; Wolfram and Schilling 2016: 385. *Ain't* cannot function as an equivalent for the full verb *have* (Cheshire 1982: 366; Anderwald 2002: 116–117). In African American English *ain't* has also been found to correspond to *did not* (Weldon 1994; Howe 1997: 273; Walker 2005; Wolfram and Schilling 2016: 386), and sometimes other forms such as *do not* (cf. Howe and Walker 2000: 119–123; Walker 2005: 2). In non-standard British English, Anderwald (2002: 146–149) found that it is quite rare, but not impossible, for *ain't* to correspond to a variety of other auxiliaries, and, sometimes, modals. Some, such as Labov et al. (1968: 178), argue that "*ain't* is merely a negative marker, with no current relation to *isn't*, *aren't*, etc. from which it is historically derived". However, it has been shown that "*ain't* is basically restricted to negating *have* and *be* and occurs largely in the present tense, just as in non-standard varieties of English" (Walker 2005: 2).

In terms of factors influencing the use of *ain't*, previous studies have focused on type of auxiliary, tense, type of subject, presence of negative concord, type of clause, and aspect. In the context of non-standard British English, Cheshire (1981: 368) has found that *ain't* is most frequent as a variant of auxiliary *have not*, followed by *ain't* as a variant of copula *be not*, and it is least frequent as a variant of auxiliary *be not*. These findings are also in part supported by Anderwald (2002: 124), whose data show that *ain't* for *have not* is more frequent than *ain't* for *be not*, auxiliary and copula taken together (see also Anderwald 2004: 186). Weldon's results show a similar pattern of occurrence for *ain't* in African American English, although she also shows that the preference for *have not* over *be not* is not statistically significant. Walker's data on Early African American English show that *ain't* for *be not* is preferred in this variety, unlike in non-standard British English (Walker 2005: 12). Howe and Walker (2000: 114) present useful data about the proportion of *ain't* for *be not* and *have not* in a number of American English varieties; their data show that *ain't* in African American English does not exhibit any pronounced preference for one auxiliary over the other, while Southern White non-standard English favours *ain't* for *be not* over *ain't* for *have not* (Feagin 1979: 226, cited in Howe and Walker 2000). Wolfram and Christian's (1976: 116) data on Appalachian speech also show a slight preference of *ain't* for *be not* over *ain't* for *have not*. All these findings taken together suggest that *ain't* for *be not* may be more frequent than *ain't* for *have not* in general American English, although, given that these studies are all based on vernacular speech, we may expect the situation to be different in the corpus data analysed in the present study. Howe and Walker (2000: 116–117) also look at how tense and aspect affect the variation of *ain't*, finding that "*ain't* is basically restricted to present temporal reference, and with respect to aspect has essentially the same distribution as *be* + *not*". In terms of person and number, Wolfram (1974: 154) found that *ain't* more frequently corresponds to *are* and *is* than to *am* in Puerto Rican speech. Weldon's African American data did not produce significant results for the type of subject and grammatical number constraint, while Walker (2005) found that both *ain't* and *not*-contraction are more likely with subjects realised by pronouns in Early African American English. Presence of negative concord was found to positively predict the occurrence of *ain't* in American English varieties (Wolfram 1974: 154; Wolfram and Fasold 1974: 162; Weldon 1994: 379; Walker 2005), while this was not the case in non-standard British English data (Cheshire 1981). Constraints relating to type of clause were found to be significant in non-standard British English, with *ain't* being more likely with tag questions, as opposed to declarative clauses, but this correlation

has not been shown in American English data. Finally, aspect has been shown to be significant in Early African American English (Walker 2005), but not in African American English (Weldon 1994); this may have been influenced by the difficulty in coding for this constraint, or the use of different definitions of stativity (Walker 2005).

All of these studies have shown that *ain't* is overwhelmingly used in non-standard varieties of English in the United States, specifically African American English. In the context of these uses, *ain't* has been found to be more likely to correspond to *isn't* or *aren't* than to *am not*, and it is favoured in clauses with negative concord (Wolfram and Fasold 1974: 162; Wolfram 1974: 154). These results are particularly interesting from the point of view of the higher acceptability of *ain't* when used in the first person singular, observed in the treatment of this feature in usage guides (see Section 5.2), because they suggest that this attitude may not necessarily be reflected in the usage patterns of *ain't*. I return to a discussion of this topic in Chapter 8. These constraints were not found to play a role in studies of *ain't* in non-standard British English, where syntactic environment was found to be the significant constraint, with *ain't* being more likely to occur in tag questions than in declarative sentences (Cheshire 1981: 369). Weldon (1994: 375), in her study of *ain't* as a negation feature of African American Vernacular English, found that, in the environments for copula *be*, there were no statistically significant constraints on the use of *ain't*, that is "*ain't* is insensitive to person-number distinctions in copular environments".

While these studies have revealed a great deal about the intralinguistic variation of *ain't*, and have provided some information about the social variation of *ain't*, they are unclear about the influence of prescriptivism on the variation patterns of this feature, although each of these studies does mention its stigmatised status. Most of these studies have not looked explicitly at any prescriptivism-related constraints on the variation of *ain't*, with the exception perhaps of Anderwald (2002: 124), who notes that the "status of high stigma might explain the relatively low occurrence of AIN'T as well as of neg concord, compared to other non-standard forms".

3.4 The discourse particle *like*

The discourse particle *like* is in many respects an atypical usage problem: it is a fairly recent target of prescriptive criticism (see Chapter 5), it is a case of robust linguistic change in progress, and it is a discourse level feature, rather than a grammatical or a lexical feature. Discourse level features are quite rare in the prescriptive canon, but the

case of *like* represents a complex process of variation and change (cf. D'Arcy 2006, 2007), which makes this word particularly salient for prescriptivists, and, indeed, for ordinary speakers. In linguistic accounts of the use of this feature, a distinction is often made between two general uses of *like*: the traditional uses, "regularly attested in dictionaries", exemplified in (8)–(10) below, and the newer, "conversational" uses in (11)–(14) (Schourup 1983: 28; D'Arcy 2006: 339).

(8) I **like** all the pretty trees. (2004, magazine, COCA)

(9) My cousin and I would fight **like** cats and dogs. (2004, magazine, COCA)

(10) She was combing her hair, just **like** you saw then. (2008, spoken, COCA)

(11) I was at my friend Ron's house and then somebody just said, **like**, guess who I did it to, and everybody was **like**, who, ... (2004, academic, COCA)

(12) I don't think he was, **like**, five times braver than me. (2007, academic, COCA)

(13) Is there some way I can, **like**, throw a bouquet to him at the hearing today? (1992, fiction, COCA)

(14) There is more, but they're not going to tell us, **like**. (2004, academic, COCA)

A significant number of linguistic studies have accounted for various aspects of the development and the sociolinguistic variation of these newer uses of *like*. In an early discussion of the discourse particle *like*, for instance, Underhill (1988: 234) argues that while this use of *like* is not considered grammatical in standard English, "it is neither random nor mindless". Underhill's account distinguishes between *like* functioning as an approximator and as a hedge, but mostly focuses on its function as a marker of focus and new information; he uses the term "discourse particle". Another function of *like* is the discourse marker function. Fuller (2003) provides a good summary of the difficulty in ascertaining the functions of the discourse marker *like*; her overview, however, seems to suggest that there are two different functions: the approximating function, to indicate "looseness of meaning", and the focusing function (Fuller 2003: 369). While some cases are clear uses of one or the other of these functions, in others the function of *like* may be more elusive. A crucial criterion in ascertaining whether a particular case of *like* functions as a discourse marker is whether it changes the propositional meaning of an utterance, which is not always clear-cut. Furthermore, the functions of certain uses of *like* are difficult to analyse from this perspective, because this use is highly subjective. A third conversational function of *like* is the

quotative *like* (e.g. Blyth et al. 1990; Romaine and Lange 1991; Dailey-O'Cain 2000; Macaulay 2001). A useful classification of these and other uses of *like* is provided by D'Arcy (2007: 387), who distinguishes between the following functions of *like*: quotative complementiser (11), approximative adverb (12), discourse particle (13), and discourse marker (14). D'Arcy (2007) calls these uses of *like* "vernacular uses".

These varied vernacular uses of *like* are also characterised by a difficulty in ascertaining clear functional boundaries between the various uses. However, all of these uses seem to be perceived as wrong and inappropriate from the point of view of prescriptive ideology. Consequently, I will not be concerned with distinguishing between these three, and will essentially focus on all these vernacular uses together; in what follows I will use the term "discourse particle" to refer to all of these uses.

In terms of sociolinguistic constraints on the use of discourse marker *like*, Fuller (2003: 370) found that in the more formal interviews in her data, speakers tended to use *like* more frequently than in her casual conversation data, contrary to expectation. She suggests that while the interviews might not have been experienced as formal speech events by the speakers, this could also be explained by the pragmatic usefulness of *like* in interviews. Fuller (2003) also found that female speakers used discourse marker *like* more than male speakers, but also that use of discourse marker *like* varies depending on the conversation context, or, more specifically, the rapport between the interlocutors. Fuller (2003: 375) concludes that "while *like* may be a marker of casual speech and age group, it is also – and perhaps primarily – part of the natural repertoire of many speakers, who use it strategically yet unconsciously in their everyday speech". Fuller (2003: 369) also mentions that discourse marker "*like* has long had the connotation of being a marker of superficiality and lack of intelligence [...], and as a stereotype of the (inferior) dialect of American English".

In the context of the usage guide tradition, the discourse particle *like* represents a newer usage problem. It is certainly not an old chestnut, as shown by the fact that it is not included in the list of usage problems in the HUGE database. However, since the social stigmatisation of the discourse particle *like* is fairly widespread, and the form is a salient problematic usage, the form was included in this study precisely for these reasons. In the context of this feature, I will try to investigate whether and how forms which are newer in the language and negatively evaluated by speakers are included in usage guides.

3.5 Non-literal *literally*

From a descriptive point of view, *literally* is "a case of semantic change in progress" (Israel 2002: 424).[2] Alongside the basic use of *literally* to mean 'in the literal sense', or 'word for word', as in example (15) below, the word has come to be used as a modifier of non-literal expressions, exemplified in (19). The function of *literally* is that of a stance adverbial (Biber et al. 1999: 767), or a metalinguistic adjunct (Huddleston and Pullum 2002: 775), used to clarify how a word or phrase is to be understood. Speakers can use *literally* as a stance adverbial "to convey their judgements and attitudes, to claim the factual nature of what they are saying, and to mark exactly how they mean their utterances to be understood" (Biber et al. 1999: 767). This subjective attitudinal component in the use of *literally* is what accounts for the variation observed in its use in present-day English (Powell 1992). Quirk et al. (1985: 618–619) cover *literally* under adverbials denoting metalinguistic comment. They make the point that "metalinguistic comment is inextricably mixed up with expressions of degree" (Quirk et al. 1985: 619). For instance, in cases such as *He almost stole the money*, "we cannot be sure whether it means that he came close to stealing it or acted in such a way that it could almost be called 'stealing'" (Quirk et al. 1985: 619). As a result, these adverbs – and, according to Quirk et al. (1985: 619), "this is especially true of *literally*" – are used as emphasisers.

(15) The Gaelic word for whiskey, Usquebaugh, **literally** means "water of life." (1990, academic, COCA)

(16) I was seated next to Al Giddings, an underwater photographer who once **literally** pulled another photographer from the jaws of a great white. (2006, magazine COCA)

(17) DeLauro is not alone, says Dr. Judy Kuriansky, who **literally** wrote the book on the subject, called "How to Love a Nice Guy." (1991, newspaper, COCA)

(18) Obsessing over healthy food is a decades-long pastime in L.A. Local grocery stores such as Erewhon and Whole Foods Market and restaurants such as Real Food Daily **literally** feed the frenzy, making the city a breeding ground for people attracted to a "pure" lifestyle. (2001, magazine, COCA)

(19) I don't even know who these people are, and suddenly they have **literally** exploded into the American consciousness. (2009, spoken, COCA)

[2]This section is an extended version of a similar section included in an article on *literally*, to appear in *English Today* (Kostadinova in press).

The variation in the use of *literally* is in a sense reflected in the existing accounts of these uses (Powell 1992; Israel 2002; Nerlich and Domínguez 2003; Calhoun 2015). Powell (1992), for instance, distinguishes between five different categories of use, while Nerlich and Domínguez (2003) distinguish between three groups of uses, further subdivided into more specific uses. Despite the differing classifications, they do reflect the general pattern of variation in the use of *literally*. The variation among classifications of the use of *literally* is also likely to be a consequence of the highly subjective and metapragmatic nature of most of the uses. In what seems to be the earliest detailed analysis of the present-day uses of *literally*, Powell (1992: 337) examines "five categories of contemporary use" and compares the basic use of *literally* – what she refers to as "folk definitional" use – to four other types of uses, which are historically speaking later, and have developed from the basic use. Crucially, Powell (1992) argues, "a normative attitude of aptness accompanies all uses". The five uses Powell distinguishes are: 1) *literally* in folk definitions, used to refer to what something means, how something is read, meant, interpreted or translated, or as an equivalent of *technically*; 2) *literally* with lexemes denoting extreme cases, including number expressions of exceptional quantity; 3) *literally* in dual readings, where it modifies a conventionally idiomatic or figurative expression, and it signals that the expression should be interpreted literally; 4) *literally* with formulary constructions as a semantic innovator; and 5) *literally* with non-literal expressions, as an aesthetic justifier. These uses are exemplified in (15)–(19), respectively. According to Powell, the most interesting cases are those where *literally* is used to comment on a non-literal reading: "contrary to what one might expect, this use is neither odd nor paradoxical; rather, it illustrates that the lexeme exhibits great continuity of function in both literal and non-literal environments" (1992: 337).

In the first use, *literally* signals that the expression it modifies should be taken in the strictest sense of the word. It is usually used to refer to what something means, as in (15) above. Historically, this use is the oldest, and was found predominantly in contexts related to the interpretation of the Bible and other sacred texts (Israel 2002; Nerlich and Domínguez 2003), where the difference between literal and figurative readings had important moral implications (Powell 1992). Even though in this use the compositional meaning of the lexeme *literally* is the strongest, a crucial component of this use is the attitudinal dimension entailed, which refers to the speaker's attitude to the fit of word-to-word or word-to-world (Powell 1992: 341). It is this attitudinal component that accounts for the development of the newer uses of *literally*, including those with non-literal expressions. The second use of *literally* is with expressions

denoting extreme cases, where "the presence of *literally* is intended to force a non-hyperbolical – that is, a literal – meaning", while simultaneously increasing the "rhetorical emphasis on the extreme case" (Powell 1992: 342). This use encompasses the frequent occurrence of *literally* with number expressions, where it "signals not only that the approximation is accurate in its range, but also that the speaker judges this number to be significant in its largeness" (Powell 1992: 342). In other words, insisting that what one means is what one is saying "is to acknowledge that it might not sound believable, and so, to emphasize that it is remarkable" (Israel 2002: 425). This use is illustrated in (16). The third use of *literally* is in cases where it forces a dual reading of an expression. Here, the expression modified by *literally* is a formulary, or idiomatic, expression which conventionally has a non-literal reading. In this case, *literally* is used to signal to the interlocutor that an expression that is conventionally interpreted figuratively should in this case be interpreted literally (Goatly 1997: 174). The sentence in (17) provides a common example. These readings are described as dual because the function of *literally* here is both to force a literal reading of an otherwise idiomatic or figurative expression, and to express that it is precisely because of that that the expression is particularly apt. The fourth use of *literally* identified by Powell (1992) seems to be in a way related to the third use, but is crucially distinguished by an element of creativity and semantic innovation. In this use, "*literally* serves both truth conditional and aesthetic functions" (Powell 1992: 344). According to Israel (2002: 425), *literally* is used "to draw attention to an apt or clever choice of words" or to emphasise "the peculiar suitability of a given choice of words for the described situation". This seems to be particularly the case for the creative uses of *literally*, as exemplified in (18). In this example, the speaker uses *literally* to modify the expression *feed the frenzy* to suggest creatively that the supermarkets support people's diet and nutrition concerns by providing suitable products. While, of course, the writer is not suggesting that frenzy is being literally fed, she uses this formulary expression in order to make an aesthetic or a creative point. The reader is thus forced to think of this expression in a new way (hence Powell's term "semantic innovator"). Finally, the fifth use of *literally* in Powell's categorisation is the use of *literally* with non-literal expressions which cannot be literally true. When used with metaphorical expressions, "the function of *literally* is to encode the speaker's aesthetic judgment that the message, as expressed, is not merely warranted by its capacity to satisfy conditions of applicability but is especially tellable" (Powell 1992: 345). When used with hyperbolic expressions, the function of *literally* is "to tell the reader that the hyperbolical mode itself is justified by the conditions it applies to and

that the lexeme that encodes it is a particularly apt one" (1992: 346). In other words, when *literally* is used with hyperbolic or metaphorical expressions, it signals that "the conventionalized non-literal meaning [of a figurative expression] ... is being used in a strict sense" (Brugman 1984: 34, cited in Israel 2002: 429). This use is exemplified by (19), where the word *exploded* is used figuratively. (Powell 1992: 346) observes that *literally* does not seem to be used with original or unusual metaphors, because in such cases it would be redundant.

Beyond more casual observations on the frequency of variant uses of *literally*, actual data on patterns of use are rather limited. McCarthy and Carter (2004) found that 91% of all occurrences of *literally* in a corpus of British English conversations are used with hyperbolic expressions. Contrary to this high rate of hyperbolic *literally*, data from the British National Corpus show that the use of *literally* with hyperbolic expressions is fairly limited, with fewer than 10% of the total number of occurrences of *literally* being used as a modifier of hyperbole (Claridge 2010: 109). These are cases in which the expressions modified by *literally* cannot be factually true. In these cases, Claridge observes, "[i]t is almost as if the hearer was invited by the use of *literally* to imagine the scene visually, creating a graphic and/or humorous, slapstick-like effect" (Claridge 2010: 110).

This process of change in the use of *literally* has been noticed and commented on by a relatively high number of language commentators and usage guide writers (Israel 2002; Nerlich and Domínguez 2003; Calhoun 2015). The word has been discussed in a series of blog posts and online articles,[3] it has been used for comic effect in popular sitcoms,[4] and its inclusion in the *Oxford English Dictionary*[5] was reported in newspapers. The word's status as a usage problem is further strengthened if we look at usage guides; my initial analysis showed that this was a salient usage problem at least as early as the first edition of Strunk's *The Elements of Style* (1918) (Kostadinova 2015), although additional analysis showed that Bierce (1909) is a usage guide which discusses the use of non-literal *literally* even earlier (cf. Section 5.2.3). In other words, the process of change that *literally* is undergoing can be described as a change above the level of consciousness. As such, *literally* lends itself to an investigation of the relationship between prescriptive accounts of language, and actual processes of language variation and change.

[3]See, for example, Jesse Scheidlower's piece for *Slate*, available online at https://slate.com/human-interest/2005/11/the-trouble-with-literally.html.

[4]Most notably in the animated television series *Archer*, created by Adam Reed.

[5]See entry on *literally* in *OED Online*, available at www.oed.com.

3.6 Negative concord

Negative concord is the linguistic term for what usage guide writers refer to as "the double negative", the phenomenon of using two (or more) negative forms in one clause to express one underlying negative meaning. In linguistic literature, the term "double negative" is reserved for cases where two negative forms are used to express two underlying negatives (cf. Seright 1966; Baker 1970: 171; van der Wouden 1997: 182). These latter cases are sometimes covered under the term "litotes" in usage advice literature. Here, I am exclusively concerned with the former kind of construction, i.e. negative concord. While negative concord, as defined by Labov (1972a), can refer to cases with two or more negative forms in the same clause, my investigation will be devoted only to cases with two negative forms, as illustrated in (20)–(23).

(20) I've got a snug estate, and **don't owe nobody** anything (1811, fiction, COHA)

(21) I don't know very much. **Nobody never** learned me. (1870, fiction, COHA)

(22) I **don't know no one** else that reads so good. (1962, fiction, COHA)

(23) I **didn't do nothing**. (2007, magazine, COCA)

Similar to *ain't*, negative concord is a salient non-standard feature, which, Walker (2005: 1–2) argues, may account for the fact that it is almost always dealt with in studies of the systems of negation in non-standard varieties of English, such as African American English. According to Wolfram and Schilling (2016: 162), both negative concord and *ain't* are among "[a] number of socially marked language variants in American English [which] transcend local communities of speakers". This is also confirmed in work on vernacular universals, which often includes negative concord (e.g. Nevalainen 2006b). According to other similar descriptions, negative concord is a "supraregional", "transnational" diagnostic feature of "substandard culture" in the English-speaking world (Gramley and Pätzold 1992: 309, 377, cited in Howe 1997: 271). Aside from it being a characteristic non-standard feature, "negative concord with indefinites has a long history in the English language", and although it has been found to be more common in historical data, negative concord constructions still occur in English (Howe 1997: 271–272), although not in standard varieties. According to Ukaji (1999: 285), "the copying of Neg into a subordinate clause seems to have fallen into disuse in Standard English in the first half of the seventeenth century",

although this appears to have been a process that started in the fourteenth century (Iyeiri 2001, cited in Nevalainen 2006b: 259). The disappearance of negative concord from the standard variety during the seventeenth century has also been confirmed by Tieken-Boon van Ostade (1982, 2008a), who shows that by the time normative grammarians started formulating the rule on negative concord, this variant had already disappeared from standard use (see also Nevalainen 2006b: 264). However, the codifiers focused emphatically on formulating this rule partly as a result of the idea that English usage should be logical, and its rules should be formulated on a rational basis (cf. Tieken-Boon van Ostade 1982), and partly as a reaction to the use of negative concord by the lower classes, which Tieken-Boon van Ostade (2008a) argues must have been fairly frequent during the eighteenth century. Negative concord has since remained a feature of many varieties of non-standard English.

The social indexicality of negative concord seems to have developed early in the history of this variant. Nevalainen (2006b) shows that the disappearance of negative concord in the period between the sixteenth and the eighteenth centuries was led predominantly by upper-class male and socially mobile speakers, while at the same time being mostly resisted by lower-class male speakers. Sociolinguistic studies of variation in non-standard varieties of twentieth-century English confirm the high frequency of negative concord, both in non-standard British English (Cheshire 1982; Anderwald 2002, 2005) and in American English (Wolfram 1969; Labov et al. 1968; Wolfram 1974; Wolfram and Christian 1976; Feagin 1979; Howe 1997; Howe and Walker 2000), although the rates of use of this feature have been found to vary. Thus, Labov et al. (1968: 267) show that the extent to which the negative concord rule applies may vary considerably among working-class speakers. On the other hand, the rule appears to be categorical for African American pre-adolescents and teenagers in New York (Labov et al. 1968: 276) and in Detroit (Wolfram 1969: 157). In terms of other non-standard varieties of English, Howe (1997: 272) provides a useful overview of findings from sociolinguistic studies reporting figures on the frequency of use of negative concord: Feagin (1979: 232) found 75% cases of negative concord among urban working-class whites in Alabama; Labov et al. (1968: 277) found negative concord at a rate of 81% in their interviews with white youth gangs from New York, and, in her study of working-class Reading English, Cheshire (1982: 65) found that negative concord was by speakers in frequencies ranging from 51% to 89%. Smith (2001) studied negative concord in the context of the relationship between non-standard British varieties and transplanted non-standard American varieties.

This brief overview confirms that while this variant continues to be socially

stigmatised, it is still often found in non-standard varieties of English, and more specifically relevant to the present study, in non-standard American English. In this respect, its inclusion here is meant to investigate further whether the attitudes to this variant in usage guides have changed in the course of the twentieth century, and whether these attitudes are somehow reflected in patterns of actual language use of the variant, as well as what kinds of attitudes can be found towards this variant among native speakers of English in the United States (see Section 4.4 for an explanation of how the analysis of this variant was operationalised).

3.7 Pronouns in coordinated phrases

The use of pronouns in coordinated phrases, also referred to as "compound subjects and objects" (Boyland 2001) or "conjoined NPs" (Denison 1998), is another area where usage varies, and has been a frequent topic of prescriptive criticism. Pronoun case more generally is frequently discussed in the usage guide literature, but it is important to distinguish the forms I am concerned with here from other points of usage which are also concerned with pronoun case, but are of a somewhat different nature. Pronoun case is treated in usage guides in three different contexts. The first involves instances in which a pronoun follows a linking verb, where the problematic distinction is between *This is he* and *This is him*. The second context in which the variation in pronoun case is addressed is the case of pronouns after conjunctions, in constructions such as *He is older than she vs. He is older than her*. Finally, the third case in which pronoun forms can vary is coordinated phrases, such as those exemplified in (24)–(27) below. In the present study, I focus only on these cases.

(24) **Schultz and I** hiked in deeper and made a small spike camp. (1994, magazine, COCA)

(25) You know, **Bernie and me** used to talk, and he'd say "Hey Jerry, I know you feel the same way". (2001, spoken, COCA)

(26) Mr. Pena charged **a friend and me** $150 for the entire day. (2008, newspaper, COCA)

(27) "What in my consciousness attracted that interaction?!" she asked **Jack and I**, who were waiting at my mother's apartment. (2007, fiction, COHA)

The use of *me* in coordinated subjects is considered incorrect, and is often

associated with children's speech. The use of *I* in coordinated objects, however, is considered an affectation, or hypercorrection, which displays a kind of hubris on the part of the speaker: in trying to sound too correct, speakers who say *between you and I* instead of *between you and me* end up committing an 'error'. Such 'errors' in pronoun case seem to be more likely to occur when pronouns are used in coordinated phrases. As Pinker (2014: 97) puts it, speakers usually "effortlessly choose the right case whenever a pronoun is found in its usual place" in the structure of the sentence, "next to the governing verb or preposition". However, "when the pronoun is buried inside a coordination phrase, writers are apt to lose sight of the governor and give the pronoun a different case" (Pinker 2014: 97). Pinker (2014: 206) explains these kind of examples by drawing on arguments related to the nature of a coordinated phrase, which is headless, and, consequently, he argues, the "harmony" between its parts may not be part of speakers' "intuitive grammars". In other words, as soon as the pronoun is separated from the governing preposition or verb by the coordinator *and*, speakers' intuitions about pronoun case may not be as strong. On the basis of historical data on pronominal usage, Tieken-Boon van Ostade (1994: 226) has hypothesised a process of natural language change in pronominal case that "may perhaps be described as a continuing loss among speakers of English of their sense of case distinction". This kind of variation in the use of English pronouns, which results in the use of pronominal subject forms where one would expect object forms, is attested in early Modern English. Hock and Joseph (2009: 192) cite instances of "an innovated system with different case marking conventions", such as these examples from Shakespeare's plays: *You know my father hath no child but I*, *Let fortune go to hell for it, not I*, and *…all debts are cleared between you and I*, where the rule about pronoun case "is getting relaxed, requiring objective marking only on pronouns that are directly preceded by the verb or preposition". This means, that "in vernacular or untutored Modern English, uninfluenced by the rules of prescriptivists … adjacency plays a role in case marking" (Hock and Joseph 2009: 192). Hock and Joseph (2009) seem to be suggesting that as soon as prescriptivism exerts some influence on the use of pronouns in coordinated subjects and objects, this influence may interfere with the more unconscious influence of adjacency in case marking. The alleged influence of prescriptivism is believed to have resulted in creating linguistic insecurity in speakers, who, as a result, may have tended to hypercorrect pronoun usage in cases such as *between you and I*. The extent of the influence of prescriptivism on the actual use of these features, however, remains unclear. Despite this influence, however, Denison (1998: 109) argues that variation in pronominal usage in coordinated phrases remains

common among educated speakers of English, despite being considered non-standard.

In a study that provides further insights into the cognitive processes that account for pronoun usage in English with a focus on "compound subjects and objects that incorporate a first person pronoun", Boyland (2001: 383) shows that hypercorrection of this sort cannot always be accounted for by sociolinguistic prestige factors alone. In a survey of the attitudes of speakers towards object *x and I*, as well as a small-scale corpus-based study of the occurrence of this variant in online language, Boyland (2001) found that cases of hypercorrections can often be accounted for by priming due to frequency effects, rather than being the result of a conscious attempt on the part of the speaker to sound more prestigious. This is an important contribution to the discussion of hypercorrect usages such as *between you and I*, because it shows that the use of this construction may not always be the result of pressures from "above" the level of consciousness, but may well be "below" the level of consciousness, and under the influence of priming. This conditioning in the use of *between you and I* may therefore not always be a conscious attempt to sound like speakers from a perceived higher social class, but a tendency to be influenced by the linguistic input from peers: "frequent exposure to a construction changes adults' intuitive judgements" (Boyland 2001: 390). In her corpus linguistic analysis, based on online forum data, Boyland analysed 227 instances of the occurrence of *x and I*, and found that only 9 instances, or less than 4%, are cases where *x and I* occurs in an object environment.

In the context of the present study, I will analyse the treatment of pronouns in coordinated phrases in American usage guides, and analyse the patterns of actual use of these variants, in order to investigate whether prescriptivism has had any measurable influence on their use.

3.8 The split infinitive

The split infinitive is a usage problem concerning the placement of an adverb which modifies a full infinitive, as exemplified in (28)–(30) below, where the first example illusrates what is perceived as an unacceptable usage, and the other two examples show the acceptable variants. The split infinitive is an 'old chestnut' among usage problems (Tieken-Boon van Ostade 2015: 57). The origins of its proscription are somewhat controversial, and go back to the nineteenth century. The origin of the split infinitive itself can be traced back to the appearance of the *to*-infinitive in Anglo-Saxon English (Bryant 1946). This allowed for the particle *to* to be separated

from the verb, and as a result of the general tendency for adverbial modifiers to immediately precede the word they modify, separating the *to* from the verb started to become common (Bryant 1946). The usage problem most likely originated in the early nineteenth century; the earliest record of a complaint against split infinitives comes from a 1834 letter to the editor, citing the Teutonic origin of English as an argument against split infinitives (Perales-Escudero 2011: 318).

(28) ...warns of "the chaos that could only result if members of the armed services were **to freely talk** on pending policy and security questions to the representatives of the press." (1968, fiction, COHA)

(29) She was now permitted **freely to study** the face. (1920, fiction, COHA)

(30) Mrs. Petrovic encourages parents to spend as much time as possible with children, and **to talk freely** about customs and situations here and back home. (2000, newspaper, COCA)

This usage problem, however, seems to have lost popularity among sticklers and in general discussions on usage correctness. As early as 1927, usage commentators pointed to the superfluity and ridiculousness of the rule, especially in the context of its accordance with the natural flow of the English sentence structure (Curme 1927; Bryant 1946). In certain cases, the position of the adverb in an infinitive phrase is not merely a matter of usage or of a style norm, but also interferes with its meaning. In such cases, the split infinitive has a meaning that is different from the meaning a similar non-split infinitive would convey, as exemplified by the opposition between *He failed completely to understand it* and *He failed to completely understand it* (Curme 1927: 341). Curme (1927) further notes that the reason why the use of the split infinitive construction has grown despite the opposition against its use is because of its "intrinsic merit" (Curme 1927: 342). Curme (1927) relates the development of the split infinitive to two different ways in which adverbs can be used to modify verbs in English. When the verb is preceded by an adverb, the stress is on the verbal activity, while when we want to emphasise some other aspect about the verbal activity, rather than the activity itself, we position the adverb after the verb. Curme (1927: 341) provides the following examples to illustrate this point: in the first case we have *He almost succéeded* while in the second case we have *When he acts he acts prómptly*, where the accents denote stress. Curme (1927: 342) points out that "this twofold position of the adverb with differentiated meaning is a marked feature of English

expression and in the last centuries has gradually been becoming more fixed" and "has furthered the development of the split infinitive". Somewhat later, Malone (1941: 52) observes that the split infinitive is "a mere matter of word-order [which could] be found as early as the fourteenth century, but did not become frequent until about 100 years ago, and even now is rare in popular speech, its use being chiefly literary [and] its professional standing has grown better".

Fischer (2000) provides important insights into the historical development of the split infinitive in the context of the reversal of the grammaticalisation process of infinitival *to*. She notes that the increased number of the split infinitives since the fourteenth century were indicative of a "disturbed" process of grammaticalisation. After the fourteenth century, with the construction becoming more frequent, it became characteristic of the personal style of certain authors (Calle-Martín and Miranda-García 2009: 347). They further note that the use of the split infinitive decreased during the sixteenth and seventeenth centuries, only to pick up again during the eighteenth century. They consider this a significant case of the variant resisting prescriptive pressure against its use. Even though they found a large majority of non-split infinitives, but they also observed "a significant decrease of non-splitting constructions" in the nineteenth and twentieth centuries, "thus coinciding with the actual spread of the construction" (2009: 351). These observations are also reflected in the work on infinitival *to* and the split infinitive of Fitzmaurice (2000a,b), who shows that negative split infinitives are becoming more frequent in present-day English. With regard to these infinitives, Fitzmaurice argues that the use of the negative split infinitive is a case of stylistic levelling, where, due to the mixing of different varieties in spoken media registers, certain features change their status from colloquial to conventional, and become unmarked (Fitzmaurice 2000a; see also Kato 2001). In addition, she considers the increase in the use of negative split infinitives to be a case of "de-grammaticalisation" of infinitival *to* (Fitzmaurice 2000b).

Additional studies looking at the split infinitive confirm to some extent the findings by both Fitzmaurice (2000a) and Calle-Martín and Miranda-García (2009). Leech et al. (2009: 263) found that the frequency of occurrence of split infinitives has increased in both British and American English corpus data, on the basis of the BROWN family of corpora, while Davies (2010a) found that the proportion of infinitives split with a *-ly* adverb has increased in the Corpus of Contemporary American English relative to infinitives immediately followed by an *-ly* adverb. Similar trends are observed in Fischer (2007), whose analysis is also based on the BROWN family of corpora. Building on the previous research, I will attempt to take

the study of the influence of prescriptivism on the use of the split infinitive one step further, and analyse whether and how the treatment of the feature has changed in usage guides, and whether and how this relates to its increased frequency of occurrence in actual use data.

3.9 Conclusion

In this chapter I introduced the HUGE database compiled at Leiden University, which contains a collection of usage guides, as well as entries on 123 usage problems. On the basis of the most commonly treated usage problems, as well as a number of other criteria, I explained how I selected the usage features for this study. Subsequently, I provided a brief background for each of these features, summarising the most important aspects of the variation in their use. These aspects are important because they may reasonably be expected to influence the treatment of these features in usage guides, as well as their patterns of use and speakers' attitudes towards them. Now that I have established the points of departure for the present study, as well as the background information on each of the features investigated, in the next chapter I will describe the details of the three-pronged methodological approach.

CHAPTER 4

Methodology

4.1 Introduction

As outlined at the end of Chapter 1, the main purpose of this study is to empirically investigate present-day prescriptivism in American English by examining a small set of language features which are generally known to be usage problems for ordinary speakers. This investigation will be carried out by approaching these features from three perspectives: (a) attitudes to these features found in American usage guides, (b) the actual patterns of variation and change in the use of these features in corpus data, and (c) ordinary speakers' attitudes towards the use of these features. In this chapter I outline the approach taken in this study, first by discussing my general approach to the study of prescriptivism in Section 4.2, and then by presenting the types of data and analysis used for the present study in Sections 4.3, 4.4, and 4.5. The study is not limited to one particular language feature; rather, it focuses on a small set of features, in order to explore in detail the extent to which different features are affected differently by prescriptivism. The selection of the features was discussed in Section 3.2. The different perspectives I will adopt in analysing the features are explored by analysing three types of data. The metalinguistic treatment of the features in usage guides will be studied on the basis of an analysis of entries on the selected usage

features found in American usage guides. The patterns of variation and change of the selected language features will be analysed using American English corpus data, and a range of methods, including a quantitative multivariate analysis of frequency patterns and constraints on the distribution of different variants in the context of one of the features described in Chapter 3, the split infinitive. Finally, speakers' attitudes will be explored with data from a survey and post-survey interviews conducted with native speakers of American English.

4.2 General approach

The overall approach to investigating present-day language prescriptivism in American English taken in this study, as mentioned in the introductory chapter above, is three-pronged; it aims to explore the three main perspectives and to account for the potential influence of prescriptivism on language use. This tripartite division is reflected in the methodology and the data used. Generally, the study applies both qualitative and quantitative methods to the analysis of the three types of data mentioned above. There are a number of ways in which the present methodological approach alleviates some of the problems and difficulties in studying prescriptivism established in previous research.

The first aspect of the approach is that the study is significantly informed by the comparison of precept and practice (Konopka 1996, cited in Auer 2009; Gustafsson 2002; Auer and González-Díaz 2005), a methodological approach often taken in studies of the influence of prescriptivism and normative linguistics on language change. As discussed in the theoretical background presented in Chapter 2 above, prescriptivism is sometimes assumed to have an influence on language, but in many of the cases where such an influence is assumed, prescriptivism is taken to be a static normative phenomenon, which is often approached as a phenomenon which does not change over time. However, numerous discussions of prescriptive ideology have shown that prescriptivism does in fact change over time, so in order for prescriptive influence to be ascertained reliably, an analysis of prescriptive ideology and institutionalised prescriptive attitudes needs to be carried out (cf. Curzan 2014).

The second aspect of the approach taken here is that I focus on multiple language features, rather than on a single feature (e.g. Auer 2006; Wild 2010; Yáñez-Bouza 2015) or a set of related features (e.g. Anderwald 2012, 2016). In terms of language features, I focus on a small set of linguistically unrelated features, described in

Chapter 3 (for the approach to analysing their variation in the context of their respective variables, see Section 4.4), but I also account for their relationship to other prescriptively targeted features (see further Section 4.4). The aim here is to find a middle ground between the detailed, exhaustive approach taken in previous studies focusing on individual language features, and a more overarching approach that attempts to account for the way in which different language features are affected by prescriptivism.

Thirdly, the study covers the period from the middle of the nineteenth century to the present day, with special emphasis on present-day American English. In other words, I look at prescriptivism during the prescription stage of the process of standardisation, a little-studied stage in the context of studies of prescriptivism and its effects. Previous research on the topic is biased towards the stage of codification and has tended to focus on the eighteenth (e.g. Auer and González-Díaz 2005; Yáñez-Bouza 2015) and nineteenth (e.g. Dekeyser 1975; Anderwald 2014) centuries. The data on prescriptive attitudes to usage, as well as on language variation and change, go further back in time, to the middle of the nineteenth century, in order to track potential changes in prescriptive attitudes. Furthermore, Auer (2009: 9) has pointed out the need for language corpora to cover longer periods of time than the precept corpora (i.e. the collection of metalinguistic, or precept, data). In this way, any identified changes in precept may be tested for their influence by looking at the language use before and after the period in which changes in precept might have taken place, allowing for a time gap for the potential influence to be reflected in language use data.

Lastly, I include data on speakers' attitudes; in the discussion of the relationship between speakers' attitudes and prescriptivism the focus will be on present-day American English. As already mentioned in Chapter 1, speakers' attitudes are crucial to a better understanding of the social influence of prescriptive ideology. In historical sociolinguistic studies, such data are fairly difficult, if not impossible, to obtain. On the other hand, in surveys of attitudes to usage conducted in the course of the twentieth century (such as Leonard 1932; Marckwardt and Walcott 1938; Crisp 1971; Mittins et al. 1970), the attitudes to usage investigated are those of language professionals, not ordinary language speakers. In dealing with the attitudes of ordinary language speakers, this study aims to contribute to our understanding of prescriptive influence and attitudes to usage. The inclusion of an analysis of speakers' attitudes has been one of the critical aspects of the broader research project which this study is a part of, as mentioned in Chapter 1.

The precept *vs.* practice approach is, of course, not without problems. While the analysis of precept has been shown to be indispensable to the study of prescriptive influence, the relationship between changes in precept and changes in practice is problematic. As discussed in Section 2.3, a few studies based on the precept *vs.* practice approach have found limited influence of prescriptivism on language variation. Such studies, however, come with an important caveat, namely that correlation is not necessarily causation (Hinrichs et al. 2015), and, because of this, Curzan (2014: 84) observes that conclusions based on relationships between prescriptivist judgements and corpus evidence should be drawn very carefully. In order to deal with this limitation, the methods used in this study will go beyond the comparison of precept and practice patterns, and will operationalise the analysis of prescriptive influence by looking at the co-occurrence patterns of a number of prescriptively targeted features, alongside the six features which are at the centre of this study. This approach is based on Hinrichs et al. (2015), who investigated the extent to which the use of *that* as opposed to *which* correlates with the use of other prescriptively targeted features such as split infinitives, passives, sentence-final prepositions, and future reference *shall* with first person subjects, on the basis of data from the BROWN family of corpora. This kind of approach allows for a more thorough investigation of the possible prescriptive influence on the use of specific variants. These additional features are discussed in detail in Section 4.4.

4.3 Usage guides: data and analysis

The analysis of attitudes to usage in American usage guides is based on an analysis of 70 guides published in America between 1847 and 2014 (see Primary Sources). Entries on the language features investigated here were collected and analysed across a number of dimensions, discussed below. The selection of the usage guides used in the analysis was carried out in large part on the basis of the HUGE database, described in Section 3.2. A search for American usage guides yielded 44 guides, of which eight are classified as both British and American in HUGE. The reason for the double classification of these usage guides is that they cover linguistic features used in both British and American English. In the HUGE database user manual, Straaijer (2015: 5) notes that "[i]n cases in which the language variety was not explicitly mentioned, the variety was usually assigned based on the country of publication and the nationality of the author". However, a closer look at the doubly classified usage guides revealed

that while some of them do address both British and American standards of usage (a good example of this is Peters 2004), others point out in their prefaces that they are mainly concerned with British usage. Such is the case for Greenbaum and Whitcut (1988), for instance, where the authors explicitly point out in their preface that they "address [themselves] primarily to those wanting advice on standard British English" (1988: xiii). Four of these guides were thus excluded from the present analysis, resulting in the use of a subset of 40 American usage guides from the HUGE database. Given that the database contains both British and American usage guides, and is less comprehensive for the twentieth century than for the nineteenth (cf. Straaijer 2015), additional research was carried out to take into account more usage guides written for American English and published in the twentieth century. This additional selection was based on the criteria used in the selection of material for the HUGE database, in order to ensure consistency in the collection of guides, *viz.* selecting those guides that treat predominantly grammatical usage problems. A further criterion was the selection of only American usage guides. This additional search process consisted of searching the digital libraries HathiTrust and Internet Archive, Google Books, and the Leiden University Library Catalogue for additional titles. The additional search produced 30 usage guides, which were added to the 40 available in HUGE, thus bringing the total number of American usage guides consulted for the purposes of the present analysis to 70.[1]

The definition of what a usage guide is, as well as the delimitation of the genre of usage guides, is not a straightforward task, because of the variation in the form and content of these works (Straaijer 2018: 12–13), an issue I raised in Section 2.2. As I explained there, the question of the genre differences in these kinds of metalinguistic works has been raised elsewhere (e.g. Connors 1983; Weiner 1988; Straaijer 2018), but it has not always been consistently applied to analyses of usage guides. It is important to point out that very few of the studies of usage guides go into much detail on the selection of materials and the definition of usage guides, or other metalinguistic reference works, such as language manuals, handbooks, and textbooks. Meyers (1995) for instance focuses on handbooks of composition, but uses the terms textbooks and handbooks interchangeably to refer to the 60 books he analyses. No further details are included as to the selection of materials. A similar gap relating to

[1] In some cases, usage guides are published in both the United Kingdom and the United States. Such is the case, for instance, for Brians (2003). In the case of Partridge (1947), the first edition of his usage guide was annotated for American English, and these notes were given in square brackets in the first British edition.

the data selection process is present in Albakry (2007), who used 18 usage books to analyse the treatment of a number of features; the criteria for selecting these 18 books, however, are not explicitly addressed. Peters and Young (1997) similarly provide no explicit criteria for determining what a usage guide is. The compilation of the HUGE database dealt specifically with the question of delimiting the genre of usage guides, and establishing the criteria for what counts as a usage guide or not (cf. Straaijer 2018); these criteria were also applied in the present study.

It is also important to discuss the extent to which these 70 usage guides are representative of general usage guide publication trends in American English. While it is impossible to come up with exact figures for the total number of guides published in the United States, it is possible to arrive at an approximate picture of the publication trends. The reasons behind the difficulty in ascertaining exact figures are that (a) it is impossible to know for certain how many usage guide titles were published in total, and (b) an attempt to ascertain this total number of usage guides would also depend largely on one's definition of usage guide, which, as I have argued above, is difficult to establish (cf. Straaijer 2018). Thus, Straaijer (2018) notes that the estimated number of usage guides ever published may be between 250 and 300 titles, depending on one's definition (see also Tieken-Boon van Ostade forthcoming). Based on the definition and genre characteristics discussed in Straaijer (2018), and a number of additional sources consulted (e.g. the bibliographies of Gilman 1989 and Garner 1998), a list of the total number of usage guides published in America was produced. The publication trends are given in Figure 4.1. An important point to make is that the considerable variation in the genre means that these figures should not be taken to represent absolute numbers of published guides per decade, but should rather be interpreted as an approximation to the real situation. There may be guides that were published, but not identified in the process. In addition, if different selection criteria were applied, the results might be slightly different.

Figure 4.1 shows the general increase in usage guide publications during the twentieth century. While I noted that it is difficult to provide exact figures, similar trends have been observed elsewhere (Straaijer 2018; Tieken-Boon van Ostade forthcoming). In addition, earlier work on handbooks of composition, a genre which can be considered related to usage guides, has shown that these types of books were also on the increase. Meyers (1995: 30), for instance, analysed "60 handbooks of composition published between 1980 and 1993 (33 between 1990 and 1993 alone)". What is interesting here is that the number of handbooks is strikingly higher in the three-year period between 1990 and 1993. Even though Meyers does not provide

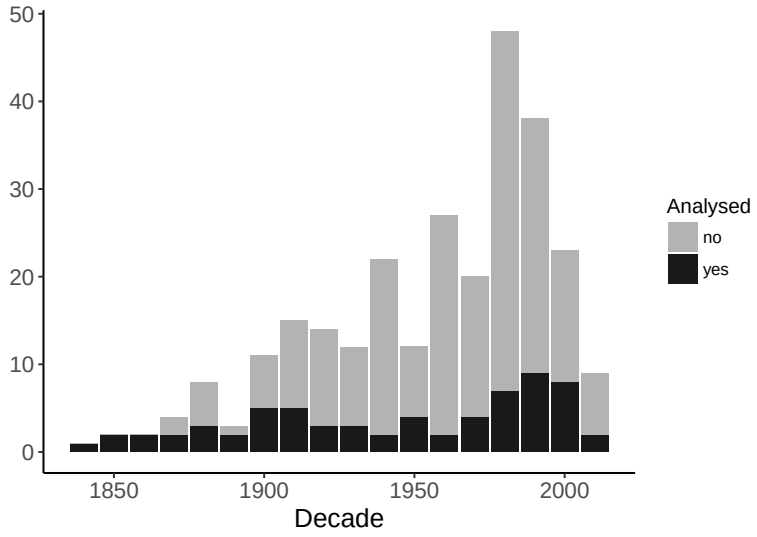

Figure 4.1: Total number of guides published in America per decade and the number of guides included in the analysis

details on the process of sampling of the analysed handbooks, or the extent to which his findings are representative of general publication trends, given the relatively high number of handbooks analysed, it might be safely concluded that the number of handbooks published grew significantly during the last decade of the twentieth century.

Figure 4.1 also shows that the number of usage guides analysed is not the same across all decades, which may be considered a limitation of the dataset. However, given the difficulty in obtaining many of these works, the inclusion of at least two usage guides per decade was considered to be adequate. When it comes to selecting a representative group of usage guides, the decision would not only depend on obtaining a representative number of guides, but also on selecting influential and popular guides. This aspect presents a different set of challenges, because determining the influence or popularity of a guide is not always straightforward. Despite these potential limitations, which will be kept in mind in the interpretation of the findings, the present collection of usage guides is significantly larger than those used in previous studies. The three most comprehensive studies of usage guides to date, for instance, analysed fewer usage guides: Creswell (1975) analysed ten usage guides, Peters and Young (1997) fourteen American usage guides, and Albakry (2007: 33) “a sample of the eighteen most popular usage books published in the United States since 1950”. In drawing on

a larger collection of usage guides, it is hoped that a clearer picture will emerge of the attitudes to usage and changes in those attitudes over time.

In most instances, the first edition of the usage guide in question was consulted, and the year of the first edition was used in the analysis of the usage guide data, as will be explained in more detail below. This also applies to subsequent impressions of the first edition. For fifteen out of the seventy selected usage guides, I was not able to obtain the first edition, nor was I able to ascertain the extent to which the editions that I consulted were changed compared to the first edition. This is a problem that has cropped up before in studies of normative grammars in historical sociolinguistics (see Wild 2010: 31, footnote 11). Yáñez-Bouza (2015: 29), for instance, argues for considering multiple editions of the texts in questions "on the assumption that different printings are likely to show modifications in the discussion of the same topic". On the other hand, with reference to eighteenth-century grammars, and specifically Lowth's grammar, Tieken-Boon van Ostade (2008b: 122) notes that new editions were sometimes advertised as corrected as a publisher's ploy to increase sales.

Although this discussion concerns eighteenth-century grammars, a similar case can be made for usage guides, which may be published in a 'revised' edition, but for which it may be rather difficult to ascertain to what extent the edition has changed, given that in most cases multiple editions of the same work were not available. For these usage guides, I decided to take the year of the edition consulted, as looking at the actual years of the revised editions showed that this may not be a critical issue in the analysis of the data. Since the data were analysed across decades, usage guides for which the edition consulted here was published in the same decade as the first edition were associated with the same time period in the analyses. This is the case with Witherspoon's *Common Errors in English and How to Avoid Them*, which was first published in 1943, while the edition I consulted was published in 1948; there were three such usage guides in total. Of the remaining twelve, most of the revised editions consulted were published within ten years of the first edition, so the difference in dating them was only one decade. In that case, the year of the revised edition was used, on the assumption that the time difference would not critically affect the analysis of the results; this was the case with eight of the usage guides. Finally, of the remaining four, the difference between the first and the consulted revised editions was two decades for three usage guides, and four decades for only one usage guide, Ebbitt and Ebbitt's *Writer's Guide and Index to English* (1978), first published in 1939.

As mentioned above, entries on the six selected usage features were identified and used in the analysis of attitudes to usage (which, as will be discussed below, are not to

be confused with speakers' attitudes). The HUGE database allows users to download entries on specific features in various formats, including xml. I used this file format to create a corpus of entries in which the text for each entry is a separate text file. This file is linked to another file which contains the metadata for each entry, such as author, title, year, and page numbers. Finally, for each entry, there was a third type of file where annotations were stored. For the usage guides which were selected additionally, the text, metadata, and annotation files for each entry were created manually. This resulted in a corpus of 281 entries in total for the six language features, given in Table 4.1.

Feature	**No. of entries**	**Average no. words per entry**
ain't	46	243.69
like	10	180.82
literally	32	173.09
negative concord	42	294.71
pronouns: object *I*	73	210.90
pronouns: subject *me*	19	347.42
split infinitive	59	381.58
total	281	268.57

Table 4.1: Number of entries across linguistic features in the corpus

On the basis of previous studies of normative grammars and usage guides outlined in Sections 2.2 and 2.3 above, a framework for analysis was created in an attempt to provide a comprehensive account of the treatment of the linguistic features under investigation and the attitudes expressed towards them. Previous studies have shown that a prescriptive approach to usage can be manifested in different ways in usage guides, including the manner in which opinions are expressed (i.e. whether they are reported as opinions held by others, or as the author's own opinions; cf. Busse and Schröder 2009), the nature of those opinions (i.e. whether they approve of problematic variants or not; cf. Albakry 2007), and the approach to using sources (i.e. whether pronouncements are based on sources or not; cf. Peters and Young 1997). Consequently, what these studies have also illustrated is that there is no single indicator of prescriptivism (cf. Peters and Young 1997), and no single way in which usage guides discuss usage, present opinions or facts, and offer advice. Building on these insights, the aspects I investigated in the usage guide entries are the following: the way usage guide authors treat the usage features, the attitudes expressed towards these features, and the dimensions of usage invoked in the treatment.

The analysis of the treatment of the six usage features in the collection of American usage guides was done by distinguishing three types of treatment: ACCEPTABLE, RESTRICTED, and UNACCEPTABLE.[2] This analysis of treatment is similar to the approach taken in previous studies of normative grammars and usage guides. Dekeyser (1975), for instance, establishes a number of methodological approaches in dealing with this issue empirically, which are reflected in some later studies. A case in point is his analysis of prescriptions found in normative grammars, using the following categories: $+$, if a grammarian supports a prescription, $-$, if a grammarian rejects a prescription to accept a problematic construction, or $\pm$, if the grammarian's opinion is between these two positions. A similar approach is used by Yáñez-Bouza (2015) for the analysis of normative grammar pronouncements on preposition stranding. This tripartite categorisation has also been used in a number of studies of usage guides.

The first extensive study to include such a classification is Creswell's (1975) analysis of the pronouncements found in the usage notes of the *American Heritage Dictionary* (1971). Motivated by the comparison of labelling practices in dictionaries by McDavid (1973), Creswell (1975) set out to investigate the extent to which dictionaries and usage guides agree in their usage judgements. He took the usage notes of the *American Heritage Dictionary* (1971) as a starting point for his comparison of the judgements on usage across a collection of handbooks of usage, and analysed them according to a number of dimensions. In the context of how various locutions were treated across different works, Creswell (1975) used the categories Not Treated ("either the word is not entered at all or, if it is, the specific problem in usage is not referred to"), Accepted ("either entered without comment, or discussed and approved, in the usage books the latter only"), and Restricted ("either assigned a restrictive label or or discussed and recommended to be completely avoided or limited to use in certain contexts") (Creswell 1975: 8). A similar approach was taken by Berk (1994) for the analysis of 26 books on usage. The distinctions in treatment she used are: "rule invoked", "rule rejected", "rule invoked for formal discourse", "rule may be overridden for rhetorical concerns", and "no discussion of rule" (Berk 1994: 111). In a study looking at the sources of evidence used in language advice literature in Australia, the United Kingdom, and the United States, Peters and Young (1997) used similar categories to those used by Creswell (1975), *viz.*, U for "unacceptable", A for "acceptable", and R for "usable in restricted contexts", where the final category

[2] I use small capitals for terms referring to analytical categories used in the present study. I use regular font when referring to the general notion of, for instance, acceptability.

"was applied whenever the usage writer explained some constraint on the usability of the structure", and was also used to "represent the rather equivocal stance of an author who admits that a certain usage practice is common but advises that 'careful writers do otherwise'" (Peters and Young 1997: 320–321). More recently, Albakry (2007) uses similar categories to analyse the treatment of five usage features in 18 usage and style guides in American English. Based on Peters and Young's categories, Albakry (2007: 34) distinguishes among "not acceptable (i.e. deemed to be incorrect and should be avoided), acceptable (i.e. deemed correct and should not be avoided), vague (i.e. commentator explains some constraints on the use of the structure or espouses an equivocal stance towards it) and not mentioned (i.e. the usage feature is not commented on in the particular usage guide)".

Given the discrepancies between these categories in previous studies, I provide more specific definitions of the three categories of treatment formulated for the purposes of this analysis, although they do not depart greatly from the basic distinctions. A treatment was classified as ACCEPTABLE when the author explicitly approves of the construction, as exemplified in (31) and (32). It is important to point out that cases in which some restrictions are mentioned, but where these restrictions are linguistic rather than social or situational, were also classified as ACCEPTABLE. Examples of this kind of entry are especially common in the treatment of the split infinitive, where the restrictions on the use of the feature have to do with the length of the element that separates the particle *to* from the verb, or with the naturalness or awkwardness of a construction, but not with whether it is socially or situationally appropriate. Such entries were classified as ACCEPTABLE, because they explicitly express acceptability of the feature across registers while sometimes also explicitly dismissing, and even disparaging, the prescriptive rule against it. The treatment was classified as RESTRICTED when the author partly approves of the construction, while noting restrictions on its use which are social or situational. This includes cases where the construction is criticised, but where it is also noted that the item is used in certain contexts, or when various opinions, both accepting and not accepting the feature, are mentioned, as in (33) and (34). Furthermore, entries that neither accept nor dismiss a feature were also put into the category RESTRICTED. This category may also be considered the most loosely defined, because RESTRICTED entries which do not contain any explicit judgement of a feature do not offer much evidence of attitudes, as in examples (35) and (36). Finally, entries in which the author explicitly disapproves of the construction and does not find it acceptable in any context were classified as UNACCEPTABLE; an example is given in (37).

(31) Do not be afraid to split infinitives or other verb forms. (Clark 2010: 255)

(32) I know of no usage authorities who believe that split infinitives are always wrong, but I take a more extreme position than most: More often than not, in my opinion, infinitives are better split. (Walsh 2004: 64)

(33) This word is a contraction of *am not* or *are not*, and can, therefore, be used only with the singular pronouns and *you*, and with the plural pronouns *we*, *you* and *they*, and with nouns in the plural. (Bechtel 1901: 119)

(34) Even though it has been universally condemned as the classic mistake in English, everyone uses it occasionally as part of a joking phrase or to convey down-to-earth quality. But if you always use it instead of the more "proper" contractions you're sure to be branded as uneducated. (Brians 2003: 6)

(35) *Literally* means "actually, without deviating from the facts," but it is so often used to support metaphors that its literal meaning may be reversed. In statements like the following, *literally* means "figuratively" and *literal* means "figurative":

The Village in the twenties [was] a literal hotbed of political, artistic, and sexual radicalism.–Louise Bernikow, New York Times Book Review

In this struggle, women's bodies became a literal battleground.–Martin Duberman, ibid.

[New York City is] literally hanging by its fingernails.–Walter Cronkite, CBS News

Literal-minded readers find such locutions absurd. (Ebbitt and Ebbitt 1978: 547–548)

(36) Writers are so often besought by rhetoricians not to say *literally* when what they mean is *figuratively* that one would expect them to desist in sheer weariness of listening to the injunction. The truth is that writers do not listen; and *literally* continues to be seen as a mere intensive that means *practically*, *almost*, *all but*. *He was literally speechless. He could only murmur: "Good God!"* This speechlessness would be literal only if he had been incapable of uttering the words we are told he murmured. [A golf cart] *literally floats over the roughest fairway*. To accomplish this it would have to be one of those vehicles that ride a few inches above the ground on a cushion of air. Since this particular cart moves with its wheels on the ground, the floating is figurative. (Follett 1966: 204)

(37) This cannot be called a contraction, and however much it may be employed it will still be only vulgarism. *I'm not* is the only possible contraction of *I am not*, and *we're not* of *we are not*. (Ayres 1911: 6)

The second dimension analysed was the attitudes to the language features expressed in the usage guide entries. The general notion of attitudes notoriously defies a straightforward definition, so it is important to distinguish between attitudes to usage in usage guides and speakers' attitudes in the context of this analysis. The issue of the multifarious nature of attitude studies was referred to in Section 2.7 above. At the end of that section, I referred to a group of studies in which the term "attitudes to usage" has been used in relation to the attitudes of grammarians, writers on language, or language authorities (cf. e.g. Finegan 1980; Sundby et al. 1991). In this sense, then, the notion of "attitudes to usage" refers to attitudes expressed in metalinguistic publications, such as grammar books, style guides, language manuals, and usage guides, and can be understood as an instantiation of metalinguistic commentary; as such, this notion of attitudes to usage should be distinguished from speakers' attitudes. While this is indeed an important distinction, the term "attitudes" seems appropriate in the context of my analysis of usage guides, because it captures the subjective and attitudinal component of the pronouncements found in usage guides.

This analysis is based on an approach which has been employed in studies of attitudes to language in normative grammars (cf. e.g. Sundby et al. 1991), and which provides a useful point of departure for the analysis of attitudes in usage guides. Important in this respect are analyses of normative or prescriptivist metalanguage, that is, the various ways in which grammarians expressed their ideas about language. Studies on normative grammar in eighteenth-century English that are important in the context of attitude analysis are mainly those that deal specifically with the labels used by grammarians in the treatment of linguistic variants. Sundby et al.'s *Dictionary of English Normative Grammar* (*DENG*) (1991) marks an important step in providing a fairly comprehensive inventory of the "prescriptive labels" that comprised the metalinguistic system of eighteenth-century grammarians. The classification presented in *DENG* provides the opportunity to use those labels to study attitudes to usage. A good example of this is Yáñez-Bouza's (2015) study of the attitudes of normative grammarians towards preposition stranding in the eighteenth century. By relying on labels used in grammars, she shows how labels found in normative grammars reveal the emergence of strikingly conservative attitudes towards preposition stranding that arose in the middle of the eighteenth century. Building on Sundby et al. (1991), Yáñez-Bouza (2015) provides a comprehensive list of terms that expressed attitudes to preposition stranding in that period, and classifies them according to whether they can be interpreted as advocating or criticising the construction. The analysis of such labels is thus part of the present approach.

The categories used here were POSITIVE and NEGATIVE attitudes, exemplified in (38) and (39) below. These two categories were taken as the most intuitive way of analysing the attitudes expressed towards the features. Attitudes to usage in usage guides are usually expressed through a kind of semi-technical language consisting of a number of metalinguistic expressions such as "vulgarism", "error", "gross linguistic gaffe", "natural", "acceptable", and so forth. More rarely, attitudes can also be expressed by the use of particular verbs, such as "avoid". In all instances of explicitly expressed attitudes, the distinction between POSITIVE and NEGATIVE attitudes is fairly intuitive, and was taken to serve as a basis for the present analysis.

(38) Despite the taint of *ain't* from its origin in regional and lower-class English, and more than a century of vilification by schoolteachers, today the word is going strong. It's not that *ain't* is used as a standard contraction for negated forms of *be*, *have*, and *do*; no writer is that oblivious. But it does have some widely established places. One is in the lyrics of popular songs, where it is a crisp and euphonious substitute for the strident and bisyllabic *isn't*, *hasn't*, and *doesn't*, as in "It Ain't Necessarily So," "Ain't She Sweet," and "It Don't Mean a Thing (If It Ain't Got That Swing)." (Pinker 2014: 204)

(39) *Ain't* is unrecognized in modern English & as used for *isn't* is an uneducated blunder and serves no useful purpose. (Nicholson 1957: 49)

The third aspect of the analysis is related to the dimensions of usage identified by usage guide writers. In his account of writing a usage guide, Weiner (1988) uses the term "sociolinguistic considerations" to refer to extralinguistic factors that usage guide writers draw on in their selection and discussion of usage problems, as well as the information provided about those extralinguistic factors in treatments of usage. It is, of course, important to note that the term "sociolinguistic considerations" does not imply that such observations made in usage guides are to be understood as objective, or as based on descriptive or empirical studies. Nevertheless, these observations are important in showing how certain sociolinguistic aspects of the features are referred to and understood by usage guide writers (cf. Tieken-Boon van Ostade 2015; Kostadinova 2018a). What Weiner (1988) refers to as "sociolinguistic considerations" can be compared to levels of usage or usage dimensions, a question usually discussed in the context of dictionaries. In dictionaries such considerations tend to be expressed through a set of labels, such as "common", "rare", "archaic", "dialectal" (cf. McDavid 1973; Creswell 1975; Card et al. 1984). Sundby et al. (1991: 38) similarly refer to the

practice of eighteenth-century grammarians of referring to a number of dimensions in their treatment of usage, such as "medium ('we never write'), genre ('hardly allowable in poetry'), frequency ('seldom used'), attitude ('rude especially to our betters'), social position ('low'), linguistic competence ('adopted by the ignorant'), territory ('peculiar to Scotland'), etc.".

These kinds of dimensions of usage are also characteristic of usage guides, and were consequently analysed separately. Employing both of these terms in the discussion of the treatment of linguistic features in usage guides presents certain problems. The term "sociolinguistic considerations" may imply a scientific sociolinguistic basis for those considerations, which, in reality, is both hard to prove and very unlikely. Even in cases where usage guide writers have attempted to consider sociolinguistic aspects of the use of particular features, sources are rarely cited (Peters 2006; Peters and Young 1997), which makes it difficult to rely on the unproven presupposition that any kind of reference to sociolinguistic factors should be understood as referring to actual sociolinguistic processes or phenomena. The problem with dimensions of usage used in dictionaries is that they tend to be fairly well formalised and strictly defined, usually in the context of one particular dictionary. In general, however, dimensions of usage are expressed through different labels in different works and in different time periods. There is thus often disagreement as to what the specific levels of usage should be. Given all of these considerations, the term "dimensions of usage" seems more appropriate than "sociolinguistic considerations", and is therefore used in the present discussion. The dimensions of usage in the entries were analysed by annotating the entries for one of the following six categories: FREQUENCY, MODE, REGISTER, SPEAKERS, VALUE, and VARIETY. Examples of these are given in (40)–(45), respectively.

The tag FREQUENCY was given to statements about the use and the frequency of use of a feature, exemplified in (40). MODE refers to mode of expression, or reference to writing or speech, or both (41). A similar category is used in Creswell's (1975: 24–26) analysis of the usage notes in the *American Heritage Dictionary*, where he found that in the usage features on which the *AHD* panel was asked to vote, in more than half of the cases it was not specified whether the questions of usage on which the panel voted referred to speech or to writing. This is perhaps indicative of the lack of inclusion of levels of usage in usage discussions and usage advice, which may in turn produce misunderstanding about the use of a particular feature, and thus make the interpretation of the votes of the panel problematic. The presence of this aspect in usage guides may thus be seen as an important refinement in the treatment

of language usage. REGISTER was used to label statements about contextual or situational aspects of using a certain feature, as shown in (42). The tag SPEAKERS is used to describe references to groups of people identified in the entries (43). VALUE is used to refer to a wider more general meaning of social value associated with the use of a feature. The example in (44) illustrates this dimension. Finally, VARIETY is used to classify all references to a specific regional or social variety of English (45).

(40) Today, split infinitives continue to appear often in Standard speech and even in Edited English. (Wilson 1993: 22)

(41) These uses of *like* are typical of informal spoken language, especially of younger people, and their occurrence in writing is limited chiefly to dialogue. (Pickett et al. 2005: 282)

(42) Some authorities feel that *ain't* would be a useful addition to informal English, particularly as a contraction for *am I not*, which has none that can be pronounced easily. (O'Conner 1998: 128)

(43) Its use is pretty much confined to users of standard English and to literary contexts. (Gilman 1989: 867–868)

(44) Like parallel fifths in harmony, the split infinitive is the one fault that everybody has heard about and makes a great virtue of avoiding and reproving in others. (Follett 1966: 313)

(45) Standard use is hard to explain, but clearly Americans have come down hardest on it, and they have made the rejection stick in Standard American English. (Wilson 1993: 22)

The entries were annotated using the annotation tool 'brat' (Stenetorp et al. 2012),[3] based on the three levels of analysis explained in detail above. This multi-level annotation allows for specific kinds of information to be extracted from the usage guides and to be analysed side by side; these kinds of information include, but are not limited to, the ways in which levels of usage are conceptualised in usage guides, and which groups of speakers or varieties of English are the ones most often referred to. Considerations of register and mode of communication may reveal a less conservative account of usage, while considerations of social value present in guides serve as an important basis for a comparison of these values with attitudes held by speakers. The results of such an analysis, as I will illustrate in the next chapter, show that usage

[3] Available online at http://brat.nlplab.org/.

guide writers, regardless of their tone, often draw on various social aspects of the use of the features they discuss. Even though these considerations may be subjective, they also serve as important indicators of more general opinions about how features are used. Finally, the inclusion of various types of references to dimensions of usage will be shown to be a crucial aspect of the development of the genre towards presenting a more varied and nuanced account of language use.

4.4 Actual language use: data and analysis

My approach to the analysis of the influence of prescriptivism on language change conceptualises language change by drawing on descriptive or usage-based accounts of language change, which analyse frequency patterns of linguistic features. Patterns of variation and change are thus determined by correlating linguistic variables with extralinguistic factors such as time, genre, register, etc. (Nevalainen 2006a: 560–561). Studies of language variation and change in this vein, associated for instance with historical sociolinguistics, have established that when it comes to language change, "the picture that emerges is one of gradual evolution rather than abrupt change" (Mair and Leech 2006: 319). Language variation and change is thus assessed on the basis of identifying "shifting frequencies of use for competing variants", one of which is prescriptively targeted (Mair and Leech 2006: 319); this allows for the importance of prescriptive influence or effect to be assessed. The approach taken in sociolinguistic and descriptive corpus-based research on language variation takes changes in frequency patterns to be indicators of language change (cf. Mair 2006: 2). Consequently, if we take this approach to language change, and if we manage to show that changes in frequencies across time are constrained by prescriptivism-related factors, we can posit an influence of prescriptivism on language change.

The data for the analysis of the actual usage patterns for each feature were extracted from the Corpus of Contemporary American English (COCA, Davies 2008–present) and the Corpus of Historical American English (COHA, Davies 2010b), both created and maintained by Mark Davies at Brigham Young University. Although both corpora are available for online use via the BYU interface, the analysis conducted for this study required the use of the full-text data, which can be purchased in three different formats: plain text, word/lemma/part-of-speech-tagged files, and SQL-database format. This makes wide-ranging exploration of the data possible, in terms of the analysis of many different features, as well as many different aspects

of the texts included in the corpora, such as the type-token ratio of each separate text in the corpus. The full-text data come with some limitations, due to texts in the corpora being affected by copyright restrictions. Consequently, the version of the corpus data available for purchase has been transformed, in order to abide by copyright restrictions. This transformation consists in the removal of ten out of every 200 words of text. This ensures legal use of the copyrighted text in such a way that the strength of the corpus and the validity of the results obtained are not seriously affected, because, as stated on the website, "95% of the data is still there".[4] This means however, that the results obtained in this analysis will be somewhat different, though not significantly so, from the same results obtained through the corpora's online interface.

The COCA corpus is described by its creator as "the first reliable monitor corpus of English" spanning the period 1990–2015 (Davies 2010a: 447; see also Davies 2008–present, 2009), and, as a monitor corpus, it is regularly updated with new data. As of March 2017, according to its website, COCA contained more than 520 million words of text, or 20 million words for each year. The full-text corpus data used for this analysis cover the period 1990–2012; thus the total number of words is lower than 520 million. The COCA corpus is divided into five sections: academic, fiction, magazines, newspapers, and spoken. Davies (2009: 161–162) describes in more detail the selection and sources of materials in COCA, and the specific subcategories for each section are given in Table 4.2. COHA can perhaps be described to some extent as the historical counterpart to COCA. It contains 400 million words, divided into four sections: fiction, magazines, newspapers, and non-fiction books (Davies 2012). These four sections are further subdivided into subgenre categories, given in Table 4.2. A complete list of sources for both corpora is available on the website, while Davies (2009, 2010a, 2012) describes in detail how the corpora have been created and how they can be used to study language variation and change in different ways.[5]

It is important at this point to raise the question of what the different corpus sections, and their subcategories, represent in terms of register variation in historical and present-day American English. The understanding of the nature of texts in the two corpora is instrumental in interpreting the results, for two main reasons. The first one relates to the comparability of the two corpora. In other words, it is important to understand in what ways the two corpora are similar, as well as where they differ, so that the interpretation of the results can take this into account. The general intention

[4]For more details see http://corpus.byu.edu/full-text/limitations.asp.

[5]I use the terms 'genre' and 'register' interchangeably when referring to the various corpus sections.

COCA	***COHA***
Academic	**Non-fiction**
History	
Education	
Geography/Social Science	
Law/Political Science	
Humanities	
Philosophy/Religion	
Science/Technology	
Medicine	
Miscellaneous	
Fiction	**Fiction**
General (Books)	Drama (Plays)
General (Journal)	Movie Scripts
Science Fiction	Novels
Juvenile	Poetry
Movie Scripts	Short stories
Magazine	**Magazine**
News/Opinion	No subgenres
Financial	
Science/Technology	
Society/Arts	
Religion	
Sports	
Entertain	
Home/Health	
African-American	
Children	
Women/Men	
Newspaper	**Newspaper**
Miscellaneous	No subgenres
International News	
National News	
Local News	
Money	
Life	
Sports	
Editorial	
Spoken	
No subgenres	

Table 4.2: Sections and subsections of COCA and COHA

underlying the use of these two corpora was that they would provide comparable evidence, at least for some registers, so that the development of usage could be traced over a longer period of time. In the case of fiction, for instance, this is quite straightforward (Table 4.2).

In terms of comparability, COCA and COHA contain similar material for the fiction, magazines, and newspapers sections in both corpora. For the historical corpus, some text types, such as movie scripts, newspapers, and magazines, are only available from the twentieth century onwards. A spoken section is only available in COCA, so my analysis of spoken data is restricted to the period 1990–2012. Finally, COCA has an academic section, while COHA has a non-fiction section. While these two are different categories, it seems a priori not unreasonable to assume that they are not dissimilar, given that they both contain a preponderance of relatively formal texts. As we will see in Chapter 6, this assumption is supported by the frequency patterns observed in these two genres. An additional reason for this consideration is that the organisation of the subcategories of texts for both the academic section of COCA and the non-fiction section of COHA is according to the general Library of Congress classification system (cf. Davies 2009: 162; Davies 2012: 124–125).

Finally, a limitation of COCA, when it comes to working with these subcategories, is that the subsections included in the spoken section are not based on actual genres of spoken language, but on the sources from where the texts were obtained. On the naturalness of the spoken data, Davies notes that they "are based almost entirely on transcripts of *unscripted* conversation on television and radio programs" (Davies 2009: 162) and that these texts are accurate and spontaneous; in other words, they provide a good representation of non-media English. The corpus data were then used to extract all occurrences of the various linguistic features with the help of Python scripts. In most cases, occurrences were extracted from the part-of-speech-tagged files of the corpora, using regular expressions. The use of the part-of-speech-tagged data greatly facilitated further analysis of each occurrence.

The data analysis can be divided into two parts. In the first part of the analysis, I look at the frequency of occurrence of a particular feature across subsections of the corpora, as well as across time periods. For the analysis of patterns of use, I will rely on the two approaches used in corpus-based linguistic analysis of variation and change suggested by Biber et al. (2016): the text-linguistic approach and the variationist approach. As Biber et al. (2016) show, these two approaches can sometimes produce different results, as an analysis of the frequency distributions of one variant is a good indication of the rate of occurrence of that variation, while the results of a

variationist analysis allow us to "report proportional preference; they do not actually tell us how often a listener/reader will encounter these structures in texts" (Biber et al. 2016: 359). Leech and Smith (2009: 176–178) also discuss the difference between variationist and text-linguistic measures of frequencies of usage, but use the terms "proportionate" method and "normalisation" method to refer to variationist and text-linguistic approaches, respectively. What a variationist account shows us is how often a variant occurs, in the context of two or more options, out of a total number of possible occurrences. In the second part of the analysis, I specifically look at whether and how additional prescriptively targeted features may predict the use of a proscribed feature. This analysis is done using the split infinitive as a case study. In what follows, I discuss each of these two types of analysis in more detail.

When establishing the variable context, it is important to distinguish between contexts in which variation is a choice, as opposed to contexts where one variant is categorically used (cf. Poplack and Dion 2009: 571: "in contexts where speakers must choose among the major variants"). With reference to the features used in this study, the assumption is made that theoretically, both the standard and the proscribed variants are possible options, given that the notion of 'incorrectness', prescriptively speaking, is linguistically arbitrary. Of course, establishing variants was not possible in the context of the discourse particle *like* and non-literal *literally* (cf. Section 2.6), since it is difficult to conceptualise a variable in which *like* is one of the variants; in addition the environments in which this variant could occur would be almost impossible to predict or determine (but see D'Arcy 2007).

The analysis of *ain't* was done on the basis of both text-linguistic and variationist frequencies. These were analysed in the context of two variables: present tense negative *be not* and *have not*. The variants are given in Table 4.3. It is also important to address a point of disagreement present in the methodological approaches employed in previous studies to determine the variables in which *ain't* is used, that is, the forms with which it alternates. Wolfram (1974: 153), for instance, considers *ain't* as alternating only with the *'m / 're / 's not* forms of *be not*, noting that *aren't* and *isn't* were not observed in his sample of speech from Puerto Rican English speakers, and were thus excluded from the analysis. Cheshire (1981) considers *ain't* as a variant of all contracted forms of present *be not* and *have not*, including both auxiliary and copular *be not*, but not the full forms of these verbs (this is not stated explicitly, but can be inferred from the data presented). Weldon (1994) considers *ain't* as a variant of both full and contracted copula in *be not* environments, while Anderwald analyses the alternations of *ain't* with *be not* and *have not*, including both auxiliary

and copula *be not*, but does not specify whether the total number of environments includes full forms. Taking into account these differing approaches, in my analysis I take both full forms and contracted forms of *be not* and *have not*, primarily because my data come from a general standard American English corpus, in which both full and contracted forms occur. Second, I do not distinguish between copula and auxiliary *be not* environments, because (a) such a distinction is not particularly relevant for the issue of the influence of prescriptivism on usage, and (b) studies of variation have shown that it is hard to identify regularities in the patterns of variation between copular *be not* and auxiliary *be not*, as opposed to *have not*, leading Anderwald (2002: 139) to conclude that "a distinction of BE into auxiliary and copular uses is perhaps not particularly warranted".

	be not		***have not***	
prescribed	*am / are / is not* *'m / 're / 's not* *aren't / isn't*	*I am not coming.* *I'm not coming.* *He isn't coming.*	*have / has not* *'ve / 's not* *haven't / hasn't*	*I have not left.* *I've not left.* *I haven't left.*
proscribed	*ain't*	*I ain't coming.*	*ain't*	*I ain't left.*

Table 4.3: The variants of *be not* and *have not*

For the analysis of the newer uses of *like*, as discussed in Section 3.4 above, I will focus only on the discourse particle *like*. While some of the occurrences of the discourse particle *like* are tagged with appropriate part-of-speech tags in the corpora used, it was noticed that the accuracy of the tagging was not very high, and that in most cases a much better indicator of whether *like* is used as a discourse particle or not was the transcription of the data. Accordingly, since the discourse particle *like* is set off with commas in the majority of the cases, this was used as a more reliable way of identifying and extracting those instances of *like*. In the cases of *like* and *literally*, for instance, a variationist analysis was not undertaken, due to the complicated issue of establishing the variable context (but see D'Arcy 2007 for a variationist analysis of vernacular *like*). In the case of *like*, only normalised frequencies of occurrence were therefore used to track the patterns of usage of this feature, based on the tags in the corpora.

In the case of *literally*, two analyses were carried out: one on the basis of the entire set of occurrences of *literally*, and another on the basis of a subset of occurrences. This decision was made in view of the fact that the process of change which *literally* is undergoing entails a significant amount of variation in its uses, and, consequently, it

is difficult to neatly disambiguate the meanings and functions of these uses, especially of the newer ones. The three uses of *literally* distinguished in the analysis are given in Table 4.4; these were distinguished on the basis of categorisations of the uses of *literally* found in previous studies (see Section 3.5).

	uses of *literally*	
prescribed	primary	*This is literally translated.*
proscribed	dual non-literal	*There were literally thousands of people.* *This book literally blew my mind.*

Table 4.4: The uses of *literally*

In its primary use, *literally* refers to what something means, how something is said or meant, or how something is translated or interpreted. In such cases, *literally* has a clear denotational meaning, as it functions as a manner adverb; here *literally* is the answer to the question how something is done (e.g. *How do you mean? I mean literally*.) In all other cases, however, the meaning of *literally* is more ambiguous and elusive, and in such cases it is almost always possible that the speaker is using *literally* to signal that something that may be understood figuratively must be understood literally (cf. creative cases in Powell 1992, discussed in Section 3.5 above). In these cases, the meaning of *literally* is highly dependent on context. In view of this, a manual analysis of all occurrences of *literally* was deemed too laborious to be worthwhile, given that it would be unlikely to reveal any major insights. A middle-ground solution was to perform both an automatic binary disambiguation of the entire set of occurrences of *literally* using Python scripts and a manual analysis classifying a subset of the occurrences of *literally* into three categories. The first type of analysis distinguishes between cases that are very clearly instances of the so-called primary use of *literally*, and all other cases (for more details on how this was done, see Appendix C). An important consideration in this decision was the fact that prescriptive attitudes are highly conservative with respect to most types of changes in the language. In the treatment of *literally*, especially, only the primary (i.e. denotational) meaning of the word was therefore accepted, while all other uses (i.e. intensifying, subjective or metapragmatic ones) are seen as 'incorrect'. This first part of the corpus analysis is based on data from both corpora, and allows us to track the usage of the proscribed variants over time.

The manual analysis was carried out on a sample of the total number of occurrences of *literally* in the corpora. This sample was extracted by selecting every

fifth case in the set of all occurrences of *literally* in both COCA and COHA. This manual analysis was done for two reasons. First, it was meant to provide additional evidence for the distribution of the three uses of *literally*, which was not possible with the automatic disambiguation of all its occurrences. Second, the manual analysis also served to identify any possible differences between the results of the automatic disambiguation and the manual analysis of the data. In this way, I hope to be able to provide a rough estimate of the level of accuracy of the automatic disambiguation.

An analysis of the frequency patterns of negative concord also requires a variationist account of the ratio of the stigmatised forms in comparison to all others; here, we cannot rely on frequency counts of negative concord constructions alone. In order to delimit the total number of possible environments in which negative concord may occur, it is important to consider the variants of negative expression and the circumstances in which negative concord can be expected. The rule of negative attraction (first formulated by Klima 1964: 267, 289, cited in Labov et al. 1968: 268; see also Wolfram 1974: 163) is the starting point for determining the contexts for negative concord. According to this rule, in standard English sentences with indefinites, there are two possible options. If the indefinite is in pre-verbal position, the negative marker is attracted to the first indefinite before the verb, which accounts for sentences such as *Nobody knows anything*, to use Labov et al.'s example (1968: 268). If the indefinite is post-verbally located, the negative marker may optionally be attached to the verb, as in *John doesn't know anything*, or to the indefinite, as in *John knows nothing*. In other words, if the indefinite comes after the verb, "the negative attraction rule may or may not apply" (Wolfram 1974: 164). In relation to these two variants, Wolfram (1974: 165) further notes that the latter is more characteristic of literary than of colloquial English. In African American English, as well as in non-standard varieties of English, the rule of negative attraction does not apply in cases where post-verbal indefinites keep the negative marker, resulting in instances such as *John doesn't know nothing*. In this case, "what takes place is a copying of the negative on as many post-verbal negatives as there are in a sentence" (Wolfram 1974: 165). In applying this rule to determining potential environments in which negative concord can occur, so that we can establish the ratio of usage of negative concord (cf. Smith 2001; Nevalainen 2006b), we can employ the variants given in Table 4.5. Since my goal here is not to give an account of negative concord in the entire language system, in order to simplify the analysis, I limited myself to investigating negative concord occurrences in sentences with the following indefinites: *anybody/nobody*, *anyone/no one*, and *anything/nothing*.

	V + INDEFINITE	
prescribed	V-neg + INDEFINITE V + INDEFINITE-neg	*I haven't seen anybody.* *I have seen nobody.*
proscribed	V-neg + INDEFINITE-neg	*I haven't seen nobody.*

Table 4.5: The variants of negative concord

Pronouns in coordinated phrases were extracted only with first person pronouns. The contexts of variation were further limited to include only certain types of cases in which pronouns are used in coordinated phrases. For instance, the analysed cases of object *I* and subject *me* are only those cases in which the pronouns *I* and *me* are found in coordinated phrases where the other phrase-constituent is a proper noun. This decision was made in light of the fact that there may be additional constraints affecting the realisation of *I* or *me*, especially if the other constituent is another pronoun. Phrases with proper nouns were seen as presenting a sufficiently uniform context in which the realisation of *I* or *me* will not be expected to be affected by the case of the other phrase-constituent. The secondary reason for this decision was of a practical nature. The identification of coordinated phrases functioning as subjects or objects in which one of the constituents in the phrase is *I* or *me* was not a straightforward task of automatic extraction from the corpora. The restriction to cases with proper nouns significantly reduced the danger of extracting a large number of false positives from the data, without influencing the quality of the data. The variants for pronouns are given in Table 4.6.[6]

	Subject		**Object**	
prescribed	*x and I* *I and x*	*Elly and I left.* *I and Elly left.*	*x and me* *me and x*	*They saw Elly and me.* *They saw me and Elly.*
proscribed	*x and me* *me and x*	*Elly and me left.* *Me and Elly left.*	*x and I* *I and x*	*They saw Elly and I.* *They saw I and Elly.*

Table 4.6: The variants of first person pronouns in coordinated phrases

The split infinitive was also analysed on the basis of both text-linguistic and

[6]The table contains *I and x* as a 'prescribed' variant simply because the form of the first person pronoun adheres to the prescription that the nominative form should be used in subject positions. However, it should be noted that there is a different set of norms against this use, mostly having to do with the impoliteness of referring to oneself first. I have not taken this into account in the analysis; however, I do not consider this a significant influence on the results, as this variant was extremely rare in the corpus data.

variationist frequencies. To establish the text-linguistic frequencies, all infinitives split by one modifier were extracted from the data. The variationist frequencies were established only on the basis of a specific variable context, which was established as infinitives modified by one lexical adverb ending in *-ly*. The reason for this was that different elements which may be placed between the *to* and the infinitive behave differently. Good examples of this are cases in which *not* is placed between *to* and the infinitive verb. In these cases, there are two possible variants: *not* + *to* + verb or *to* + *not* + verb, which is a pattern of variation different from infinitives modified by a *-ly* adverb, because *-ly* adverbs can be placed after the verb, in addition to before the verb and before *to*. The variants thus established are given in Table 4.7.

	MODIFIED INFINITIVE	
prescribed	*to* + INFINITIVE + modifier	*to improve significantly*
	modifier + *to* + INFINITIVE	*significantly to improve*
proscribed	*to* + modifier + INFINITIVE	*to significantly improve*

Table 4.7: The variants of the modified infinitive

On the basis of the variables outlined here, the occurrences of these linguistic features were extracted from the corpus data. The sizes of the various datasets differed; Table 4.8 gives an overview of the total number of occurrences of each feature in the corpus data. More specific information on sample sizes, as well as the number of occurrences of prescribed, as opposed to proscribed, variants are given in the relevant section in Chapter 6.

Feature	**COHA**	**COCA**
ain't	39,348	12,228
like	634	10,020
literally	6,848	14,946
negative concord	10,041	8,530
object *I*	194	380
subject *me*	456	819
split infinitive	10,062	63,079

Table 4.8: Raw frequencies for each of the features extracted from COHA and COCA

The second part of the corpus analysis which I mentioned above focuses specifically on investigating the relationship between proscribed variants of a particular feature and the use of other prescriptively targeted features. This analysis

is based only on the COCA data, and it is conducted using the split infinitive as a case study. The general approach is based on Hinrichs et al. (2015), who distinguish between intralinguistic and extralinguistic constraints on the variation in the realisation of the relative pronoun in restrictive relative clauses. In addition to the intralinguistic and extralinguistic constraints, Hinrichs et al. (2015) also included a number of prescriptivism-related constraints, expressed as frequency of occurrence of other prescriptively targeted features. This was done on the basis of the assumption that if a feature's variation is affected by prescriptive constraints, then these constraints would also be noticeable in the context of other features that might be expected to be affected by prescriptivism. They thus included four features in their multivariate analysis. In the present analysis, the same principle was used; however, the number of additional prescriptivism-related predictors was increased. The goal of this analysis is to investigate the potential influence of prescriptivism at the level of individual texts and on the basis of a number of different language features. In this way, it is hoped that this analysis will supplement the separate analyses on the six language features which this study focuses on.

The dataset used for this analysis was composed of all the occurrences of infinitives modified by a single *-ly* adverb extracted from COCA. The dependent variable was defined as MODIFIED INFINITIVE, with two levels: SPLIT and NON-SPLIT. Thus, each occurrence of a modified infinitive in the dataset was classified at either of the two levels. A set of additional predictors were defined, as outlined in Table 4.9; each occurrence of a modified infinitive in the dataset was additionally coded for each of these predictors. In terms of internal predictors, I distinguish the semantic class of the adverb and the length of the adverb compared to that of the verb, i.e. the independent variables ADVERB TYPE and ADVERB LENGTH, respectively. The ADVERB TYPE for each occurrence of a modified infinitive was determined on the basis of the semantic classification of adverbs in Biber et al. (1999: 552–560). The adverbs modifying the infinitive were classified into the following categories: ADDITIVE-RESTRICTIVE adverbs (e.g. *especially*), DEGREE adverbs (e.g. *almost*), LINKING adverbs (e.g. *therefore*), MANNER adverbs (e.g. *happily*), STANCE adverbs (e.g. *probably*), and TIME adverbs (e.g. *recently*). The length of the adverb was operationalised as the number of syllables and as a categorical variable with three levels: SHORTER, if the adverb is shorter than the verb measured in number of syllables; EQUAL, if the adverb has the same number of syllables as the verb; and LONGER, if the adverb has more syllables than the verb. External predictors included in the analysis are YEAR and GENRE. YEAR is a continuous variable with

values from 1990–2012, associated with the source year of each corpus text in which relevant instances of modified infinitives were identified. The predictor GENRE is a categorical variable with five levels, corresponding to the five corpus sections: ACADEMIC, FICTION, MAGAZINE, NEWSPAPER, SPOKEN.

Predictors	**Levels**
Internal predictors	
ADVERB TYPE	ADDITIVE-RESTRICTIVE DEGREE LINKING MANNER STANCE TIME
ADVERB LENGTH	LONGER EQUAL SHORTER
External predictors	
YEAR	1990–2012
GENRE	ACADEMIC FICTION MAGAZINE NEWSPAPER SPOKEN
Prescriptivism-related predictors	
And/*But* *data is* *hopefully* *less* + plural nouns *these kind of*/*these sort of* *none are* passives *shall* *try and* *whom*	frequency per 1,000 words

Table 4.9: Predictors used in the analysis of prescriptive constraints on the use of split infinitives

Finally, prescriptivism-related predictors are a number of additional features that are also often proscribed in usage guides. The additional features were selected on the basis of their frequency of occurrence in the prescriptive literature, as well as their relative ease of analysis. The HUGE database, described in more detail in the previous section, was the tool used to assess the most commonly treated features in

usage guides. On the basis of a search in the HUGE database, a list of features can be extracted and ordered by the number of guides in which the particular features are treated. Features were then further selected in view of their relative ease for automatic disambiguation. The additional features thus selected are: the use of *and/but* at the beginning of the sentence, singular *data*, *hopefully*, *less* with plural nouns, *these kind/sort of*, *try and*, plural *none*, passives, *shall*, and, finally, *whom*. In addition to these features, *ain't*, *like*, *literally*, and negative concord were also included in the dataset. For each text in the corpus data, the frequency of each of these features was established; the raw frequencies were normalised per 1,000 words. In order to obtain a more uniform dataset, texts from the corpus whose total number of tokens was too low (e.g. 500 words) or too high (e.g. 90,000 words) were excluded from this analysis, because the frequency of occurrence of the features used as predictors would be affected by the differences in size of the texts, especially when it comes to short texts, as this produced many zeros in the dataset. The final dataset contained 4,925 occurrences of a modified infinitive across the same number of texts. The analysis applied to this dataset was fixed-effects binomial logistic regression. Logistic regression is a technique applied in cases where the outcome variable, or the dependent variable is a one of two possible values – in this case SPLIT or NON-SPLIT infinitive (see Baayen 2008: 195). The technique is used to estimate the probability of one of these two possible values in comparison with the other, given the set of predictors – in this case all the predictors described above. In other words, the technique allows us to investigate questions such as: do texts in which, for instance, the frequency of *ain't* is high predict the probability of a modified infinitive being realised as split as opposed to non-split? The assumption for conducting such an analysis is that the observations are independent from each other. In order to satisfy this assumption, and make sure the cases are independent, only one case of a modified infinitive from each text was used.

4.5 Attitude survey: data and analysis

Finally, with a focus on present-day English specifically, data on the attitudes of speakers of American English towards the six linguistic features investigated in this study were collected and analysed. The data on speakers' attitudes were collected during a two-month fieldwork stay in Los Angeles, in 2014. The choice of Los Angeles was determined partly by the fact that it is one of the biggest metropolitan areas in

the United States, where I expected to find a great deal of variation in speech, and speakers, and partly because of contacts I had established there, which allowed me to recruit respondents more easily, and to enter communities as a 'friend-of-a-friend'. The main goal was to collect data from speakers who do not belong to the category of 'language professionals' and are not solely university students, as is the case with some previous studies on attitudes to usage (cf. Leonard 1932; Mittins et al. 1970; Albanyan and Preston 1998). Through a friend who worked at a start-up company in Beverly Hills, I interviewed a number of young adult professionals, who were chosen for my research in the attempt to collect data from respondents other than university students as well. Through my friend's family members, I interviewed a number of older professionals, and an additional number of respondents were recruited through the contacts that were established via these interviews. The aim was to arrive at a sample which is varied in terms of age. An additional number of respondents also came from Santa Monica College, where I distributed flyers to recruit potential respondents. Here, the focus was on recruiting first-generation students. The respondents received ten dollars for their participation, which lasted between 30 and 60 minutes. Table 4.10 shows the make-up of the sample of respondents. One limitation of the sample in this study is that it does not form a representative random sample of the population of Los Angeles. Not only was a fully random stratified sample beyond the scope of this data collection process because of time constraints, it would have also entailed the determination of categories of speakers a priori, an approach which raises its own methodological issues. In other words, such a sample would have meant that certain social categories or social variables would have needed to be defined in advance

The respondents were selected using a 'friend-of-a-friend' technique, resulting in a convenience sample of 79 respondents in total; their responses were used in the analysis of attitudes in Chapter 7. Table 4.10 shows that the sample of respondents is skewed towards younger adults. This limitation of the sample of respondents makes it difficult to carry out a comparison across all the age categories, but it does provide insights into how the attitudes of these speakers differ across usage features and contexts of use. The age categories presented in the table break down the sample by 10-year groups, but these were not used in the analysis of the data, where I divided informants into two age groups (see below, and Chapter 7).

Age group	Female	Male	African American	Other	White	Total
19–29	22	24	14	18	14	46
30–39	7	11	0	8	10	18
40 +	7	8	3	3	9	15
Total	36	43	17	29	33	79

Table 4.10: Distribution of respondents according to age, gender, and ethnicity

The meeting with each respondent consisted of two parts. The first part was a survey in which the respondent was asked to rate sentences containing the usage problems investigated in this study. The sentences used in this part of the survey were selected from COCA, described in the previous section of this chapter, and were used in different contexts, based on the corpus section in which they were found. This was done to ensure naturalness of the stimuli. Rather than presenting the respondent with a written choice of register (e.g. 'acceptable in formal writing', etc.), which is the approach taken in previous studies on attitudes to usage (cf. Mittins et al. 1970; Ebner 2017), the variable context of use was included as part of the stimulus. Thus, there were different types of sentence stimuli for each of the linguistic features investigated, in different contexts: 'spoken informal', 'spoken formal', 'written informal', and 'written formal'; the contexts for each of the sentences are given in Table 4.11. As the table shows, not all language features were included in the survey in all four contexts, simply because certain features are highly unlikely to occur in some of the contexts. For instance, the discourse particle *like* was only included in 'spoken informal' and 'spoken formal', because it is very unlikely to be encountered in written contexts.[7]

Each sentence had to be rated by respondents on four criteria: 'correctness', 'acceptability', 'goodness', and 'educatedness', using five point semantic-differential scales: CORRECT-INCORRECT, ACCEPTABLE-UNACCEPTABLE, GOOD ENGLISH-BAD ENGLISH, and EDUCATED-UNEDUCATED, as exemplified in Figure 4.2. For spoken stimuli, the survey contained a link to an audio file, followed by the same kind of structured response as the one exemplified in Figure 4.2.[8] The ratings for these four semantic-differential scales were taken as evidence for the ways in which attitudes of speakers towards the use of the six linguistic features differed across a number of variables, explained below. These variables were established on the basis of a number

[7] An exception here might be online language use, or specifically the language used on social media and in discussion groups. However, even in these contexts there is little evidence as to the extent to which the discourse particle *like* would be used. If it does occur, this is likely a fairly new development.

[8] The entire survey is available at https://bit.ly/2xWraST.

Feature	Context	Stimulus sentence
ain't	Spoken informal (conversation)	I ain't going to see them next month.
ain't	Spoken formal (interview)	In school they ain't pushing me, they are encouraging me.
ain't	Written formal (newspaper article)	You won't move forward in your career if you ain't brave enough.
like	Spoken informal (conversation, male)	Didn't you, like, all like go to, erm..., like a boot camp?
like	Spoken informal (conversation, female)	I've like done a couple of like summer camps in like languages and accounting.
literally	Spoken informal (conversation, male)	I literally died from boredom on my date last night!
literally	Spoken informal (conversation, female)	There is story in this book that literally blew my mind!
literally	Written informal (social media)	This book literally blew my mind.
negative concord	Spoken informal (conversation)	I'm strong minded and I'm not going to let nobody lead me off in the wrong direction.
negative concord	Written informal (text message)	I'm sorry. But I'm not going to argue with nobody.
negative concord	Written formal (novel)	I thanked the good lord that I had not killed nobody.
pronouns: object *I*	Spoken informal (conversation)	I think this has been the trouble between you and I.
pronouns: object *me*	Written formal (academic article)	These findings have been very important for my colleagues and me.
pronouns: object *I*	Written informal (social media)	This trip has been a great adventure for my parents and I.
pronouns: object *I*	Written formal (professional email)	The collaboration with your company has been a great pleasure for my workers and I.
pronouns: subject *me*	Spoken informal (conversation)	Me and my husband went to a party with several other young couples.
pronouns: subject *I*	Written formal (professional email)	My colleagues and I will look into this and get back to you as soon as possible.
pronouns: subject *me*	Written informal (text message)	Me and dad are on our way home!
pronouns: subject *me*	Written formal (professional email)	My team and me are working to resolve your problem as soon as possible.
split infinitive	Spoken formal (radio interview)	So, I would encourage young men and women to seriously consider a career in law enforcement.
split infinitive	Written informal (social media)	Trying to decide if there is anything interesting to further explore in my new town.
split infinitive	Written formal (magazine article)	This therapy has been shown to significantly reduce the risks of heart attacks and strokes
split infinitive	Written informal (social media)	Trying to find out if there is anything interesting to explore further in my new town.

Table 4.11: Stimuli sentences used in the survey

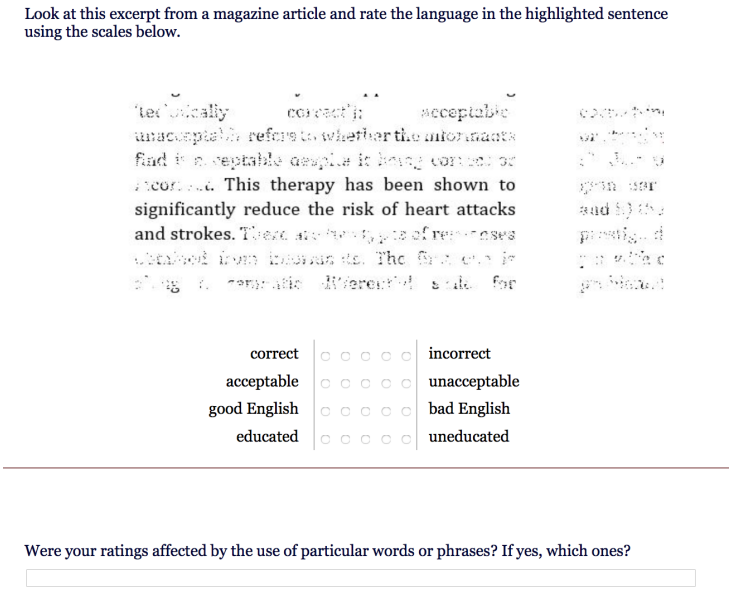

Figure 4.2: An example of the way stimuli sentences were presented to participants in the survey

of general questions about the respondents' backgrounds, which were also included in the survey. The second part was a follow-up unstructured interview with each of the respondents, after the completion of the sentence evaluation. The main purpose of the interview was to allow respondents to reflect on the survey, as well as to communicate thoughts and observations they may have felt were impossible to address in the survey. The interviews were thus fairly unstructured, but the topics covered were naturally related to the respondents' attitudes to language use, as well as the usage problems covered in the survey.

The variables included in the analysis of attitudes are the following. The dependent variables are the ratings of the stimuli sentences in different contexts. The three independent variables are: AGE, GENDER, and ETHNICITY, established on the basis of relevant questions in the survey. Age was operationalised as a nominal variable with two levels 29 AND BELOW and 30 AND ABOVE. The information about the gender of the respondents was obtained by asking an open question ("What is your gender?"). This produced binary data: all of the respondents chose either MALE or FEMALE. Consequently the gender variable was operationalised as a binary variable, although there is an increasing tendency in sociolinguistic research to operationalise

it as a categorical variable, i.e. with multiple levels, such as *male*, *female*, and *other*. Information about ethnicity was similarly obtained with an open question ("How do you describe your ethnic background?"). On the basis of the answers, the respondents were grouped into three groups in terms of ethnicity; this resulted in a categorical variable with three levels: AFRICAN AMERICAN, OTHER, and WHITE. The effects of these variables on the ratings of the stimuli sentences will be explored using non-parametric tests for inter-group comparisons of the ratings of the stimuli sentences. The ratings produced ordinal data which are not normally distributed, so testing between the ratings of two groups was done with the Wilcoxon (Mann-Whitney) test for independent samples (Levshina 2015: 108–113). Because multiple comparisons were conducted on the same data, the significance level was Bonferroni corrected (Levshina 2015: 181), and differed for each of the features, as a different number of tests were conducted for each feature. These aspects of the analysis are addressed in more detail in Chapter 7, which discusses the results of the analysis of the speakers' attitude data.

4.6 Conclusion

In this chapter I have presented the general approach to my analysis of the relationship between prescriptive attitudes to usage in American English, patterns of actual language use, and speakers' attitudes. For each of these perspectives, I have outlined the data used, how the data were collected, and the analytical approaches used to explore these data. In the following chapters, the analysis of each of these datasets is presented and discussed.

With respect to the analysis of the data, I have used a number of software programs and tools. I mention each specific tool in the chapter where I discuss the analysis for which I have made use of that tool. To provide a brief overview of these tools: for the analysis of the usage guide data I used the Hyper Usage Guide of English database (Straaijer 2015) and the 'brat' annotation tool (Stenetorp et al. 2012). I used R (R Core Team 2013) for all statistical analyses and visualisations. `Ggplot2` (Wickham 2009) was used to produce most of the plots in Chapters 5 and 6 and the `Likert` package (Bryer and Speerschneider 2017) was used to produce the plots in Chapter 7. The package `rms` (Harrell 2018) was used to conduct logistic regression analyses.

CHAPTER 5

Metalinguistic commentary in American usage guides

5.1 Introduction

In this chapter, I will present a fine-grained analysis of the metalinguistic commentary on the six linguistic features discussed in Chapter 3 in a collection of American usage guides, as described in Section 4.3. The two main goals of this analysis are (a) the identification of attitudes expressed in American usage guides towards the linguistic features, and (b) the identification of possible changes in those attitudes. As explained in Section 4.3, attitudes to usage in metalinguistic works such as usage guides can be expressed in various ways, which called for an analysis of more than a single indicator of potentially prescriptive attitudes. I will therefore look at three aspects of metalinguistic commentary in American usage guides: the treatment of usage features, the attitudes expressed towards these features, and a number of additional dimensions of usage which are referred to in the treatment of these features.

Three levels of analysis are distinguished: treatment, attitudes to usage, and dimensions of usage. The first level of analysis involves determining the overall treatment of the features in the usage guide entries by classifying each entry into

one of three mutually exclusive categories, ACCEPTABLE, UNACCEPTABLE, and RESTRICTED, as explained in Section 4.3. The results of the analysis of treatment are discussed in Section 5.3. The second level of analysis, attitudes to usage, concerns the identification and classification of explicit expressions of POSITIVE or NEGATIVE attitudes in the treatment of usage features. This analysis is presented in Section 5.4. The third level of analysis, dimensions of usage, identifies the dimensions of usage which usage guide writers refer to in their discussion of the features; the categories analysed at this level are: FREQUENCY, MODE, REGISTER, SPEAKERS, VALUE, and VARIETY. This part of the analysis is covered in Section 5.5. Before I present and discuss the findings of these three levels of analysis, in Section 5.2 I will provide a general description of the entries, focusing specifically on differences in the frequency with which the features are covered in the collection of usage guides consulted. Finally, in Section 5.6 I summarise all types of findings, and discuss their relevance for the first set of research questions formulated at the outset of this thesis. In doing so, I address the importance of these results in the context of previous studies of usage guides, and their importance for the study of actual language use and speakers' attitudes.

5.2 Coverage of the language features in usage guides

In Section 4.3 I explained how I compiled the corpus of entries on which the following analysis is based. The analysis of the corpus revealed a considerable degree of variation in terms of both the length of the entries and the content of their discussion. Entries vary in length from one line to paragraph-length entries, exemplified with the two entries on *ain't* given in (46) and (47) below; in Gilman (1989: 60–64), for instance, the entry on *ain't* is five pages long. This aspect of the entries seems to change over time, with nineteenth-century entries being in general significantly shorter than those from the second half of the twentieth century; this is true for all of the linguistic features analysed.

(46) "That ain't just," should be, That is not just. (Anonymous 1856a: 68)

(47) "Ain't" has a long and vital history as a substitute for "isn't," "aren't," and so on. It was originally formed from a contraction of "am not" and is still commonly used in that sense. Even though it has been universally condemned as the classic "mistake" in English, everyone uses it occasionally as part of a joking phrase or

> to convey a down-to-earth quality. But if you always use it instead of the more "proper" contractions you're sure to be branded as uneducated. (Brians 2003: 6)

With regard to content, a close reading of the entries reveals interesting and important changes over time in the way in which these usage features are discussed, as well as in the kinds of information included in the discussion. It is, for instance, worth noting at this point that there is some variation in exactly what is thought of as problematic when it comes to the various usage problems (an issue already pointed out in Section 2.6). For instance, the term 'double negative' is used to refer both to the double marking of negation in sentences such as *I haven't seen nobody* (i.e. negative concord) and to examples such as *She is not unhappy* (i.e. litotes) (cf. the entry in Brians 2003: 62, which covers both features under the same heading). These constructions are different from a linguistic point of view, and are also problematic in different ways: negative concord is more strongly stigmatised and more salient than litotes. The former structure is considered grammatically incorrect from a prescriptive point of view, on the basis of the argument that two negatives make a positive, while the latter is considered grammatically correct, but stylistically problematic because it is often considered intentionally misleading (cf. Randall 1988: 241–243). Similar distinctions and peculiarities are also found in the treatment of other usage features. In the context of the usage problem *ain't*, various nuances of 'problematicity' can be distinguished: *ain't* is considered less problematic when used with first person pronouns, or in first person question tags, compared to *ain't* in the third person. *Ain't* used in place of *be not* is likewise generally considered less problematic than *ain't* used in place of *have not*. In what follows, I discuss the various patterns of coverage of the features, as well as some observations made on the basis of a close reading of the entries.

5.2.1 *Ain't*

Ain't is discussed in 46 entries, and appears in 41 guides in total in the collection. Figure 5.1 shows the proportion of usage guides which discuss *ain't*, out of the total number of usage guides per decade. As the figure shows, *ain't* is fairly frequently covered in usage guides, especially in those published in the twentieth century, with about half of the usage guides analysed per decade discussing the feature. For the 1990s and the 2000s, fewer than half of the consulted guides discuss *ain't*; this may be an indication either that the form is slowly becoming less problematic or, as I will

argue below, that its use is limited to certain contexts in which *ain't* is accepted and therefore merits little discussion. Even though in the majority of the cases *ain't* is discussed as a variant for present *be not* forms, some variation was observed in this respect. In 16 of the 46 entries, *ain't* is treated as a variant of both present tense *be not* and present tense *have not*. Out of these 16, only four make a clear distinction in acceptance between *ain't* for *be not* and *ain't* for *have not*, explicitly judging the latter to be much less acceptable than the former. For instance, Utter (1916: 29) declares that "there is no defense possible for the vulgar use of *ain't* for *hasn't* and *haven't*".[1] Another topic identifiable in the treatment of *ain't* is the distinction between *ain't* as a substitute for *am not*, especially in questions, as opposed to the use of *ain't* as a variant for other forms of present tense *be not*. Fewer than half of the entries make this distinction, and the consensus is that *ain't* is more acceptable as a contraction for *am not*, that is, when used in the first person singular, than it is for all other present tense forms of *be not*.

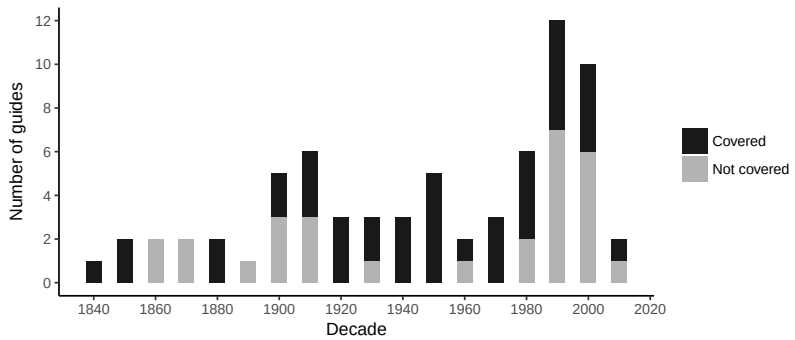

Figure 5.1: Number of usage guides with entries on *ain't* per decade (n, guides = 70; n, *ain't* entries = 41)

Ain't was first commented on in Hurd (1847) and in exactly half of the other usage guides published in the nineteenth century, as shown in Figure 5.1. These comments are, without exception, very short entries, which list *ain't* as an error and give the correct options, such as *am not* / *are not* / *is not*, or *have not* / *has not*, in cases where *ain't* is used in the environment of present tense *have not*. As already mentioned, the

[1] It is important to note that the variant *hain't* is also mentioned in some usage guides in a separate entry – for instance, Hurd (1847: 40), Ayres (1911: 112), Vizetelly (1920: 102), and Wilson (1993: 219). These entries, however, were not taken into account in the present analysis because *hain't* is considered to be a separate feature from *ain't*, and because it is much less frequent, both in usage guide data and in corpus data.

use of *ain't* for *have not* is not discussed to the same extent as the use of *ain't* for *be not*; for instance, the use of *ain't* for *have not* is not mentioned in the usage guides from the nineteenth century. Bechtel (1901) appears to be the first usage guide writer to give a lengthier discussion of *ain't* by referring to its uses as a variant of *be not*, as well as to its etymological predecessor *an't*. He does, however, note that *ain't* is "an inelegant word" and that it would be a "blessing" if the word is no longer used in English (Bechtel 1901: 119–120). Despite the reference to the etymology of *ain't*, Bechtel's treatment of the feature is decidedly prescriptive, as are those found in usage guides published in the twenty years following his publication. In addition to Bechtel, Fernald (1907: 5) and Vizetelly (1920: 8) also describe *ain't* as "inelegant", and fervently proscribe its use. Krapp (1927: 27–28) is the first usage guide writer to note that even though *ain't* is "low colloquial", "many educated people permit themselves the habit, even though they reprehend it as careless". Krapp (1927: 27–28) further notes that "[o]nly the enforcement of a strong academic authority prevents **ain't** from becoming universal colloquial usage".

The situation with the treatment of *ain't* in usage guides changes even more significantly after the 1950s, when we see a slow turn towards accounting for genre influence on the use of *ain't*, a distinction between the use of *ain't* for *be not* and the use of *ain't* for *have not*, as well as a tendency to present attitudes to the feature indirectly by referring to what is generally thought of the use of *ain't* or what authorities on usage say about it. In other words, there is an increase in "reported opinions" as opposed to *ipse dixit* statements (cf. Busse and Schröder 2009: 80). An example can be found in what seems to be the first relatively balanced account, where Evans and Evans (1957: 23) note that "[t]his word may mean *am not*, *is not*, *are not*, *have not* or *has not*" and that "[i]t is not considered standard in any of these cases, with the possible exception of *am not* in a question, that is, *ain't I?*". They also observe that "a few bold spirits insist on using it because the language needs an expression of this sort" (Evans and Evans 1957: 23).

Entries from the later decades of the twentieth century express a more explicit awareness of the proscription against *ain't*, but this is accompanied by a recognition that the word may be legitimately used in deliberately informal usage (de Mello Vianna et al. 1977: 155). Thus, *ain't* is considered acceptable when used intentionally to achieve certain effects, to be jocular, or to index dialectal speech. When its use is shown to be deliberate or specialised, *ain't* is considered part of the speech of educated speakers. An illustrative example of this is found in Wilson (1993: 22), who notes that "[c]onsciously jocular uses are acceptable, but using *ain't*

in circumstances that do not suggest deliberate choice may brand you as a speaker of Vulgar English". More recent usage guides also note that despite the criticism of *ain't*, the word has a "vital history" (Brians 2003: 6) and it is "still going strong" (Pinker 2014: 204). In general, the entries reveal an incremental increase in degree of sophistication in the treatment of this feature, and a movement towards including a more descriptive account of its usage. This is instantiated by references to various degrees of acceptability and contexts of use, as well as to its status as a usage problem, by discussing the popular attitudes to *ain't*.

5.2.2 The discourse particle *like*

The treatment of the discourse particle *like* in usage guides is an important example of how changes in the language, especially those that become socially stigmatised and are characterised by high and increasing levels of metalinguistic awareness among speakers, are treated in usage guides. *Like* is covered in ten entries, found in ten usage guides in my collection. Figure 5.2 shows that the inclusion of *like* in the group of usage problems is fairly new compared to the other five features, as it was only during the 1970s that it started to be discussed in usage guides.

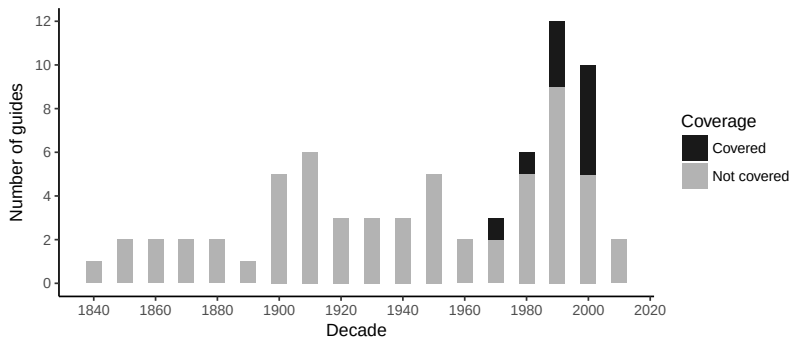

Figure 5.2: Number of usage guides with entries on the discourse particle *like* per decade (n, guides = 70; n, *like* entries = 10)

The first treatment of the discourse particle *like* is found in Shaw (1975: 142), who describes the word as "a filler, a throwaway word used constantly in the speech of many persons, especially young people". Shaw's observation serves as a good illustration of the overall treatment of *like* found in nine other usage guides from the second half of the twentieth century and the beginning of the twenty-first century. Apart from the word "filler", used repeatedly to describe *like*, the word "(bad) habit"

is also used (Brians 2003: 126; Batko 2004: 222), alongside more strongly negative terms, such as "a verbal tic" (Johnson 1991: 361; Garner 1998: 410) or "verbal hiccup" (Brians 2003: 126). *Like* is almost always associated with teenagers or younger speakers. In terms of other groups of people associated with this usage, Brians (2003: 126) observes that the word is "especially associated with hipsters". The use of *like* is mostly associated with uncertainty, poor expression, nervousness, and ignorance.

5.2.3 Non-literal *literally*

Literally is covered in 32 entries in the usage guides consulted. Figure 5.3 shows that *literally* became a usage problem at the beginning of the twentieth century, and has steadily been more frequently covered in usage guides. Commentary on this form first appears in Bierce (1909), which is earlier than the treatment of *literally* found in Strunk (1918), which I initially thought was the earliest mentions of *literally* in an American usage guide (Kostadinova 2015). Apart from Bierce (1909), Strunk (1918), and Krapp (1927), all other entries come from usage guides published in the second half of the twentieth century.

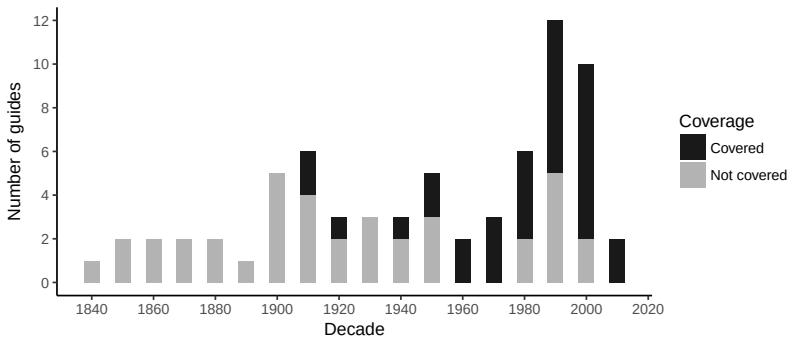

Figure 5.3: Number of usage guides with entries on *literally* per decade (n, guides = 70; n, *literally* entries = 32)

The consensus in the treatment of *literally* is informed by the observed increase in frequency of the intensifying use of the word, which is described as a "misuse" (Johnson 1991: 365) or an "overuse" (Brians 2003: 128). Randall (1988: 2010) sums up the various negative attitudes to *literally* by saying that "such usage may be regarded as loose, superfluous, erroneous, or exasperating". There are two authors whose discussion of *literally* departs from the predominantly negative accounts found in usage guides. The first one is Gilman (1989: 607), who notes that *literally* is "neither

a misuse nor a mistake", but merely a development of a new use of the word. Peters (2004: 326) gives the second more balanced account of *literally*, by pointing out that the word is used much less than is usually assumed, while when it is used it functions grammatically as an intensifier.

5.2.4 Negative concord

Negative concord is covered in 42 of the 70 usage guides consulted. The distribution of the entries across time is given in Figure 5.4. The figure shows that negative concord is consistently covered throughout the period investigated, with roughly half of the usage guides in each decade discussing the feature. Similar to the treatment of *ain't*, all of the nineteenth-century usage guides disapprove of negative concord. These entries tend to be short and concise: the proscription against negative concord is stated simply by using the imperative (e.g. *say x* or *don't say x*), without any arguments or attitudes expressed. Occasionally, the construction is described as "a very common mistake" (Anonymous 1856b: 184) or as "incorrect" (Ballard 1884: 35). Bechtel (1901) presents the first lengthy treatment of negative concord, with eleven subheadings referring to various types of two negatives occurring in the same sentence; he describes negative concord as "inelegant" (1901: 195) and "incorrect" (1901: 196). Similar negative statements are found in usage guides published throughout the twentieth century, and well into the 1990s, as exemplified by Carter and Skates' (1990: 404) pronouncement that "a double negative is redundant and incorrect", and Booher's (1992: 179) remark that there are "no exceptions to this taboo". Similar negative qualifications of negative concord can be found in more recent usage guides as well; an illustrative example is found in Lovinger (2000: 101).

Alongside these negative pronouncements, somewhat less negative discussions of negative concord can also be identified. Hall (1917), for instance, presents the first treatment of negative concord which draws on historical arguments. He notes that negative concord had a prominent place in the history of the language, describing the feature as "out of vogue but not ungrammatical" (Hall 1917: 76–77). He is also the first to qualify the construction as "natural", by noting that "the double negative springs from the desire for emphasis and seems to be natural to human language" (Hall 1917: 77). Hall's pronouncements are thus an important exception compared to other usage guides of the same period. Hall's assessment of the negative concord is reiterated in later usage guides, such as Perrin (1950), Corbin and Perrin (1963), and Ebbitt and Ebbitt (1978). While there are usage guides that describe negative concord in perhaps

less negative terms, use of the construction is consistently advised against. Even those who maintain that it is part of the idiom of English note that it is "widely perceived as a rustic or uneducated form" (Gilman 1989: 365), that it is "inappropriate in spoken and written Standard English" (Wilson 1993: 154), or that it is "socially stigmatized in both American and British English" (Peters 2004: 369). The social stigma against negative concord is a recurrent theme its treatment in usage guides.

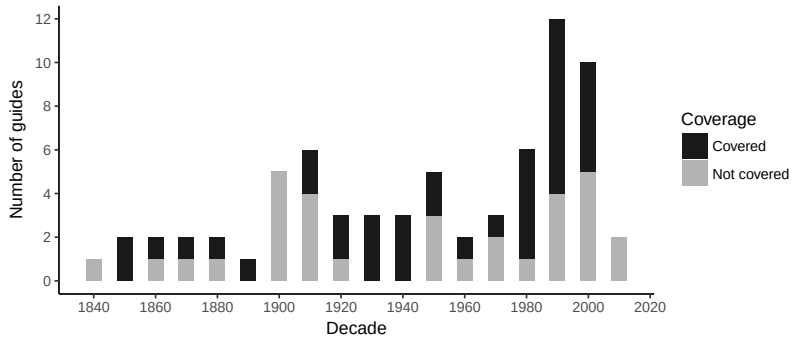

Figure 5.4: Number of usage guides with entries on negative concord per decade (n, guides = 70; n, negative concord entries = 42)

5.2.5 Pronouns in coordinated phrases

The use of subject and object pronouns in coordinated phrases functioning syntactically as subjects or objects is often discussed in usage guides, but there is a marked difference in the extent to which object *I* is covered compared to subject *me*.[2] On the basis of the data analysed, this is undoubtedly influenced by the fact that a particular instance of object *I*, the phrase *between you and I*, is especially salient as a problematic usage feature, and a significant number of the entries identified as entries on object *I* in the usage guides consulted deal explicitly with *between you and I*. With 73 entries on object *I* identified in 47 guides in the collection, this is the most commonly discussed feature among the features included in this study. Figure 5.5

[2] As shown in Table 3.1, object *I*, or *I* for *me*, is covered in 61 of the usage guides included in the HUGE database, while subject *me*, or *me* for *I*, is covered in 49 usage guides. It is important to note that the entries included in the database do not make a distinction between pronoun forms in coordinated phrases and pronoun forms following linking verbs, as in *This is I vs. This is me*. The latter contexts are not taken into account in the present study, so the difference in terms of the number of entries which cover pronoun forms in coordinated phrases functioning as subjects or objects established on the basis of my collection of 70 usage guides is strikingly larger.

shows the number of guides which cover object *I*, out of the total number of guides per decade in the collection. The figure shows that object *I* has been more or less consistently covered in usage guides since the middle of the nineteenth century. In the last two decades of the twentieth century and the first decade of the twenty-first century, the number of usage guides in my collection which discuss this usage feature is fairly high. This suggests that the usage problem status of object *I* has not changed over time in terms of coverage.

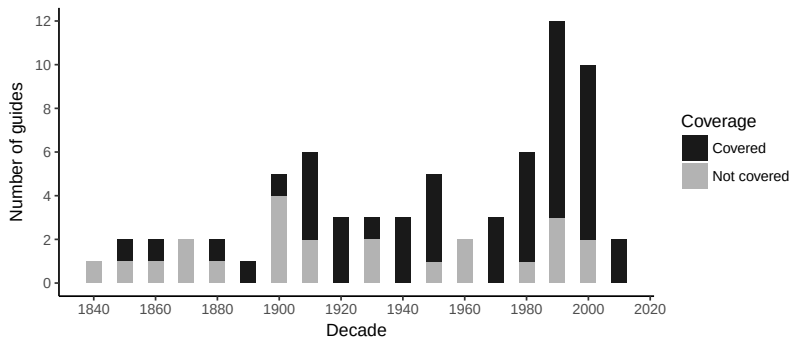

Figure 5.5: Number of usage guides with entries on object *I* per decade (n, guides = 70; n, object *I* entries = 47)

A closer look at the entries, however, reveals that nineteenth-century usage guides are usually both strongly negative in their treatment of object *I*, and very short. While confusion in the use of pronoun case forms tends to be considered ungrammatical and erroneous by the majority of usage guide writers, there are also usage guides in which the nature, or seriousness, of this perceived mistake is questioned. In the spirit of the former view, this variant is described as "as erroneous an expression as any" (Anonymous 1856b: 83), "grossly incorrect" (Utter 1916: 139), a "blunder" (Copperud 1980: 47), and a "gross linguistic [gaffe]" (Garner 1998: 345). Many of the entries also refer to the frequency of occurrence of the object *I* by noting that it is frequently or commonly heard (Perrin 1950: 451; Ebbitt and Ebbitt 1978: 421; Bryson 1984: 27; Pinker 2014: 205). While this variant tends to be associated with carelessness and lack of knowledge of grammar, it is also often observed that educated speakers use the construction; an example is Garner's (1998: 345) observation that "it is perennially surprising how many otherwise educated speakers commit [these gross linguistic gaffes]". An example of the tendency to be more critical of the objections to the use of object *I* is the entry found in Evans and Evans (1957: 60), where it is

noted that even though the expression *between you and I* is not standard, it "has a long and honorable history and has been used by so many great writers that it cannot be classed as a mistaken attempt to speak 'elegant' English". Overall, while the variant is certainly predominantly dismissed as an error, and its usage is advised against, there is an identifiable variation in the treatment, due to considerations of actual language use and language history, as is the case with the entry in Evans and Evans (1957: 60). In fact, a comparison between the oldest entry in the collection, where object *I* is described as a "heinous fault" (Anonymous 1856b: 82), and the latest one, where the author observes that "[w]riters are well advised to avoid *between you and I*, since it makes many readers bristle, but it is not a heinous error" (Pinker 2014: 207), serves to illustrate how the treatment of this variant has changed over time.

Subject *me* is commented on much less than object *I*. As represented in Figure 5.6, subject *me* is only sporadically discussed in usage guides in the second half of the nineteenth century and the first half of the twentieth; the number of usage guides which cover the feature is somewhat higher in the 2000s, with about half of the usage guides analysed discussing it. Most of the entries are found in usage guides published around the turn of the twenty-first century. Discussed explicitly in 19 entries across ten usage guides, subject *me* is described as a "gross error" (Bache 1868: 73), a "vulgar error" (Vizetelly 1920: 109), or a "lapse" (Bryson 1984: 77). In less strong qualifications, the feature is described as simply "incorrect" (de Mello Vianna et al. 1977), or an "error" (Beason and Lester 1996). Johnson (1991: 21) and Gilman (1989: 628) point to its occurrence in children's speech. Peters (2004: 341–342) and Pinker (2014: 97) give the most descriptive account of the use of subject *me*; nevertheless, they still advise against its use in formal edited writing.

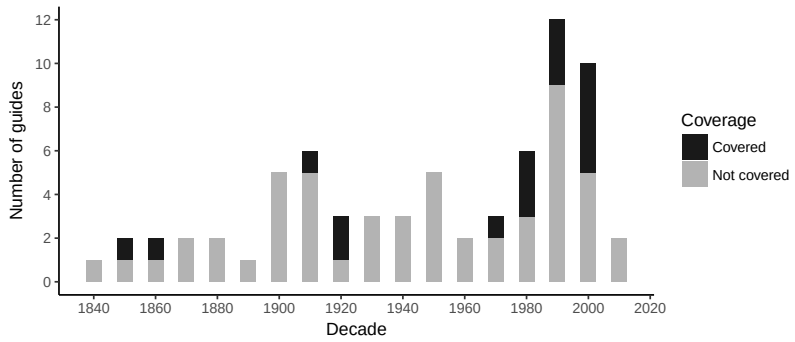

Figure 5.6: Number of usage guides with entries on subject *me* per decade (n, guides = 70; n, subject *me* entries = 10)

5.2.6 The split infinitive

The split infinitive is the second most frequently discussed variant in the collection of usage guides analysed, after object *I*, with 59 entries found in a total of 52 guides in my collection; Figure 5.7 shows the distribution of the entries per decade in the collection of usage guides. The split infinitive is described as early as 1856, and the early treatment of the feature is usually strictly prescriptive. According to the majority of nineteenth-century and early twentieth-century authors, the separation of the particle *to* from the verb is not considered a practice of careful writers and speakers, and should be avoided (Anonymous 1856b; Gould 1867; Bache 1868; Ballard 1884; Bechtel 1901; Fitzgerald 1901). In terms of observations about its frequency of use, Ayres (1911) and Bache (1868) note that examples of the split infinitive are commonly attested, and Ayres (1911: 298) is the first author to note that sometimes a split infinitive may be used by someone due to "some special reason for doing so", which is usually related to the need for emphasis or to considerations of clarity of expression. In other words, Ayres is the earliest example in my collection of usage guides of an author making a distinction between acceptable and unacceptable split infinitives. This distinction becomes more common in usage guides in the course of the twentieth century.

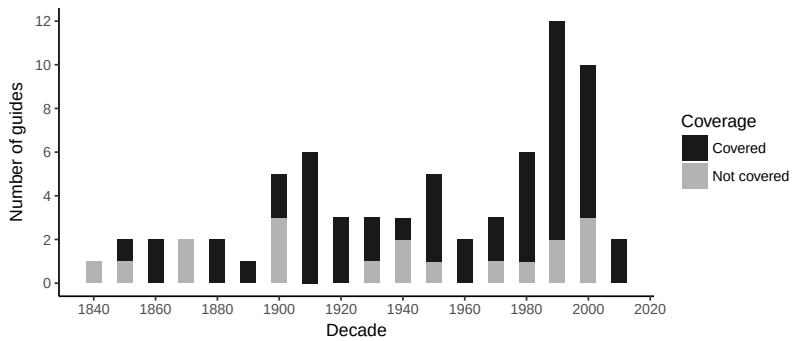

Figure 5.7: Number of usage guides with entries on the split infinitive per decade (n, guides = 70; n, split infinitive entries = 52)

It is also worth noting that the widespread condemnation of the split infinitive and its status as a usage problem is first mentioned by Bierce (1909: 66), who notes that the condemnation was "pretty general", but also fairly recent at the time. In the early twentieth century, Payne (1911: 42) observes that "this much discussed construction seems to be growing in favour, but still it is awkward in most cases of everyday

speech", while Turck Baker (1910: 170) questions the traditional grammarians' censure of the split infinitive by arguing that "there is no reason why [the particle *to*] should not be separate from its verb by the introduction of the adverb, especially when by this position the meaning is more clearly or emphatically expressed". Furthermore, and especially in contrast to the other variants discussed here, split infinitives are more strongly approved of, to the point where the rule against its use is dismissed almost consistently by usage guide writers throughout the twentieth century. For instance, Pinker (2014: 199) calls it a "bogus rule", while Walsh (2004: 64) argues that "more often than not infinitives are better split". The strikingly un-prescriptive treatment of the split infinitive observed here is confirmed in the analyses of treatment and attitudes, to be presented in the next two sections.

5.2.7 Summary

Looking at the results of the distribution of entries in usage guides across decades side by side, and as represented in Figure 5.8, it can be seen that the split infinitive, *ain't*, negative concord, and object *I* are the more frequently treated features, compared to *literally*, subject *me*, and the discourse particle *like*, in terms of both the length of the period in which they are discussed and the number of guides that discuss these features. *Literally* can perhaps be considered to be somewhere in the middle, in that usage guides do not cover *literally* before the beginning of the twentieth century; however, in the second half of the twentieth century, it is covered in the majority of the guides analysed. This attests to its firm place in the usage problem canon today. The split infinitive, *ain't*, negative concord, and object *I* are, in a sense, 'old chestnuts', and consequently, this picture is hardly surprising. What is perhaps more interesting here is the emergence of the coverage of *literally*, which took place around the beginning of the twentieth century, as well as that of the discourse particle *like*, during the 1970s.

The analysis discussed so far in this section reveals two important insights into how features are discussed in usage guides. The first one is that the variation in the treatment of different usage features depends on a number of factors. These factors tend to be associated most significantly with the use of each particular feature, its social connotations, and its place in the history of language advice literature and usage commentary. Negative concord, for instance, is not covered in usage guides as often as *ain't* or the split infinitive, and while some authors do mention the historical development of that feature in English, most of their treatment is focused on the feature's non-standardness. Entries on negative concord rarely refer to its status as

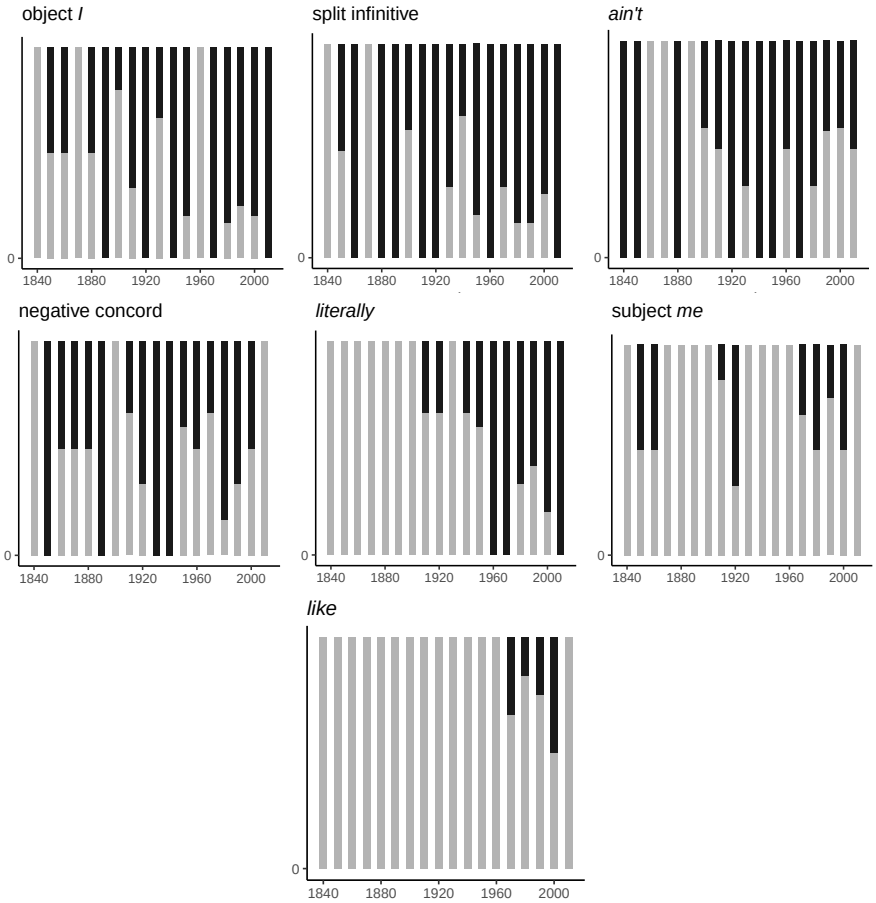

Figure 5.8: Proportions of usage guides with entries on each of the features per decade

a usage problem. This is also the case with more recent usage problems, such as *like* or *literally*. *Ain't* and the split infinitive, on the other hand, are different usage problems, in that the former is considered a non-standard feature, while the latter is considered a stylistic usage problem. Despite this difference, however, the usage problem status of both features is regularly discussed in the entries, resulting in a great deal of prescriptive meta-discourse (cf. Curzan 2014: 48). The second insight is that, despite these differences, we can draw parallels across features in terms of the aspects that are typically discussed in usage guide entries. These aspects include pronouncements on whether the use of a feature is advised or discouraged, whether or not the use of a feature is associated with specific groups of people, whether or not a

feature is a serious grammatical error, and whether or not it is commonly used. These multiple aspects of the discussion of usage features in usage guides are all important in gaining insight into how usage guides treat usage problems, and how this treatment has changed over time. In order to tease out these dimensions, as I mentioned above, I distinguish between treatment, attitudes to usage, and dimensions of usage in the analysis of the entries. Each of these levels of analysis is discussed in the sections that follow.

5.3 Treatment of the language features in usage guides

In the previous section I discussed how the coverage of different features in usage guides varies in degree and across time. Different patterns of coverage can show how usage guides in certain cases continue promote long-established usage problems, such as *ain't*, while in other cases usage guides appear to respond to, or engage with, processes of language variation and change, as in the case of *like*. These patterns, however, tell only one part of the story. In addition to the variation in metalinguistic commentary in terms of the coverage and frequency of discussion of usage features, usage guides vary in how they treat usage features, which is another aspect that is critical to understanding how prescriptive ideas about language features change over time. In order to analyse changes in the treatment, as already explained in Section 4.3, I distinguish three mutually exclusive categories of treatment: whether the usage feature is approved of or accepted (ACCEPTABLE), whether its use is neither explicitly accepted nor explicitly criticised, or accepted only in specific registers of use (RESTRICTED), or whether its use is explicitly dismissed (UNACCEPTABLE). Each entry was classified into one of these three categories (cf. Section 4.3).

The results from the classification are shown in Figure 5.9, which shows the number of each category of entries (ACCEPTABLE, RESTRICTED, or UNACCEPTABLE) out of the total number of entries per feature. This allows for a comparison of the various kinds of treatment, and an analysis of how they differ across language features regardless of the fact that some of them occur in a smaller number of entries than others. As the figure shows, *literally*, negative concord, *like* and subject *me* only have RESTRICTED or UNACCEPTABLE entries, with the latter being proportionally more frequent. This is especially the case with *literally*, *like*, and subject *me*; for negative concord, the proportion is somewhat more even. Entries on *ain't* and object *I* are split between UNACCEPTABLE and RESTRICTED, with a small proportion of entries being

classified as ACCEPTABLE. Object *I* is perhaps treated somewhat less negatively than *ain't*, in that a larger proportion of the entries were classified as RESTRICTED rather than UNACCEPTABLE. Finally, the split infinitive is the feature for which there was a fairly even distribution of entries across the three categories. This shows that the split infinitive was the most positively treated feature, followed by *ain't* and object *I*.

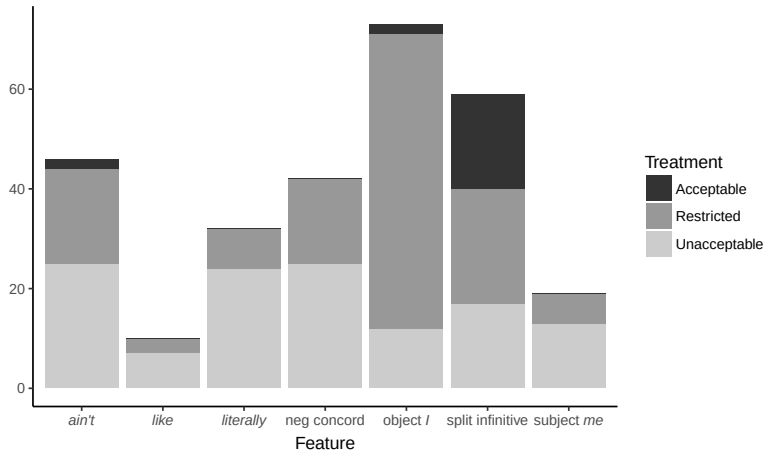

Figure 5.9: Number of entries across treatment categories per feature

While Figure 5.9 shows the general differences in treatment across language features, it does not provide any clues as to possible changes over time. Looking for such patterns across time provides additional evidence of possible changes in precept. In order to identify changing patterns, as well as the statistical significance of these changes, logistic regression modelling was used to investigate how time predicts the likelihood of an entry being a particular category of treatment as opposed to another; this was done using the `mlogit` R package (Croissant 2013), and following the procedures outlined in Levshina (2015: Chapter 12 for binomial regression, and Chapter 13 for multinomial regression). In this kind of modelling, the categories of treatment, i.e. ACCEPTABLE, RESTRICTED, and UNACCEPTABLE, are the three levels of the categorical outcome variable TREATMENT, while YEAR is the predictor variable, which is a continuous variable. Each entry was thus classified as one of the three levels of TREATMENT; the value for YEAR for each entry was the year in which the usage guide was published. Using logistic regression modelling, the likelihood of each of the three categories of treatment is modelled as a function of the year of publication of the usage guide in which the entry was found. This kind

of modelling provides insights into how the probabilities of different categories of treatment have changed over time, and whether this change is significant. I used the multinomial logistic regression technique for the features for which all three categories of TREATMENT were identified, while those features for which only two categories of TREATMENT were present were analysed using binomial logistic regression models. The difference between the two techniques is in the number of levels in the outcome variables. Both the binomial and the multinomial logistic regression techniques allow us to estimate the chances that one of two or three possible outcomes will occur, in this case the three treatment categories, given a specific value of the predictor, in this case the year of publication. In the case of usage features in which only two categories of treatment were identified in the data, the binomial logistic regression model was used for this reason. In what follows, I discuss in detail the results from these tests for each feature in turn.

5.3.1 *Ain't*

Figure 5.10 shows that *ain't* was not found acceptable in usage guides until the twenty-first century. The nineteenth-century usage guides all disapprove of the item, and this situation remains the same in the first half of the twentieth century. The real change happened during the second half of the twentieth century, as evidenced by the number of RESTRICTED entries on *ain't*. Figure 5.10 shows the predicted probabilities for the three categories of treatment, based on the categorisation of the entries described above. After the 1950s, the probability for an entry being UNACCEPTABLE is lower than 50%, while the likelihood that an entry is RESTRICTED increases. Around the turn of the twenty-first century, there is yet another change, with the probability that an entry is ACCEPTABLE increasing, while the probability that an entry is RESTRICTED decreases slightly. The coefficients in a multinomial logistic regression model which compares the odds that an entry is ACCEPTABLE as opposed to UNACCEPTABLE, or the odds that an entry is RESTRICTED as opposed to UNACCEPTABLE, given in Table 5.1, show that the increase in the likelihood that an entry is ACCEPTABLE is not significant, while the increase in the likelihood that an entry is RESTRICTED has increased significantly. In other words, for every year, the odds of an entry on *ain't* being RESTRICTED increase by 1.05. This suggests that while the treatment of *ain't* has indeed changed during the twentieth century, this change does not mean that *ain't* is now considered acceptable; rather, it has become acceptable in restricted contexts, as evidenced by the increased likelihood that an entry

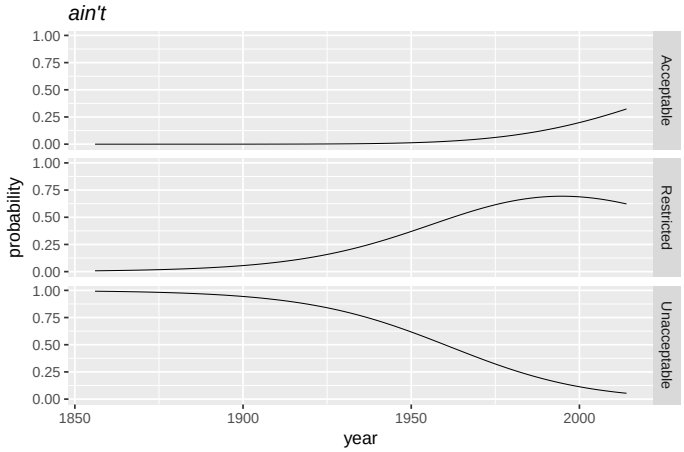

Figure 5.10: Probabilities of the occurrence of categories of entries on *ain't* as predicted by a multinomial logistic regression model

is RESTRICTED in the usage guides analysed.

	b	OR	Std. errors	*t*-value	*p*-value
Acceptable (intercept)	−1532.842	0	1418.500	−1.081	0.2799
Restricted (intercept)	−96.569	1.15e-42	28.890	−3.343	0.0008***
Acceptable:year	0.765	2.150	0.708	1.082	0.2794
Restricted:year	0.049	1.051	0.015	3.343	0.0008***

Table 5.1: Multinomial logistic regression model for the three categories of treatment of entries on *ain't*; reference level is UNACCEPTABLE

5.3.2 The discourse particle *like*

Entries on *like* are split between RESTRICTED and UNACCEPTABLE, with an interesting change happening in the twenty-first century. The entries found in usage guides published in the second half of the twentieth century are not explicitly UNACCEPTABLE, while those in guides published in the first decade of the twenty-first century tend to be more explicitly negative about this feature. Figure 5.11 plots the likelihood of UNACCEPTABLE entries as opposed to RESTRICTED. As was the case with subject *me*, it is impossible to make observations with any certainty on the basis

of these data, because the confidence intervals are too large, as a result of paucity of data. The slight increase in the probability of an entry being UNACCEPTABLE rather than RESTRICTED is not statistically significant. A parallel can thus be drawn between the usage guide treatment of *like* on the one hand, and the increase in negative social evaluation of the item (addressed in more detail in Section 7.3) among native speakers on the other; both of these may be seen as a reaction to the increase in its frequency of use.

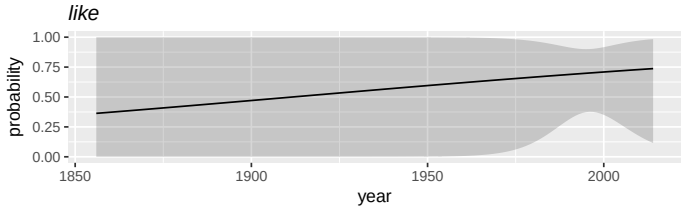

Figure 5.11: Probabilities of the occurrence of categories of entries on the discourse particle *like* as predicted by a binomial logistic regression model; reference level is RESTRICTED, predicted level is UNACCEPTABLE

5.3.3 Non-literal *literally*

Literally starts to be treated as a usage problem in guides at the beginning of the twentieth century, when the majority of the entries are UNACCEPTABLE. Some authors of guides published later in the twentieth, or in the twenty-first century, are vague in their pronouncements, and refrain from expressing an explicit stance towards *literally*, beyond the observation that *literally* is sometimes used as an intensifier. However, these guides are a minority in comparison to those that treat this feature as UNACCEPTABLE. As Figure 5.12 shows, there is a slight decrease in the probability of an entry being UNACCEPTABLE, as opposed to RESTRICTED, in the course of time. This decrease is, however, not statistically significant. This puts *literally* in the same category as *like* and object *I*/subject *me*, as a strongly dispreferred feature.

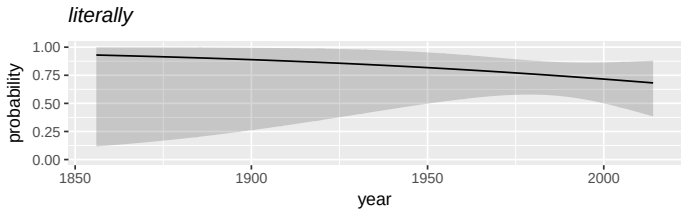

Figure 5.12: Probabilities of the occurrence of categories of entries on the discourse particle *literally* as predicted by a binomial logistic regression model; reference level is RESTRICTED, predicted level is UNACCEPTABLE

5.3.4 Negative concord

Negative concord is considered unacceptable overall, and this is highlighted by the fact that there were no ACCEPTABLE entries. However, the number of RESTRICTED entries, i.e. entries that specify some kind of acceptability in certain contexts or that are vague about the acceptability of the feature, has grew during the second half of the twentieth century and at the beginning of the twenty-first. This change in the probability for the two categories of treatment identified in the entries in negative concord is shown in Figure 5.13. A binomial logistic regression model of the two categories of entries, *viz.* RESTRICTED and UNACCEPTABLE, which predicts the odds of an entry being UNACCEPTABLE as opposed to RESTRICTED as a function of the year in which it was published, is given in Table 5.2. The model shows that over time the odds of an entry being UNACCEPTABLE decrease (b = -0.023); according to the odds ratio, for each one-unit increase in time, i.e. for one year, the odds of an entry being UNACCEPTABLE decrease by 1.023. This change is significant ($p = 0.014$).

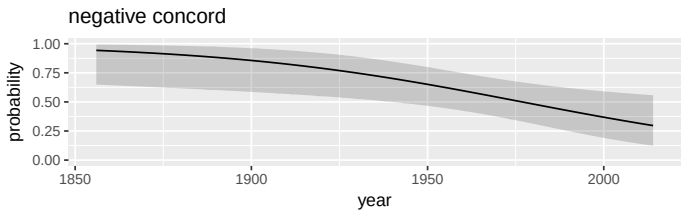

Figure 5.13: Probabilities of the occurrence of categories of entries on negative concord as predicted by a binomial logistic regression model; reference level is RESTRICTED, predicted level is UNACCEPTABLE

	b	OR	Std. errors	*z*-value	*p*-value
Intercept	−45.913		18.70	−2.455	0.014*
year	−0.023	1.023	0.010	0.015	0.014*

Table 5.2: Binomial logistic regression of the two categories of treatment in entries on negative concord; reference level is RESTRICTED, predicted level is UNACCEPTABLE

5.3.5 Pronouns in coordinated phrases

Entries on object *I* are predominantly RESTRICTED, as shown in Figure 5.9 above. The large number of RESTRICTED entries is the result of the acceptance of this feature only in restricted contexts, such as informal, colloquial, or familiar settings, but its use in formal and written contexts is decidedly proscribed. For instance, in discussing the phrase *between you and I*, Pickett et al. (2005: 380) note that, while variation in pronoun case forms has long been characteristic of English usage, the proscription of the use of nominative pronominal forms after prepositions was first put in writing around the middle of the nineteenth century. While guides from that period acknowledge that "the phrase occurs quite often in speech", they also warn the reader that it is "widely regarded as a sign of ignorance", and point out that formal writing requires *between you and me*. Evans and Evans (1957: 60) are the authors whose treatment of object *I* is more positive than the other usage guides analysed, which, as already mentioned in Section 5.2.5, do not consider this feature as a mistake, due to its history of use. Looking at the change in treatment over time, the change in the probability of the three categories of treatment is plotted in Figure 5.14. The figure plots the change in probability in the categories of TREATMENT analysed. In terms of change over time, Figure 5.14 shows that there is a very slight change in the probability of an entry being RESTRICTED, which is about 25% during the second half of the twentieth century. The multinomial logistic regression model reveals that these differences are not significant, suggesting that the treatment of this feature has not changed over time, and has remained RESTRICTED.

The use of *me* in subject position is similarly disapproved of, with 13 out of the 19 entries being analysed as UNACCEPTABLE (Figure 5.9). Since only two categories of treatment were identified, the change over time was modelled using binomial logistic regression, which is shown in Figure 5.15. Since this feature is one of the less frequently discussed, the figure shows that there is little certainty as to which trends

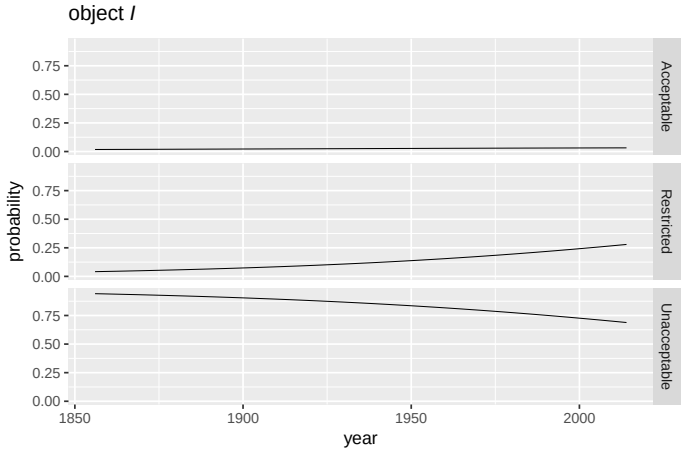

Figure 5.14: Probabilities of the occurrence of categories of entries on object *I* as predicted by a multinomial logistic regression model

can be identified, due to the small number of data points and the large confidence intervals. A trend with smaller confidence intervals can be identified for the latter half of the twentieth century, but not earlier. This change is not statistically significant, which means that on the basis of these results, it is difficult to identify any change in the treatment of subject *me*. Since the results of the model are not significant, I can only conclude that subject *me* is treated as an unacceptable usage feature in usage guides across the twentieth century on the basis of the majority of UNACCEPTABLE entries, as well as the fact that no ACCEPTABLE entries were identified for this feature.

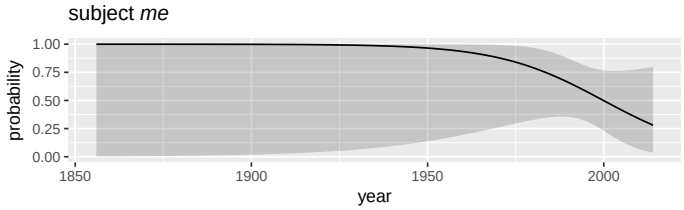

Figure 5.15: Probabilities of the occurrence of categories of entries on subject *me* as predicted by a binomial logistic regression model; reference level is UNACCEPTABLE

5.3.6 The split infinitive

The treatment of the split infinitive provides the most striking exception to the patterns observed for the usage features discussed so far. Figure 5.9 above shows that entries on the split infinitive include the highest proportion of ACCEPTABLE entries compared to the rest of the features analysed, and the lowest proportion of UNACCEPTABLE entries. The change in treatment itself is also significant, statistically and generally speaking. The contrast is quite substantial, with nineteenth-century usage guides decidedly disapproving of the split infinitive, while the treatment of the feature in twenty-first-century usage guides is predominantly ACCEPTABLE. Figure 5.16 and Table 5.3 show the change in the probability of the three categories of treatment.

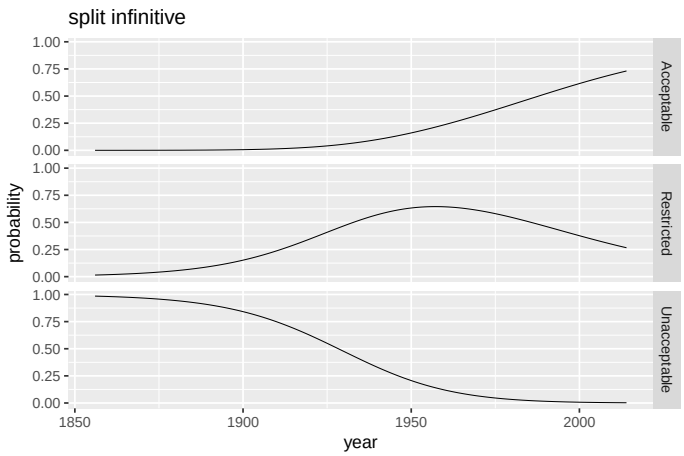

Figure 5.16: Probabilities of the occurrence of categories of entries on the split infinitive as predicted by a multinomial logistic regression model

	b	OR	Std. errors	*t*-value	*p*-value
Acceptable:(intercept)	−183.395	2.25e-80	45.294	−4.049	5.15e-05***
Restricted:(intercept)	−109.667	2.36e-48	32.800	−3.344	0.001***
Acceptable:year	0.094	1.098	0.023	4.055	0.000***
Restricted:year	0.057	1.058	0.017	3.345	0.001***

Table 5.3: Multinomial logistic regression of the three categories of treatment of entries on the split infinitive; reference level is UNACCEPTABLE

5.3.7 Summary

In summary, what these trends in change in the treatment of the various features show is that usage guides vary in their judgements across time and across language features. For some features, usage guides appear quite accepting of changes in the language, while for others the judgements remain conservative over time. In the case of conservative treatment, it is interesting to note that there is rarely universal approval or disapproval of a feature. The variation that exists between the categories UNACCEPTABLE and RESTRICTED may be due to variation within the genre of usage guides. To draw on Algeo's (1991b) typology, which distinguishes between subjective and objective usage guides, subjective usage guides tend towards more strongly expressed negative treatment of a feature, while more objective usage guides fall in the category of RESTRICTED precisely because they refrain from overt negative value judgements. The results observed in the current analysis provide further evidence in support of this typology.

5.4 Expressions of attitudes to usage in usage guides

In this section, I discuss the results of the second level of analysis, i.e. the occurrence of POSITIVE and NEGATIVE attitude expressions in the entries. The number of such expressions is taken as an indication of how usage features are predominantly discussed, but also of how the discussion has changed over time. In addition to the quantitative analysis of the number of occurrences of POSITIVE and NEGATIVE attitudes in the treatment of the various features, I will also discuss other important dimensions of attitudes observed in the data, on the basis of a qualitative analysis. It is important to keep in mind, as I already pointed out in Sections 2.2 and 4.3 above, that these attitudes have to be distinguished from speakers' attitudes.

The frequency of occurrence of attitude expressions differs across the entries, depending on the language feature and the time period with which the entries are associated. These differences serve as further indicators of how the treatment of different features varies across time. Figure 5.17 shows the number of attitudes in total identified for each feature and each entry in the collection of entries analysed. It also shows the number of POSITIVE attitudes compared to that of NEGATIVE attitudes. The subfigures in Figure 5.17 are ordered from the split infinitive, the feature with the highest number of POSITIVE attitude expressions, to subject *me*, a feature for which

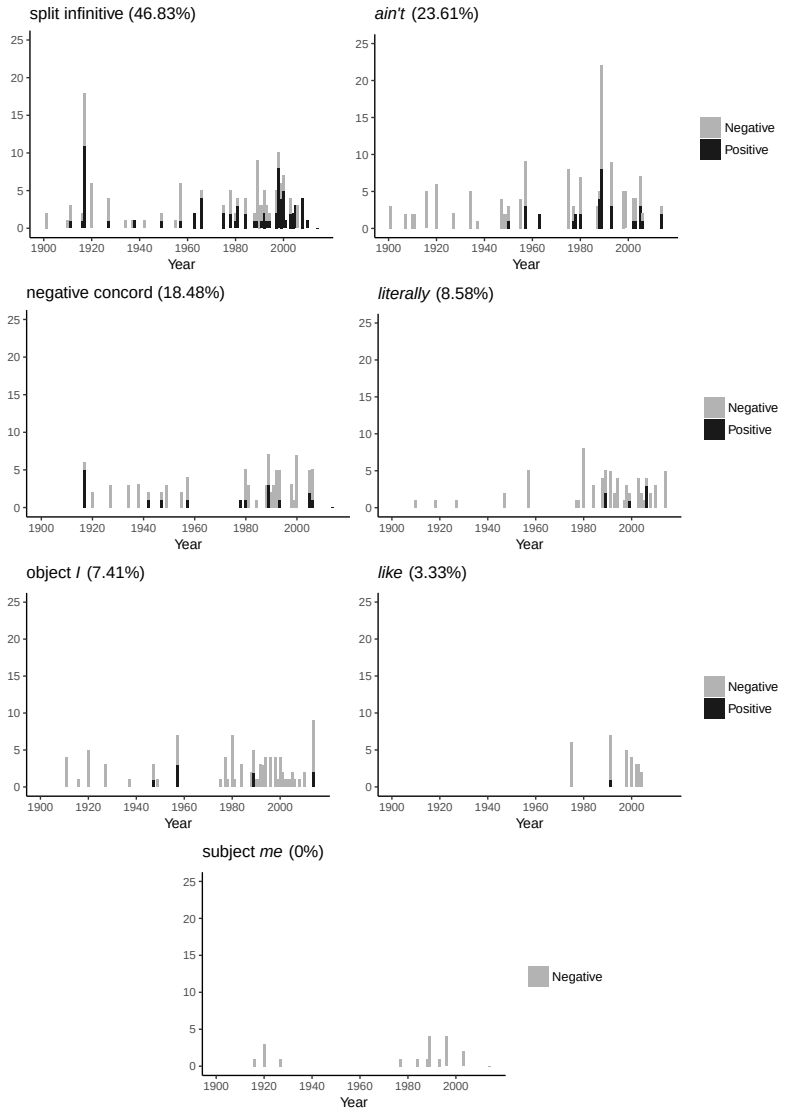

Figure 5.17: Number of POSITIVE and NEGATIVE attitude expressions per feature across time

no POSITIVE attitudes were identified in the sample of entries. The split infinitive can be clearly distinguished from the other features by the number of attitudes expressed explicitly, and, particularly, by the number of POSITIVE attitudes expressed. This pattern is in line with what was observed in the classification of the various types

of treatment according to which the split infinitive is the most accepted feature. There were 84 NEGATIVE attitude expressions, and 74 POSITIVE attitude expressions (46.83%) in the entries on the split infinitive. The attitudes expressed explicitly towards *ain't* were POSITIVE in 23.61% of the cases, i.e. a POSITIVE attitude was expressed 34 times, as opposed to 110 expressions of NEGATIVE attitudes. The treatment of negative concord is also marked by a high number of NEGATIVE attitudes expressed in the entries. There are 75 expressions of NEGATIVE attitudes, as opposed to 17 expressions of POSITIVE attitudes towards this feature, which is 18.48% of the cases. Attitudes to *literally* were POSITIVE in six cases, and NEGATIVE in 64, which translates to 8.58% of POSITIVE attitudes. Attitudes to object *I* are predominantly NEGATIVE, with 100 NEGATIVE attitude expressions as opposed to only eight expressions of POSITIVE attitudes, which is 7.41% of the total number of attitude expressions identified in the entries on object *I*. Attitudes to *like* are also predominantly NEGATIVE, with 29 NEGATIVE expressions and 1 POSITIVE expression, or 3.33%. Finally, all 21 attitudes expressed towards subject *me* were NEGATIVE. Figure 5.17 also shows that, with some exceptions, POSITIVE attitudes are characteristically found in guides published after the 1950s.

These trends provide some indication as to how the attitudes expressed differ across features and change over time, but they do not provide more specific information regarding the kinds of attitudes expressed. As Yáñez-Bouza (2015) has shown, in studies of precept it is important not only to look at general changes, but also to account for more specific types of attitudes expressed. In other words, the ways in which the acceptability of a construction is expressed may differ with respect to linguistic features, usage guides, or time periods. Understanding these nuances provides a more comprehensive understanding of how precept in general, and attitudes in particular, have changed over time. In order to explore this question in more detail, I will now turn to the actual expressions identified in the entries for each feature. I will discuss each feature in the order of proportion of POSITIVE attitudes, starting with the split infinitive. Tables containing all the annotations for expressions of attitudes to each usage feature can be found in Appendix C.

5.4.1 *Ain't*

Similarly to the split infinitive, it is possible to identify a number of recurring themes, or dimensions, in the expressions of POSITIVE or NEGATIVE attitudes towards *ain't* (see Table D.1 in Appendix C). Two of these dimensions can be identified in the

POSITIVE attitudes expressed: references to the naturalness or usefulness of the contraction in the English language, and the appropriateness of *ain't* in signalling casual or colloquial speech. Such observations are found in Corbin and Perrin (1963), Randall (1988), Parrish (2002), and Pinker (2014). The other dimension found in the POSITIVE attitudes to *ain't* refers to its "down-to-earth" quality (Ebbitt and Ebbitt 1978; Brians 2003). An interesting observation about the POSITIVE attitudes towards *ain't* is that the majority appear in usage guides after 1950. This may indicate that a change in treatment has taken place, from a more uniformly negative to a more balanced treatment.

NEGATIVE attitudes are far more common, as shown in Figure 5.17, and there are three themes that can be distinguished here. The first is the perceived grammatical incorrectness of *ain't* (Turck Baker 1910; Vizetelly 1920; Gilman 1989; Mager et al. 1993; Brians 2003). The second theme is the association of *ain't* with "illiterate" speech (Wood and Stratton 1934; Partridge 1947; Witherspoon 1948; Nicholson 1957; Morris and Morris 1975; Copperud 1980; Shaw 1975; Allen 1999; Parrish 2002) or "uneducated" speech (Morris and Morris 1975; Brians 2003). In addition to this social meaning of *ain't*, the feature is also associated with non-standard or dialectal speech. "Vulgar", and related forms, such as "vulgarism", are also often found (Hurd 1847; Ayres 1911; Utter 1916; Vizetelly 1920; Wood and Stratton 1934; Treble and Vallins 1937; Stratton 1949; Perrin 1950; Perrin and Smith 1955; Gilman 1989; Wilson 1993; Pickett et al. 2005). These strongly negative social connotations are most common in the entries on *ain't* as opposed to the other features analysed. The third dimension identified in the attitudes to *ain't* is its status as a usage problem, which suggests its strong presence in prescriptivist discourse. This is exemplified by qualifications such as "bugbear", "condemned", "displeases many people", "frowned upon", and "stigmatised", which reflect what Curzan (2014: 48) calls "prescriptive meta-discourses, or conversations about the prescriptive conversations about language". In other words, what many of these entries contain are not only discussions of the use of *ain't*, but also of the history of the proscription, as well as other sources discussing the feature. Finally, the debate surrounding the status of *ain't* as a word (cf. Curzan 2014) is reflected in the presence of qualifications of *ain't* both as "a word" (Wilson 1993) and as a "nonword" (Garner 1998). This final dimension is interestingly reflected in the study of speakers' attitudes reported in Chapter 7.

5.4.2 The discourse particle *like*

Like is one of the features for which no POSITIVE attitudes were identified in the entries (see Table D.2 in Appendix D). The striking characteristic of the NEGATIVE attitudes to *like*, which sets this usage problem apart from the rest, is the description of *like* not as an error or a mistake, but as a "filler" (Shaw 1975; Garner 1998; Parrish 2002), a "verbal crutch" (Lovinger 2000), "verbal tic" (Johnson 1991; Garner 1998), or "verbal hiccup" (Brians 2003: 126). Furthermore, the use of the discourse particle *like* is seen as a "habit" (Brians 2003; Batko 2004) that "infests every sentence" (Johnson 1991) and needs to be eradicated. The third dimension that is characteristic of the attitudes towards *like* is its strong association with negative personal characteristics such as nervousness (Shaw 1975) or indecisiveness (Johnson 1991). The case of *like*, as already mentioned above, serves as a good illustration of how usage guide authors address recent changes in the language, as well as of the evaluative connotations they use in their treatment of this relatively recent addition to the usage problem canon, such as age and personal characteristics of those who use new forms.

5.4.3 Non-literal *literally*

POSITIVE attitudes to *literally* relate to the variation in its use; such an attitude is expressed through the observation that the intensifier use of *literally* is "neither a misuse nor a mistake" (Gilman 1989: 608) but a development natural to language (see Table D.3 in Appendix D). The only other POSITIVE attitude to *literally*, which was identified in the usage guide entries analysed, comes from Peters (2004: 326), who notes that it "lends impact to quantitative statements" and "invites readers to savour the aptness of the writer's terms of reference". Both Gilman and Peters are descriptive usage guides, and their observations of usage are based on solid empirical evidence. These POSITIVE attitudes cannot therefore be considered to be indicative of a change in attitudes to *literally* in usage guides in general, but merely represent variation within the usage guide genre (cf. Algeo 1991b). The majority of the attitudes to non-literal *literally* are, however, NEGATIVE, and these relate most often to its supposed incorrectness (Strunk 1918; Krapp 1927; Randall 1988; Gilman 1989; Batko 2004), as well as its lack of (logical) meaning (Evans and Evans 1957; Ebbitt and Ebbitt 1978; Johnson 1991; Brians 2003; Pinker 2014). The third dimension discernible in the NEGATIVE attitudes towards *literally* relates to its superfluousness (Copperud 1980; Randall 1988; Johnson 1991; Wilson 1993). Finally, *literally* is seen

as being overused (Sutcliffe 1994; Brians 2003).

5.4.4 Negative concord

POSITIVE attitudes to negative concord are found far less often than NEGATIVE ones (see Table D.4 in Appendix D). The predominant theme which can be identified in these attitude expressions is the naturalness of the feature (Hall 1917) and its place in the English idiom (Ebbitt and Ebbitt 1978). A smaller group of expressions centre on the feature's function in strengthening or emphasising a negative statement. More explicitly, a number of entries explain that the use of negative concord is "not a backsliding from the idiom of more formal English" (Ebbitt and Ebbitt 1978: 483–484), but a "natural" and "normal way of strengthening a negative" (Evans and Evans 1957: 143–144). Speakers are even advised that they do not need to eradicate the feature if it is part of their everyday language use. Negative concord is also seen as emphatic (Hall 1917), as "powerful" (Wilson 1993: 154), and as "an effective construction in writing dialogue or striking a folksy note" (Pickett et al. 2005: 148–149). NEGATIVE attitudes to negative concord are found mostly to refer to the alleged grammatical incorrectness of the feature, and to its inappropriateness of use. The attitude of grammatical incorrectness of negative concord is expressed by labels such as "mistake" (Anonymous 1856b), "error" (Meredith 1872; Stratton 1949; Johnson 1991; Lovinger 2000), and "incorrect" (Ballard 1884; Turck Baker 1910; Carter et al. 1990; Booher 1992; Mager et al. 1993; Pickett et al. 2005). The second dimension identifiable in the NEGATIVE attitudes expressed towards negative concord is the social stigma associated with its use. The feature is seen as "an immediate indication that the speaker's or writer's diction is substandard" (Johnson 1991: 332), while Evans and Evans (1957: 143–144) observe that "no one who values public opinion can afford to" use negative concord.

5.4.5 Pronouns in coordinated phrases

Pronouns in coordinated phrases are also treated in a predominantly negative way, with a very small number of POSITIVE attitudes expressed (see Tables D.5 and D.6 in Appendix D). These POSITIVE expressions are found only for object *I*, while all the attitudes expressed towards subject *me* are NEGATIVE. Object *I* is seen positively as a historically sound construction which is part of spoken English (Evans and Evans 1957; Gilman 1989). NEGATIVE attitudes to object *I* focus almost entirely on its grammatical incorrectness. Classifications of this feature as "incorrect", "wrong", an

"error", a "mistake", or a "fault", are found in the great majority of the entries analysed (Anonymous 1856b; Bache 1868; Ayres 1911; Payne 1911; Vizetelly 1920; Hadida 1927; Treble and Vallins 1937; Stratton 1949; de Mello Vianna et al. 1977; Copperud 1980; Bryson 1984; Randall 1988; Gilman 1989; Carter et al. 1990; Johnson 1991; Booher 1992; Wilson 1993; Sutcliffe 1994; Beason and Lester 1996; Garner 1998; Lovinger 2000; Batko 2004; Clark 2010; Pinker 2014). A small number of entries also refer to the negative social associations of object *I* by labelling it "illiterate" (Nicholson 1957: 55), "half-educated" (Ebbitt and Ebbitt 1978: 421), or as being "regarded as a sign of ignorance" (Pickett et al. 2005: 380). The most common NEGATIVE attitude towards subject *me* is related to its grammatical incorrectness.

5.4.6 The split infinitive

The split infinitive is the feature for which the largest number of POSITIVE attitudes were expressed (see Table D.7 in Appendix D). There are four dimensions observable in these attitudes: grammaticality, stylistic value, social value, and encouragement to use split infinitives. The first dimension can be identified in the recurrent emphasis on the fact that split infinitives are not grammatically incorrect and do not violate grammatical rules. Illustrative cases can be found in Compton (1898: 39–40), who notes that split infinitives are not "not a violation of any rule of grammar"; Gilman (1989: 867–868), who notes that "there is nothing grammatically wrong" with split infinitives; and Allen (1999: 548), who observes that the split infinitive is not "a grammatical blunder". The second dimension refers to the positive stylistic values associated with the use of split infinitives. These kinds of expressions stress that the split infinitive is "natural" (Hall 1917; Nicholson 1957; Bryson 1984; Mager et al. 1993; Walsh 2004), and that it improves the clarity of expression (Hall 1917; Krapp 1927; Ebbitt and Ebbitt 1978; Vermes 1981; Bryson 1984; Lovinger 2000). The third dimension in the POSITIVE attitudes expressed towards the split infinitive relates to the fact that it is not seen as indexical of particular negative social aspects of its users. These include observations that "the use of the split infinitive does not necessarily put us among the illiterates, ignoramuses, and violators of English undefiled" (Hall 1917: 96), or that it is "acceptable" and "perfectly proper" (Garner 1998: 616). Finally, the last characteristic of the POSITIVE attitudes expressed towards the split infinitive relates to calls for using split infinitives freely, which can be found often in the entries. Fogarty (2008: 11), for instance, notes that "it's OK to split infinitives", and advises her readers not to "let anyone tell you that it's forbidden". These attitudes to the split

infinitive show that not only is the feature presented as being acceptable, it is also considered acceptable on multiple levels, both linguistic and social. In other words, split infinitives are accepted in formal speech and writing, especially if they are seen as an emphatic or effective way to express a thought. The stress on the stylistic acceptability of the split infinitive is in line with Curzan's (2014) account of different types of usage problems, or, more specifically with identifying split infinitives as usage features considered problematic on the stylistic level. Since many of the usage guide writers take issue with the prescriptivist discourse prohibiting split infinitives, their attitudes can be seen as reactions to the notion that split infinitives are stylistically unacceptable.

Nevertheless, only about half of the total number of attitudes expressed are POSITIVE; to be precise, 46.83% of the attitude expressions found in entries on the split infinitive are positive (Figure 5.17). The rest of these expressions are NEGATIVE, and in line with what one might expect to find in prescriptive language advice literature. These NEGATIVE attitudes can be seen as directly contradicting the four dimensions of POSITIVE attitudes I mentioned above. For instance, while, according to some authors, split infinitives do not violate a grammatical rule (Compton 1898), others assert that splitting infinitives violates the rule of never inserting an adverb between *to* and the verb (Anonymous 1856b). Grammatically speaking, split infinitives are also seen as mistakes, errors, or faults (Bache 1868; Bechtel 1901; Ayres 1911; Stratton 1949; Follett 1966; Bryson 1984; Allen 1999). In terms of stylistic aspects, NEGATIVE attitudes associated with split infinitives are awkwardness (Payne 1911; Hall 1917; Partridge 1942; Ebbitt and Ebbitt 1978) or clumsiness of expression (Hall 1917; Wilson 1993), as well as sloppiness (Johnson 1991). Split infinitives are associated with ignorance (Johnson 1991) or carelessness (Garner 1998). Finally, on the basis of these NEGATIVE attitudes, in many of the entries readers are advised to avoid using split infinitives.

5.4.7 Summary

In summary, the split infinitive is the most positively treated feature of the six features studied here, followed by *ain't*. The analysis of attitudes across time, however, also shows that POSITIVE attitudes towards the split infinitive are found in the earliest usage guides, which might indicate that the split infinitive was perhaps never a strongly dispreferred feature (cf. Albakry 2007) in comparison to other features. For example, *ain't* is described positively only in usage guides published after 1950. This

indicates that there has been a change in the treatment of *ain't* in the middle of the twentieth century. Such a change is not observed for the other features, the treatment of which remains mainly negative. Furthermore, there is also variation in the indexical nature of the language features. *Ain't* is very strongly associated with negative social characteristics, while object *I* is predominantly seen as an error, and relatively few usage guides mention its association with "illiterate" speech. The other features are all referred to fairly negatively, with very few expressions of POSITIVE attitudes. These NEGATIVE attitudes are expressed with a range of typical prescriptive labels (cf. Sundby et al. 1991; Yáñez-Bouza 2015), the most frequent of which are "illiterate", "inelegant", "uneducated", "ungrammatical", "vulgar", and "vulgarism". This is in turn indicative of the strong prescriptive influence still observable in the genre of usage guides.

Despite the differences in the numbers of POSITIVE and NEGATIVE attitudes expressed towards each of the features analysed, it is noteworthy that the overall number of attitudes in general has increased over time. This is important because it shows that the treatment has changed over time. The entries become incrementally more sophisticated and, especially in the case of so-called shibboleths of usage, such as *ain't* or the split infinitive, there is a pronounced tendency to include both sides of the usage debate in the discussion, resulting in the expression of both POSITIVE and NEGATIVE attitudes towards the feature. In other words, although there is generally a predominance of NEGATIVE attitudes expressed in the entries taken together, the identification of POSITIVE attitudes is important in pointing to two aspects of the way in which these usage features are treated. The first aspect that the presence of POSITIVE attitudes points to is a change in the treatment of some of the features. The second aspect is the variation in treatment observed in the entries, which may suggest a development in the usage guide genre itself.

5.5 Dimensions of usage

Alongside expressions of attitudes to the usage features, the entries also contain various types of references to dimensions of usage; i.e. any aspects of the usage of a particular feature which usage guide authors refer to in their discussion. I distinguish six categories of dimensions of usage: FREQUENCY, MODE, REGISTER, SPEAKERS, VALUE, and VARIETY (see Section 4.3 for explanations and examples of each of these categories). In this chapter I will discuss the dimensions of usage referred to

in the entries for each language feature. The discussion that follows is based on the identification of references to dimensions of usage in the entries for each of the features. The identification of these references was done by annotating each of the entries using 'brat' (see Section 4.3); the complete list of annotations can be found in Appendix D.

	Freq.		Mod.		Reg.		Speak.		Val.		Var.		Total	
	A	E	A	E	A	E	A	E	A	E	A	E	A	E
ain't	37	18	32	15	78	23	43	19	27	1	41	13	258	46
like	6	5	6	4	6	0	12	0	6	6	2	2	38	10
literally	21	16	4	4	4	2	12	9	0	0	1	0	42	32
neg. conc.	24	13	15	11	25	10	40	10	13	10	27	14	144	42
obj. *I*	38	28	17	8	15	8	25	13	11	9	8	7	114	73
subj. *me*	7	7	8	5	6	3	9	5	1	1	4	3	35	19
split inf.	35	14	13	10	19	8	24	11	5	5	5	5	101	59
total	168		95		153		165		63		88		732	281

Table 5.4: References to dimensions of usage (A = number of annotations; E = number of entries in which these annotations were made)

Table 5.4 contains the counts of references to these dimensions of usage for each separate feature. For each feature, the table contains the number of annotations made (the A columns in the table) in the identification of references to each of the six dimensions, FREQUENCY, MODE, REGISTER, SPEAKERS, VALUE, and VARIETY, and the number of entries (the E columns in the table) in which these annotations were made. The counts show that entries on *ain't* have the highest number of references to dimensions of usage, followed by object *I*, negative concord, and the split infinitive. The rest of the features contain smaller numbers of references to dimensions of usage. This is partly, though not entirely, the result of the different number of entries for each feature, as is shown by the total number of entries for each feature. Thus, while *ain't* is not the most frequently treated feature, with 46 entries, it does contain a strikingly high number of references to dimensions of usage, i.e. 258; there are thus on average 5.6 annotations per entry for *ain't* (Table 5.4). A similar mismatch between the number of annotations and the number of entries can be observed in the treatment of the discourse particle *like*, which does not contain the lowest number of references to dimensions of usage, despite being covered in only ten entries. This distribution provides further evidence of the significant variation in treatment across different language features, which parallels the patterns observed in the attitudes expressed, discussed in the previous section. Table 5.4 also shows that the more frequent references to dimensions of usage concern the FREQUENCY of use, categories of SPEAKERS, and REGISTER, followed by the less frequent references to MODE of expression, language VARIETY,

and social VALUE associated with the language feature.

5.5.1 *Ain't*

The entries on *ain't* contain a variety of dimensions of usage; the annotations of the entries are found in Table D.8 in Appendix D. The large number of references itself is an indication of their importance in the treatment of *ain't*, as well as in understanding better how this treatment has changed over time. The judgements on the frequency of *ain't* range from it being described as "common" (Perrin 1950; Gilman 1989; Parrish 2002; Brians 2003) or as "frequently heard" (Bechtel 1901) to being "occasionally" used (Randall 1988; Brians 2003). There is a difference in these observations, depending on the register of use: when authors describe the occurrence of *ain't* as a 'mistake', they also observe that the word is used with high frequency; such a connection is explicitly made in a third of the entries in which *ain't* is evaluated negatively. On the other hand, when they describe the frequency of use of *ain't* in colloquial speech, they describe it as occasional. In terms of register of use, *ain't* is found to be used predominantly in informal and colloquial contexts (Bechtel 1901; Perrin 1950; Corbin and Perrin 1963; Shaw 1975), or in casual speech (Lovinger 2000). In addition, usage guides often refer to a number of more specialised uses of *ain't*, such as its use in proverbs, clichés, or fixed expressions such as *if it ain't broke don't fix it*, or *it ain't over till it's over* (Randall 1988; Pinker 2014), as well as its uses in lyrics (Pickett et al. 2005; Pinker 2014) or in literature as a marker of dialectal or regional speech (Morris and Morris 1975). It is especially with reference to these specialised uses that *ain't* is considered irreplaceable and, hence, acceptable. In terms of MODE, *ain't* is considered to occur predominantly in spoken language, while its use in writing is noted as limited. For *ain't*, 15 of the 46 entries draw on distinctions between speech and writing in its use. Importantly, the use of *ain't* in writing is found to be more acceptable in relation to its more specialised uses described above. In terms of VARIETY, *ain't* is predominantly viewed as non-standard (Corbin and Perrin 1963; Copperud 1980; Gilman 1989), or as regional and dialectal (Shaw 1975; Randall 1988), but it is not related to a particular social or regional group of speakers. When it comes to references to SPEAKERS, it is often noted that *ain't* occurs in the speech of educated or cultivated speakers (Shaw 1975; Lovinger 2000; Peters 2004; Pickett et al. 2005), despite "not conforming to the standard of the language variety of educated speakers and writers" (Randall 1988: 9–10). Finally, the VALUE of using *ain't* is strongly related to the prescriptive pressure against its use, and, thus, using

ain't carries a negative social connotation. The use of *ain't* is found to be acceptable only by speakers whose status as educated is established, while speakers who use *ain't* regularly are seen as uneducated or ignorant (Copperud 1980). In addition, the strong social and educational pressure against *ain't* is seen as the greatest obstacle to using it (Perrin 1950; Corbin and Perrin 1963).

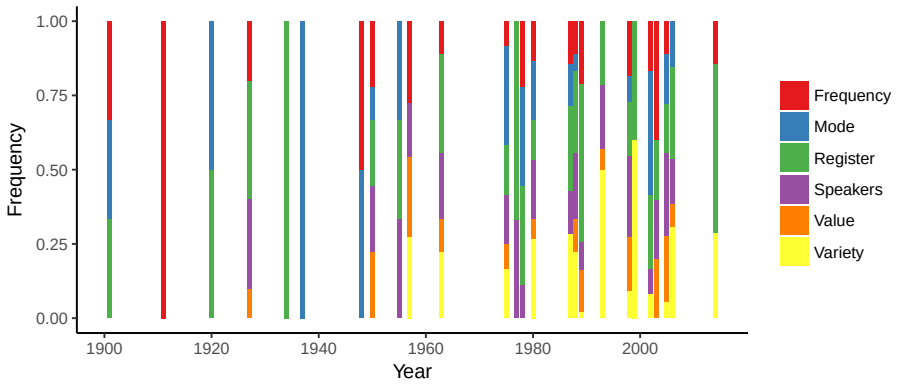

Figure 5.18: Dimensions of usage in the treatment of *ain't* (n = 258)

In addition to the variation in dimensions of usage referred to in the entries on *ain't*, a change can be observed over time in the use and frequency of these references to dimensions of usage. The distribution of the annotations of these references to dimensions of usage identified in the entries on *ain't* is plotted in Figure 5.18. Each column shows the proportion of the six types of dimensions of usage in entries published in that year (usually one or two entries, corresponding to one usage guide published in that year). The vertical axis shows the frequency of the different types of the dimensions of usage as proportions of the total number of references to dimensions of usage, in order to illustrate the changes more clearly. The figure shows that in the second half of the twentieth century, usage guides relied on many of these references to dimensions of usage. On the basis of the figure, it can also be observed that references to VARIETY, indicated in yellow, also started appearing only in the second half of the twentieth century. This suggests that in the course of the twentieth century the treatment of *ain't* became more balanced, more nuanced, and perhaps more descriptively informed.

5.5.2 The discourse particle *like*

Discourse particle *like* has the second highest number of references to usage in its entries, despite the low raw numbers, which is to a large extent a consequence of the relatively small number of entries in total (for the annotations, see Table D.9 in Appendix D). The average frequency of annotations is 3.80 per entry, which puts *like* close to *ain't* and negative concord in this respect. In terms of FREQUENCY, *like* is found to be common, or even "ubiquitous" (Garner 1998), and it is mostly associated with speech, as evidenced from the references to MODE found in four of the ten entries. In terms of REGISTER, it is associated with informal colloquial speech, and the references to SPEAKERS confirm its association with young people or teenagers (Shaw 1975; Johnson 1991; Wilson 1993; Garner 1998; Parrish 2002; Brians 2003; Pickett et al. 2005). According to the two references to VARIETY, *like* is associated with substandard speech and with the variety of English spoken in California (Wilson 1993). Finally, in terms of VALUE, connotations of using *like* are mainly associated with its indexing of young people's speech, which is, as a result, associated with being "faddish" (Lovinger 2000) or as signalling "arrested development" in adults (Garner 1998).

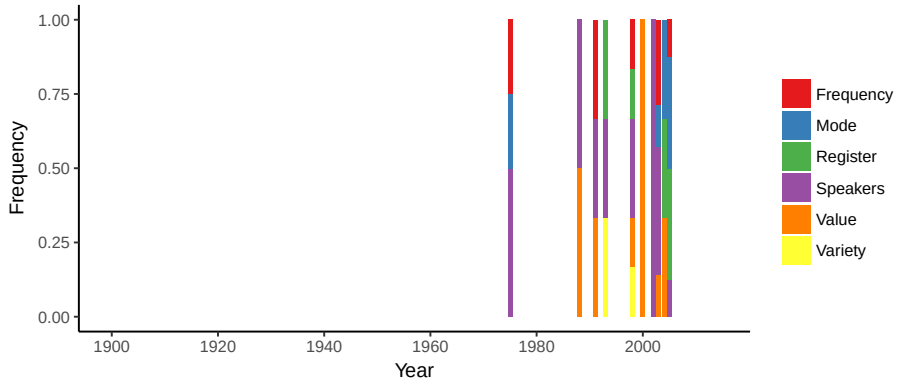

Figure 5.19: Dimensions of usage in the treatment of *like* (n = 38)

5.5.3 Non-literal *literally*

With 1.30 annotations per entry, non-literal *literally* was the feature for which I identified the lowest number of references to dimensions of usage (see Table D.10 in Appendix D). The distribution of these references across time is given in Figure 5.20. The figure shows that the majority of the entries contain a reference to FREQUENCY,

which might be related to the perception of *literally* as an 'overused' word. Entries which state that this feature is often or commonly used are found in many of the usage guides analysed (Strunk 1918; Partridge 1947; Morris and Morris 1975; de Mello Vianna et al. 1977; Bryson 1984; Gilman 1989; Johnson 1991; Sutcliffe 1994; Clark 2010; Pinker 2014), and this perception does not appear to have changed over time. References to MODE are present in four entries only. References to REGISTER and VARIETY are similarly scarce, and no explicit references to VALUE related to the use of *literally* were identified. Groups of SPEAKERS referred to are predominantly writers (Follett 1966; de Mello Vianna et al. 1977; Copperud 1980; Bryson 1984; Sutcliffe 1994; Peters 2004; Clark 2010).

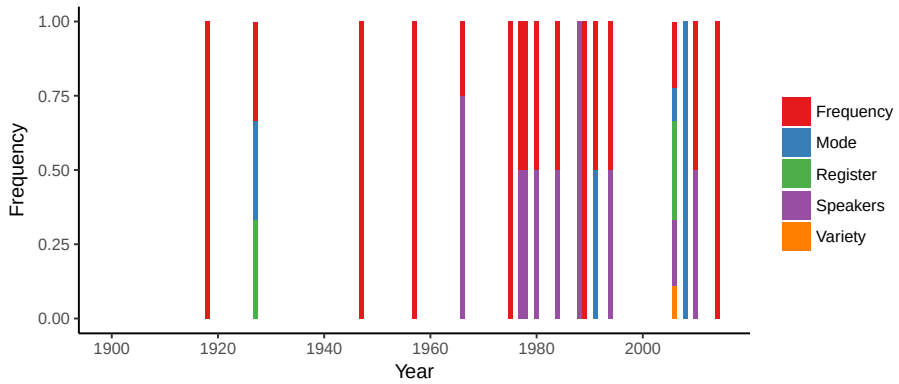

Figure 5.20: Dimensions of usage in the treatment of *literally* (n = 42)

5.5.4 Negative concord

Entries on negative concord contain the third highest number of references to dimensions of usage, with on average of 3.42 annotations per entry (Table 5.4; the complete list of annotations is given in Table D.11 in Appendix D). The distribution of the references to dimensions of usage is shown in Figure 5.21. Here again we can observe a trend similar to the one noted for *ain't* (Figure 5.18). With time, these references become more frequent and more varied. This confirms the trend observed for the treatment of the other usage features. In terms of frequency and use, the consensus seems to be that negative concord is indeed commonly used (Anonymous 1856b; Mathews 1876; Perrin 1950; Ebbitt and Ebbitt 1978; Gilman 1989; Wilson 1993), but that the context of use is non-standard English. Negative concord is found to be uncommon in formal standard English (Perrin 1950; Ebbitt and Ebbitt 1978), or

as used "by educated people" (Corbin and Perrin 1963).

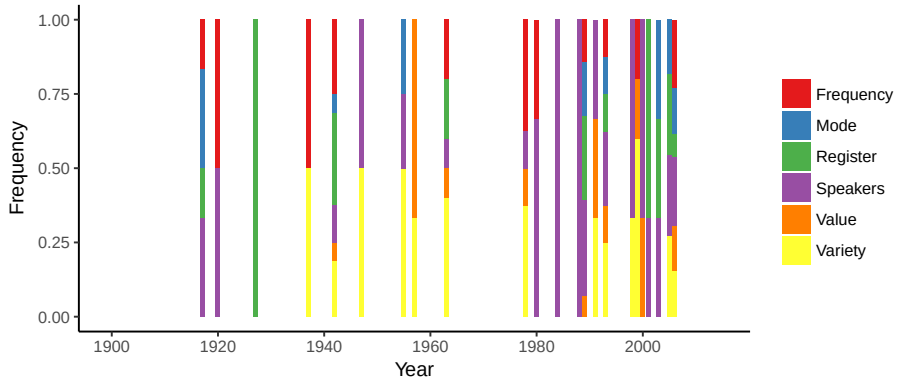

Figure 5.21: Dimensions of usage in the treatment of negative concord (n = 144)

These observations of FREQUENCY also relate to those on MODE and REGISTER. In terms of MODE, in 15 of the 42 entries, a distinction is made between the use of negative concord in speech as opposed to that in writing; this is not a high proportion, but it is nevertheless noteworthy that these kinds of distinctions are present. In terms of REGISTER, negative concord occurs in "conversation" (Krapp 1927; Gilman 1989; Peters 2004) or in jocular (Corbin and Perrin 1963; Wilson 1993) or literary (Gilman 1989) contexts. Groups of SPEAKERS are distinguished primarily on the basis of (a rather vague notion of) level of education, or with reference to language-related professions, such as writers, grammarians, or English teachers. In addition to references to language professionals, the groups most often referred to are the educated (Perrin 1950; Corbin and Perrin 1963), the less educated (Gilman 1989), and the "illiterate" (Vizetelly 1920). References to VARIETY of English in the entries on negative concord are predominantly references to the distinction between standard and non-standard English, which are present in the majority of the entries (Perrin and Smith 1955; Corbin and Perrin 1963; Ebbitt and Ebbitt 1978; Wilson 1993; Garner 1998; Peters 2004; Pickett et al. 2005). Only three entries refer to social varieties of English using terms such as "vulgate English" (Perrin 1950) or "the present idiom of the educated" (Perrin 1950). The types of VALUE referenced in entries on negative concord relate to the negative effects of its use in social situations. Negative concord is seen as not being in "fashion" among educated speakers (Perrin 1950; Corbin and Perrin 1963; Ebbitt and Ebbitt 1978). Furthermore, using negative concord is said to mark a speaker as uneducated (Evans and Evans 1957; Gilman 1989; Johnson 1991;

Wilson 1993; Peters 2004), and its use is consequently not recommended.

5.5.5 Pronouns in coordinated phrases

First person pronouns in coordinated phrases exhibit the same pattern of contrast observed in other levels of analysis, i.e. coverage (cf. Section 5.2.5) and treatment (cf. Section 5.3.5). Object *I* is more often discussed in usage guides, and there are consequently more references to dimensions of usage than for subject *me* (see Tables D.12 and D.13 in Appendix D). Figures 5.22 and 5.23 show the striking difference between the dimensions of usage found in the treatment of the two cases of pronoun use. Pronouncements related to the FREQUENCY of object *I* include observations that the feature is common and frequently used (Anonymous 1856b; Bache 1868; Ballard 1884; Payne 1911; Vizetelly 1920; Treble and Vallins 1937; Perrin 1950; Nicholson 1957; Morris and Morris 1975; de Mello Vianna et al. 1977; Ebbitt and Ebbitt 1978; Copperud 1980; Bryson 1984; Randall 1988; Booher 1992; Sutcliffe 1994; Beason and Lester 1996; Trask 2001; Clark 2010; Pinker 2014). Object *I* is found to be used in speech (Krapp 1927; Randall 1988; Gilman 1989; Wilson 1993; Pickett et al. 2005), and references to the distinction between speech and writing are present in eight of the 73 entries, which indicates that for this feature, the majority of the entries do not draw on this distinction in their treatment of this feature. REGISTER references mostly refer to colloquial use (Krapp 1927; Morris and Morris 1975; Copperud 1980; Gilman 1989), or to formal writing and literature (Ayres 1911; Copperud 1980; Pickett et al. 2005), and serve mainly to point out that object *I* is more characteristic of the former. References to SPEAKERS often mention "educated" speakers (Anonymous 1856b; Johnson 1991; Garner 1998; Brians 2003), as in the case of all other features. There are also recurrent references to language professionals, such as writers, authors, editors, and English teachers (Ayres 1911; Randall 1988; Sutcliffe 1994; Pinker 2014). References to VARIETY are found only in seven out of the 73 entries, and these are references to "dialectal speech" (Krapp 1927), "standard English" (Wilson 1993; Brians 2003), and "educated varieties of English" (Gilman 1989). In terms of VALUE, object *I* is seen as signalling "an attempt at refinement" (de Mello Vianna et al. 1977) or "a form of overrefinement" (Lovinger 2000). It is also "widely regarded as a sign of ignorance" (Pickett et al. 2005), or as a construction that can have negative influence on how its users may be perceived by other speakers (Gilman 1989).

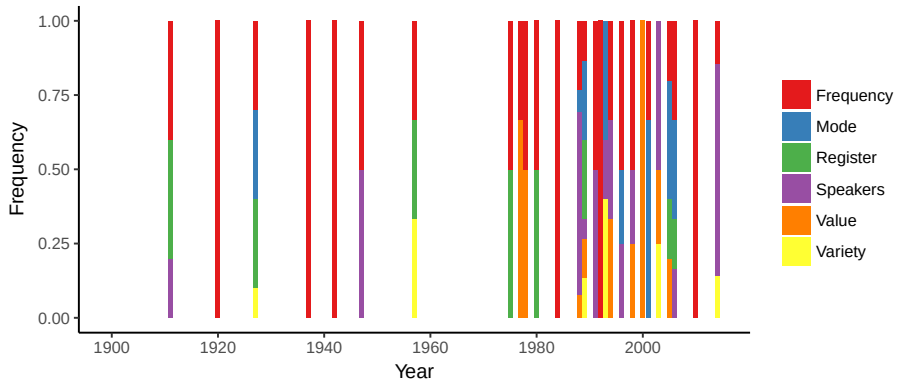

Figure 5.22: Dimensions of usage in the treatment of object *I* (n = 114)

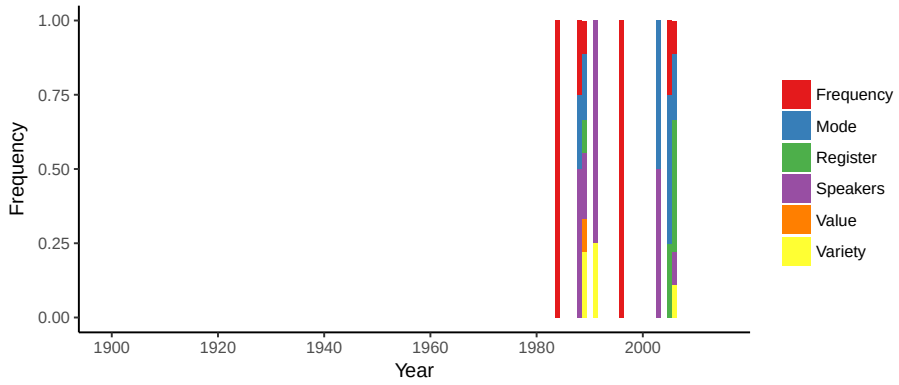

Figure 5.23: Dimensions of usage in the treatment of subject *me* (n = 35)

Compared to object *I*, subject *me* is not seen as "common", but as a feature that occurs "sometimes" (Bache 1868; Randall 1988), or is "quite rare" (Beason and Lester 1996). Distinction in MODE is found in five entries of the 19, all of which observe that this feature is usually associated with speech. References to REGISTER are less frequent, occurring in three entries only. According to these, subject *me* is used "facetiously" (Gilman 1989), or in informal "conversation" (Peters 2004). Groups of SPEAKERS referred to in entries on subject *me* are editors, proofreaders, educated people, and children. References to VARIETY take the form of mentioning what can be described as social varieties of English, such as "less educated English" (Gilman 1989) and "non-mainstream varieties of English" (Gilman 1989). Finally, in one of the entries, subject *me* is "associated with the speech of children" (Gilman 1989), which

can be considered to belong to the category VALUE.

5.5.6 The split infinitive

The split infinitive is somewhat unusual, in that the entries which discuss this feature are the longest, with an average length of 381.58 words per entry (see Table 4.1). However, references to dimensions of usage are not as numerous as those for *ain't*, *like*, or negative concord. I identified 1.71 references to usage on average in the entries on the split infinitive in my analysis. These are plotted across time in Figure 5.24. This figure is somewhat less clear than those for the other features in terms of a trend towards increase and diversification of the references to dimensions of usage over time. While a generalisation can perhaps be made that references become more frequent and more diversified, there are notable exceptions to this observation throughout the period investigated; note, for instance, the last two data points, where only references to FREQUENCY appear. The figure also shows that there are gaps in the occurrence of references to dimensions of usage, most strikingly for the 1940s. One reason for this may be the fact that there is only one entry covering the split infinitive in the usage guides published in that decade in my collection, found in Stratton (1949). An additional examination of this entry revealed that it deals with the intralinguistic constraints on the split infinitive, such as cases in which split infinitives are better than non-split infinitives, but it does not touch upon extralinguistic dimensions of usage.

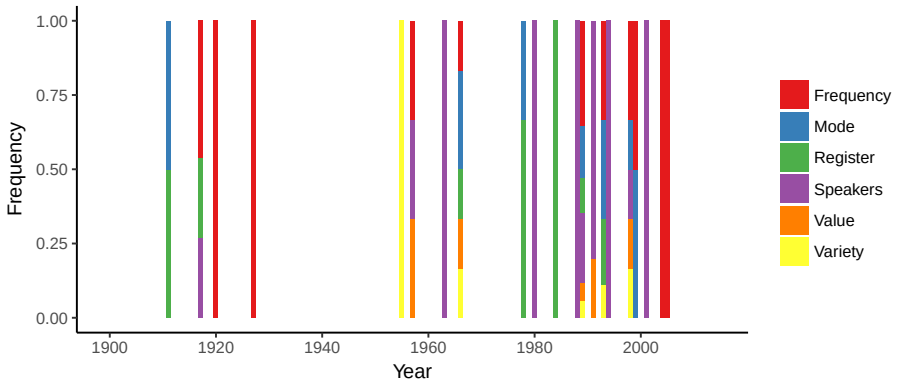

Figure 5.24: Dimensions of usage in the treatment of the split infinitive (n = 101)

Entries on the split infinitive most often contain references to FREQUENCY, followed by references to SPEAKERS. The other categories are less frequently found.

The references to frequency of use of the split infinitive show a lack of consensus. While many of the entries note that the split infinitive is frequently found in all types of registers and contexts, some entries observe that it is nevertheless rare compared to the regular infinitive (Hall 1917; Follett 1966). It should, however, be noted that Hall (1917) is a study of usage based on an empirical analysis of a corpus of literary works. It perhaps comes as no surprise that the observations made by Hall (1917) are contradictory to what is stated in other entries from the same period. Speech *vs.* writing distinctions, i.e. references to MODE, are found in ten of the 59 entries, supporting the idea that usage guides treat features without sufficient regard for differences between speech and writing, or differences between levels of usage. The references to register, which are somewhat more frequently found than references to mode, show that the split infinitive is found not only in colloquial or informal speech, but also in "general English". That this is a stylistic usage feature is also confirmed by the description of this usage problem as "a question of style" (Bryson 1984). Groups of SPEAKERS often referred to are "writers" (Compton 1898; Hall 1917; Corbin and Perrin 1963; Copperud 1980; Johnson 1991; Garner 1998). Varieties of English referred to in entries on the split infinitive are "standard English" (Perrin and Smith 1955; Wilson 1993). Finally, the VALUE associated with splitting infinitives is that it is seen as a "deplorable breach of etiquette" (Nicholson 1957), or as associated with "the speech of the less educated" (Gilman 1989). Interestingly, references to VALUE also refer to the connotations of not splitting infinitives, which is seen as "showing off" (Johnson 1991).

5.6 Conclusion

The first observation to be made on the basis of the results presented in this chapter is that the usage guide treatment of various points of usage is not always characterised by consensus. When it comes to attitudes to usage, the lack of consensus is more striking in the context of the split infinitive, and to some extent *ain't*. Furthermore, the lack of consensus is most notable in the context of references to dimensions of usage. A similar lack of consensus was identified for eighteenth-century English normative grammars by Leonard (1929), and notably also for French normative grammar pronouncements on the future temporal (Poplack and Dion 2009). However, this is only one aspect of the treatment of usage features. For instance, with respect to the frequency of negative concord, we find both observations that its use is frequent,

as well as observations that it is not; such observations are intertwined with references to REGISTER and VARIETY: negative concord is found to be frequent in colloquial, dialectal speech, or non-standard speech, but not in written standard English. In addition, where written standard English is mentioned with reference to negative concord, it is to observe that the feature is not common, but does occur in specific contexts, such as dialogue or fiction. Furthermore, this perspective also provides some evidence against the claim that treatment of usage in usage guides does not take into account aspects of language variation. This analysis suggests that usage guide writers sometimes do take into account aspects of variation in language, even if, in some cases, that variation may be criticised. Even though the number of guides which draw on facts of language variation is still rather small, it is a significant aspect of the usage guide tradition, which merits further investigation.

The second observation is that every usage problem has its own history (cf. Chafe 1984; Tieken-Boon van Ostade and Ebner 2017). There are striking differences in the attitudes expressed towards each feature, which is in line with the hypothesis of Busse and Schröder (2010: 100) that "long-standing cases of disputed usage are treated differently from more recent ones". While both Busse and Schröder (2010) and the present study show that there are indeed such differences, the results of my analyses do not always support their finding that more recent disputed usages are treated more descriptively than older ones, which Busse and Schröder (2010) concluded on the basis of comparing the treatment of *hopefully* with the treatment of *different to / from / than*. In the present study, the picture we see by comparing older and newer usage problems is more complex. A comparison, for instance, of the treatment of the split infinitive with that of *like*, or *literally*, reveals that in both cases more recent usage problems are treated more negatively. A generalisation would thus be difficult to make solely on the basis of the 'age' of the usage problem concerned; additional aspects, such as the extralinguistic associations or personal qualities which it indexes, can play a role in the attitudes to different usage problems expressed in usage guides.

The third observation, related to the previous one, is that prescriptively targeted language features cannot be lumped together in discussions or analyses of prescriptive influence. The most important case here is the split infinitive, as it is usually cited as one of the archetypal usage problems, and is assumed to have been heavily proscribed. As a result, as discussed in Section 2.3, in some discussions of language change prescriptive influence on the use of the split infinitive is readily postulated (cf. Fitzmaurice 2000b; Calle-Martín and Miranda-García 2009; Perales-Escudero 2011), on the basis of the assumption that prescriptive ideology strongly affects

this feature. However, the present analysis of treatment of usage features in usage guides suggests that prescriptive ideology has changed over time, and that the split infinitive is becoming increasingly more accepted. This, as I will demonstrate in the next chapter, has important implications for the analysis of prescriptive influence on language variation and change.

The last observation is related to the importance of these results for the study of usage guides in general. The multi-level analysis of treatment conducted here may be reasonably assumed to indicate a change in the genre of usage guides. The increase in RESTRICTED and ACCEPTABLE entries, and the increasing use of positive attitudes, as well as references to dimensions of usage, show that even for strongly stigmatised non-standard features usage guide entries are increasingly likely to offer a more balanced, perhaps descriptively informed, account of those features. This balanced account is crucially characterised by a distinction among various levels of acceptability; the lack of such a distinction in prescriptive accounts of language use is often cited as one way in which those accounts fail to represent or discuss language use more realistically. My analysis thus provides counter-evidence to this observation, and suggests that, taken as a whole, the usage guide genre in American English may be moving, slowly but surely, towards a more balanced account of language use. Additionally, this also provides evidence that there is a clear group of usage guides that offer a more neutral, balanced, or impartial account, which may be related to Algeo's (1991b) typology of usage guides. Furthermore, the results seem to be in line with the findings of the comparative analysis of British, American, and Australian usage guides conducted by Peters and Young (1997: 318), which showed that American usage guides tend to have higher instances of referencing to other sources than British usage guides. Peters and Young (1997: 321) also found that, despite the great variation in types of usage guides in both the American and the British traditions, American usage guides tend to "accept" more than British usage guides. Ultimately, all this taken together could also indicate that in the future, usage guide treatment of language use and contested language features may become ever more influenced by descriptions of language use patterns.

CHAPTER 6

Patterns in actual language use

6.1 Introduction

In Chapter 4, I explained the general approach taken in this study to exploring the question of whether prescriptive metalinguistic discourse affects usage patterns both across time and across register. I also explained that this will be approached by comparing patterns of change observed in the treatment in usage guides of the six linguistic features investigated, i.e. precept (see Section 4.2), with patterns of variation in the actual use of those linguistic features, i.e. practice. Having analysed the precept data in Chapter 5, I now turn to the patterns of actual use of each of the six features, i.e. *ain't*, the discourse particle *like*, *literally*, negative concord, pronouns in coordinated phrases (i.e. object *I* and subject *me*), and the split infinitive.

The data on actual use are taken from the two large-scale corpora introduced in Section 4.4, COCA and COHA. In that section, I also explained that the patterns of language use will be explored on the basis of two analytical approaches, or two types of metrics (cf. Biber et al. 2016). First, I look at the patterns of variation by identifying the text-linguistic frequency of occurrence of linguistic variants considered problematic to varying degrees from a prescriptive point of view, i.e. *ain't*, the discourse particle *like*, the non-literal use of *literally*, negative concord, object *I* and

subject *me*, and split infinitives. Secondly, I use the variationist approach to analyse the proportion of the use of some of these variants in the context of their linguistic variables by identifying the proportion of use of the unacceptable linguistic variant out of the total number of environments in which it could occur. For example, I look at the proportion of *ain't* for *be not* out of the total number of environments in which a *be not* variant is used, or the proportion of split infinitives out of the total number of infinitives modified by a single adverb, both split and not split. For more details on the identification, extraction, and disambiguation of the occurrences for each of the features, see Section 4.4 and Appendix C. Sections 6.2 – 6.7 discuss the patterns of occurrence of each of the six linguistic features across time periods and the various corpus genres: academic, fiction, magazines, newspapers, and spoken (see also Table 4.2).

In addition to this, I present an analysis which aims to empirically identify the potential influence of prescriptivism on the use of the split infinitive. Using this feature as a case study, I conduct a multifactorial analysis, in order to identify the extent to which the use of split infinitives is associated with the use of other prescriptively targeted features, at the level of individual texts. Section 6.8 presents the results of this analysis. In the final section, I bring these findings together, and discuss the issue of the influence of prescriptivism on language use.

6.2 *Ain't*

did, or with other auxiliaries and modals.

As explained in the previous section, for the purposes of this analysis I rely on two types of metrics in order to analyse the patterns of usage of *ain't* across time periods and genres in the corpora. The first account of the patterns of use of *ain't* is the normalised frequency of use of all occurrences of *ain't* in the corpus, irrespective of their function. The reason that this may be considered a good indicator of the changing patterns of usage of *ain't* is that, regardless of the *function* of *ain't*, the *form* is generally stigmatised. The second metric measures the proportion of *ain't* used for *be not*, in the context of all possible environments of *be not*, as well as *ain't* used for *have not*, in the context of all possible environments of *have not*. The reason for the second type of metric is that, despite the general stigmatisation of *ain't*, there is a sense of *ain't* for *be not* being somewhat more acceptable than *ain't* for *have not*. In order to explore the extent to which such ideas identified in the precept data relate to

patterns of actual language use. The variables used in the analysis are given in Table 4.3.

In Section 3.3, I briefly outlined the major findings from previous studies on the variation in the use of *ain't* in American English, in terms of linguistic and sociolinguistic constraints. The complex variation in the use of *ain't* is reflected in the data analysed for this study. First, with reference to the linguistic variation in the use of *ain't*, the analysis showed that alongside the predominant uses of *ain't* in environments of *be not*, in examples (48) and (49), and *have not*, as in (50), there were a number of cases where *ain't* is used as a variant of *didn't*, as in (51), with modals such as *mustn't*, as in (52), and possibly, in a small number of cases, with *wasn't* (cf. Anderwald 2002). Finally, what is an interesting and, I believe, significant finding resulting from the corpus data was the discovery of a number of occurrences of a metalinguistic mention of *ain't*, in (53), in which the word is criticised or implicitly associated with the proscription against its use. These will be discussed in more detail in the final part of this section.

(48) He thinks he **ain't** a man any more. (1987, fiction, COHA)

(49) He **ain't** saying that to my face. (2006, spoken, COCA)

(50) You **ain't** said yes yet. (1932, fiction, COHA)

(51) Why y'all **ain't** call me? (2011, magazine, COCA)

(52) You must be joking, **ain't** you, Mr Luther? (1940, fiction, COHA)

(53) Language of this sort could be terrifying to someone who only the week before at Miss Burke's had been sent to detention for saying **ain't**. (1959, fiction, COHA)

The data thus confirm previous accounts of the variation in the uses of *ain't*; however, it also confirms that most of these uses, exemplified in (48)–(53), are fairly rare, even in non-standard spoken data. Since the corpus data used for the present analysis reflect the standard American language variety, it is not surprising that these variants are very rare. This means that, despite the existence of the different variants, the greatest majority of *ain't* uses are found in the environments for *be not* and, to a lesser extent, *have not*. As a result, all other cases were excluded from the variationist analysis presented here.

The normalised frequency distributions of all occurrences of *ain't* across time periods are shown in Figure 6.1. The figure contains two subfigures, one for the

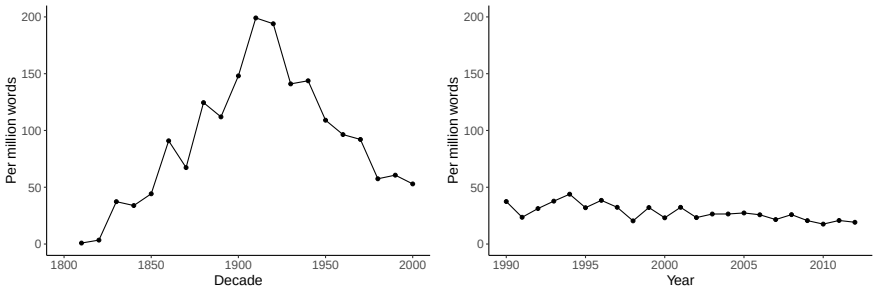

Figure 6.1: Text-linguistic frequencies of all occurrences of *ain't* across time (COHA: n = 39,348; COCA: n = 12,228)

rate of occurrence of *ain't* in COHA, and the other for the rate of occurrence of *ain't* in COCA.[1] Due to the make-up of the corpora, as well as the proportionally different time scales they cover, the time periods used for COHA are decades, while those for COCA are years. The second subfigure can thus be seen as zooming in on the last two decades in the period under investigation. As the graphs show, the frequency distribution of *ain't* undergoes a striking increase until the 1910s, followed by a similarly dramatic decline in the course of the twentieth century. Since the year 2000, the frequency of *ain't* has remained steadily low. While these results might lead us to postulate that prescriptivism may have had some effect on the use of *ain't*, it is important to consider other factors first.

One of those factors is register variation, which I also explore using both text-linguistic and variationist metrics to establish the normalised frequencies and proportions of *ain't* across the subsections of the two corpora used. The results from the text-linguistic analysis are given in Figure 6.2, which shows the normalised frequencies of occurrence of all cases of *ain't* across sections of the two corpora. The vertical axis represents the number of occurrences of *ain't* per million words across the major genre sections of the two corpora, i.e. fiction, magazine, newspaper, and non-fiction in COHA, and academic, fiction, magazine, newspaper, and spoken in COCA, which are plotted on the horizontal axis. The two plots show that the frequency of occurrence of *ain't* is highest in fiction in both corpora, with the fiction section in COHA containing the highest rate of occurrence of *ain't*.

[1] As evident from the graphs, the two corpora overlap for the period 1990–2000. There is some overlap in the materials included in the two corpora for the final decade of the twentieth century. For transparency, I represent the figures in their entirety, as well as separately, due to the fact that the make-up of the corpora is not entirely the same.

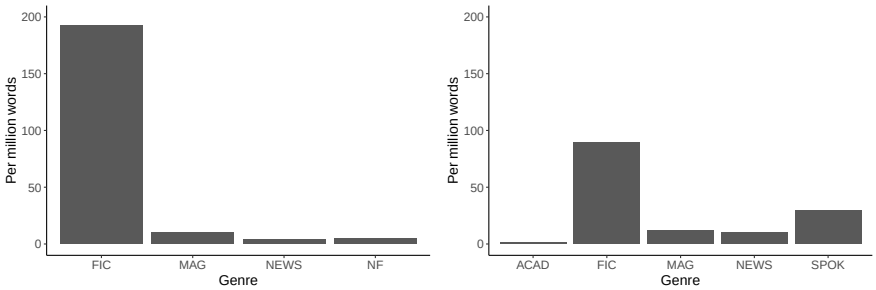

Figure 6.2: Text-linguistic frequencies of all occurrences of *ain't* across corpus sections (COHA: n = 39,348; COCA: n = 12,228)

Since the results of the effects of genre on the use of *ain't* show that the form is especially frequently found in fiction, I also plotted the trends for the occurrence of *ain't* in all other corpus genres taken together, excluding fiction. These results are given in Figure 6.3. There is a clearly even trend, with almost no difference whatsoever in the normalised frequency of occurrence over the course of the entire period investigated. There is a very slight increase at the end of the twentieth century, which could perhaps partly be explained by the presence of spoken data in COCA. A comparison between Figures 6.1 and 6.3 confirms the fact that the large-scale increase observed over time in the frequency of occurrence of *ain't* in COHA is an effect of its increase in fiction.

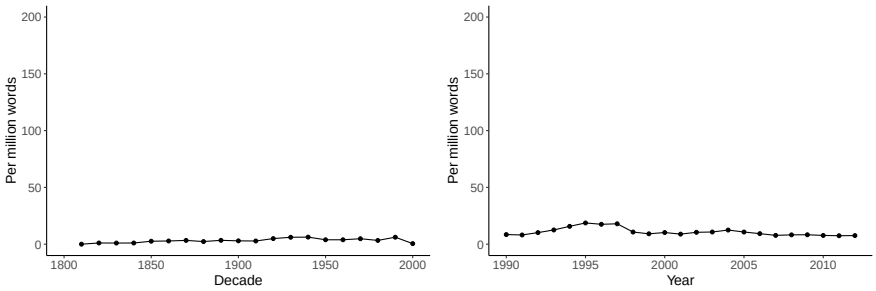

Figure 6.3: Text-linguistic frequencies of all occurrences of *ain't* across time, excluding fiction (COHA: n = 1,373; COCA: n = 4,751)

The question then is whether this increase and subsequent decrease in the rate of occurrence in fiction is a change in the use of *ain't* in this particular genre, or whether there are other explanations for the trend observed in Figure 6.1, such as the make-up of the fiction section. The latter scenario was investigated with further exploration of

the make-up of COHA, focusing specifically on the subgenres included in the fiction section (see Table 4.2). This analysis shows that the percentage of drama texts out of all fiction texts is the highest for the 1910s (i.e. 33.90%) and the 1920s (i.e. 33.10%; see Table C.1 in Appendix C for the percentage of drama texts for the other decades in the corpus, which is lower than for the 1910s and the 1920s). Similarly, almost 30% of all occurrences of *ain't* in fiction in those two decades come from drama texts. This means that *ain't* is a feature which is characteristic of fiction in general, and of plays in particular. This in turn also suggests that the increase and decrease in the rate of occurrence of *ain't* in fiction is more likely to be related to the higher percentage of drama texts for those two decades, rather than being a consequence of changing patterns of usage.

Having established that there has been no change in the rate of occurrence of *ain't* in American English since the beginning of the nineteenth century, and that the variation patterns observed are the effect of register, I now turn to the question of how this finding relates to the change in treatment of *ain't*. I already pointed out in the discussion of the treatment of *ain't* in usage guides (see Section 5.3) that during the course of the twentieth century this feature was increasingly viewed as acceptable in restricted contexts. On the basis of these two analyses, it could of course be the case that there is no relationship between language use and usage guide treatment, and that the two developments identified here are independent of each other. However, given the salience of *ain't* both as a dialectal feature and as a usage problem, this seems unlikely. Rather, it seems more likely that usage guides have changed their treatment of *ain't* as a consequence of the low frequency of the form in general standard American English, as well as its stable place as a dialectal feature, mostly used in fiction, and especially drama. In order to explain how this relates to the usage guide treatment of *ain't*, it is important to look more closely at the kind of acceptability of *ain't* that is expressed in usage guides. We can observe that, while usage guide writers, especially in the second half of the twentieth century, tend to be more accepting of *ain't*, this acceptability is still restricted to a few contexts. These contexts include specific functions of *ain't* in marking non-standard or dialectal speech in works of fiction, the use of *ain't* in set phrases and idioms, and its use in popular songs. These functions, it seems, have become more stable over the course of time, resulting in the low overall frequency of *ain't*. It is precisely this kind of regularisation of the contexts of use of *ain't* that may have allowed for its higher acceptability in restricted contexts in usage guides. The use of *ain't* in drama may therefore be understood as the reason for the acceptance of *ain't* in restricted contexts. In other words, once the

feature became very limited in frequency in general language use, and its use in fiction became stable, the need to proscribe *ain't* slowly disappeared. This also implies that there is a time lag in the change in treatment.

The discussion so far has been based only on the text-linguistic frequencies of occurrence of *ain't* across corpus sections. In order to gain a better understanding of the use of *ain't* in the context of the variables *be not* and *have not*, I turn to the variationist analysis of *ain't*, looking not only at how *ain't* is used in particular types of texts or periods of time, but also at how it is used in relation to the other variants for *be not* and *have not*. Figure 6.4 shows the proportion of cases realised with *ain't*, as opposed to all other cases of *be not*, realised by both full and contracted forms, across decades in COHA and years in COCA. Figure 6.5 shows similar proportions for *ain't* functioning as *have not*, as opposed to the total number of cases of *have not*.

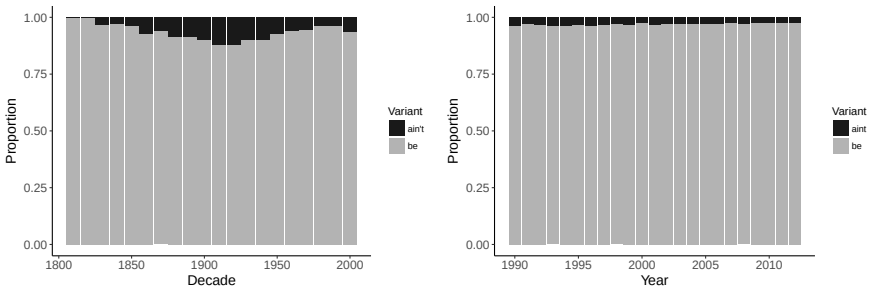

Figure 6.4: Proportion of occurrences of *ain't* (COHA: n = 30,106; COCA: n = 10,154) across time out of the total number of *be not* environments (COHA: n = 415,677; COCA: n = 637,133)

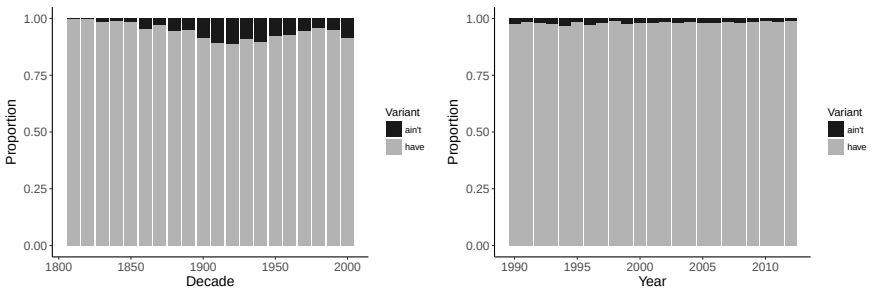

Figure 6.5: Proportion of occurrences of *ain't* (COHA: n = 6,762; COCA: n = 2,061) across time out of the total number of *have not* environments (COHA: n = 102,584; COCA: n = 110,890)

The variationist analysis of *ain't* for *be not* shows that the percentage of *ain't* was somewhat higher in the beginning of the twentieth century, as shown in Figure 6.4, reflecting the increase in frequency observed in the overall distribution of *ain't* in Figure 6.1. The figures for *ain't* for *have not*, given in Figure 6.5, are somewhat lower than *ain't* for *be not* in the historical data, and not much different in the contemporary data.

Turning to the patterns of variation across genre sections of the corpora, Figures 6.6 and 6.7 plot the proportions of *ain't* occurrences from the total number of possible environments in the context of the variables *be not* and *have not*, respectively. The distribution of uses of *ain't* for *be not* and *ain't* for *have not* across genres shows that fiction is the genre where almost all uses of *ain't* are found. The proportion of *ain't* is slightly higher in the COHA data, but both plots show that the overall proportion of *ain't* is fairly low.

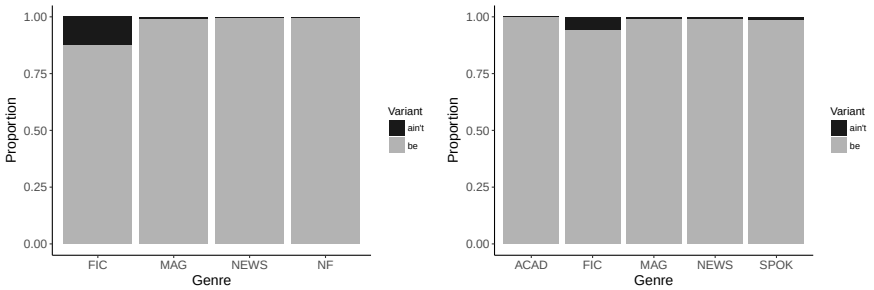

Figure 6.6: Proportion of occurrences of *ain't* (COHA: n = 30,106; COCA: n = 10,154) across corpus sections out of the total number of *be not* environments (COHA: n = 415,677; COCA: n = 637,133)

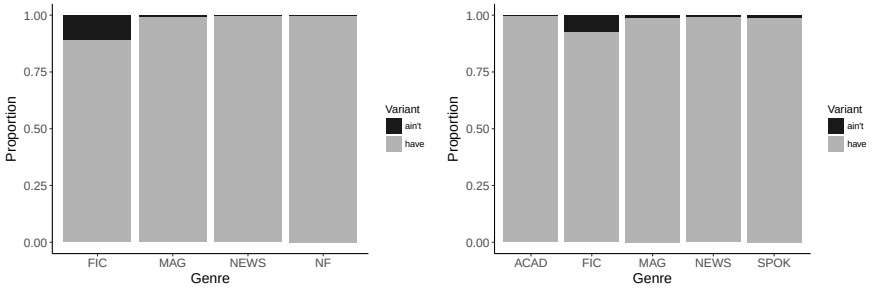

Figure 6.7: Proportion of occurrences of *ain't* (COHA: n = 6,762; COCA: n = 2,061) across corpus sections out of the total number of *be not* environments (COHA: n = 102,584; COCA: n = 110,890)

Having examined both the text-linguistic and the variationist frequencies of occurrence of *ain't*, I now turn to a brief discussion of the types of pronouns which *ain't* is most commonly used with. I explore this question in order to investigate whether there is any empirical basis for the high acceptability in usage guides of *ain't* with the first person singular, as opposed to its use with the third person singular. To illustrate this, I will focus only on the use of *ain't* for *be not* in COCA. In this dataset, in 38% of the cases *ain't* is used with something other than a personal pronoun, i.e. with an noun phrase headed by a noun or a proper noun. Of the remaining 62%, *I* is used in 17% of the cases, *it* in 19% of the cases, and *he* and *she* in 6% and 3% respectively. *I* thus appears to be only the second most frequent pronoun, after *it*. This evidence suggests that the prescriptive ideology concerning the acceptability of *ain't* is not supported by its actual use.

Finally, as mentioned at the beginning of this section, I found that in the corpus data used for the present study *ain't* also occurs in metalinguistic contexts, illustrated in examples (54)–(57) below. These examples testify to the stigmatised status of *ain't* and its association with non-standard speech. In some sense, then, these examples provide evidence at least for the cultural influence of prescriptivism, and certainly of the status of *ain't* as a usage problem.

(54) He looked up, clear-eyed to her pleasure, and wounded her with delight in the way he said, "I **ain't** done anything." "That's right," she said, nodding firmly. "Don't say **ain't**, just because I fergit now and then when I'm working hard, and haven't time for the fancies and the rights of this and that. But I don't want my baby-boy t'get habit of speaking wrongly." (1936, fiction, COHA)

(55) Language of this sort could be terrifying to someone who only the week before at Miss Burke's had been sent to detention for saying **ain't**. (1959, fiction, COHA)

(56) Or Lynn Smith Jr., a rancher who wears a cowboy hat, tucks pants into boots and still says '**ain't**'. (2000, newspaper, COCA)

(57) She wiped her eyes and gave Pelton a withering look. "Don't say '**ain't**'!" There is no such word... (1996, newspaper, COCA)

6.3 The discourse particle *like*

The frequencies of occurrence of the discourse particle *like* in COHA and COCA show a definite increase in the use of this feature over time. A variationist analysis of *like*

was not attempted in this study, due to the difficulty of ascertaining the variable context in which the feature occurs and establishing the total number of potential environments (see Section 4.4). On the basis solely of text-linguistic frequencies, it can be noted that the discourse particle *like* has indeed seen a striking increase in occurrence in the corpus data analysed. This increase is particularly salient in the COCA data, which contain a spoken language section. Comparing this distribution to the usage guides' coverage and treatment of *like*, it is clear that the more likely phenomenon we are observing is that usage guide writers are reacting to a robust process of language change, which has also been accompanied by social stigmatisation, in some sense independent from the usage guide tradition.

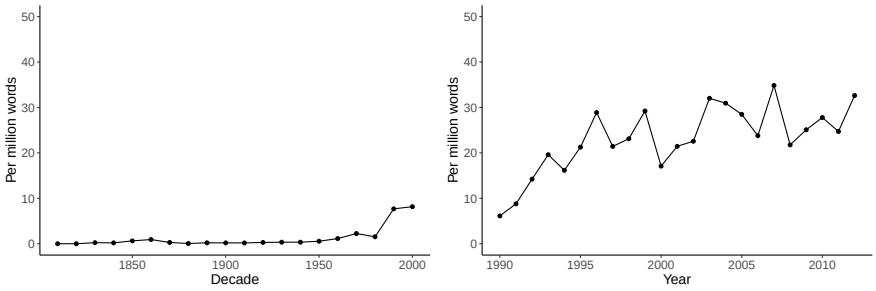

Figure 6.8: Text-linguistic frequency of the discourse particle *like* across time (COHA: n = 634; COCA: n = 10,020)

In terms of genre, Figure 6.9 shows that spoken data contain the highest number of instances of the discourse particle *like*, followed by fiction. Another important observation is the much higher frequencies observed for the COCA data, which reach almost 100 occurrences per million words, compared to 2.5 occurrences per million words for the highest frequency per genre observed in COHA. These distributions are hardly surprising, as the feature is a typical spoken language feature.

Both of these patterns of occurrence suggest that the case of the discourse particle *like* is a robust language change in progress, and is being led by the spoken language, as has been confirmed in many previous studies (see Section 3.4). This in turn provides further evidence that usage guides are responding to this development, which may suggest that *like* is on its way to becoming a usage problem. A crucial factor in this process, however, is the social stigmatisation of *like*, which preceded its treatment in usage guides. I return to these aspects of the use of the discourse particle *like* in Chapter 7 of this thesis.

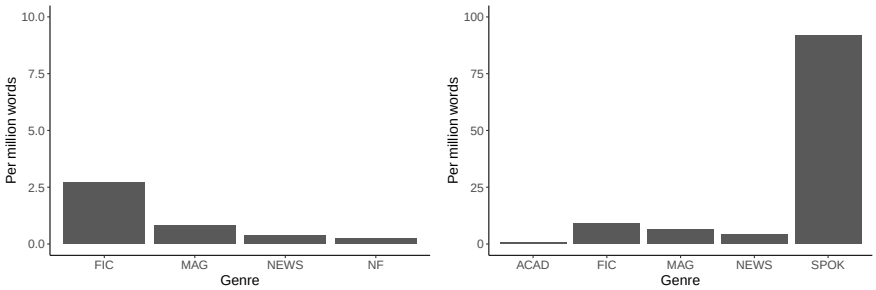

Figure 6.9: Text-linguistic frequencies of all occurrences of *like* across corpus sections (COHA: n = 634; COCA: n = 10,020)

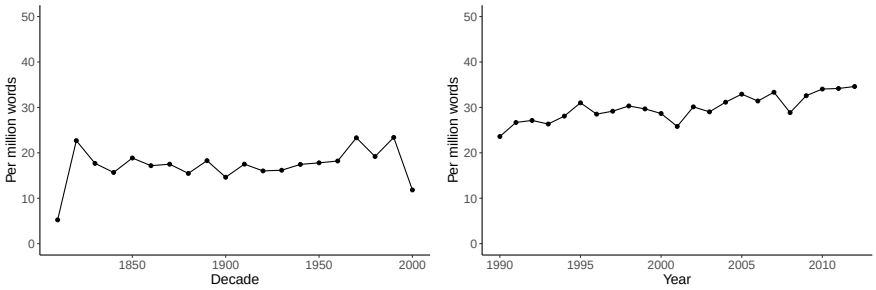

Figure 6.10: Text-linguistic frequency of *literally* across time (COHA: n = 6,848; COCA: n = 14,946)

6.4 Non-literal *literally*

The overall frequency of use of *literally*, as shown in Figure 6.10, has been increasing very slightly over the course of the last twenty years, from around 20 to a little more than 30 occurrences per million words, hardly a substantial increase. What these figures show, however, is that the notion that the word has come to be 'overused' is clearly not borne out by the data. In addition to plotting the overall frequency of occurrence of *literally*, two additional steps were taken in the analysis in order to arrive at a better understanding of the distribution of its three uses, as explained in Sections 3.5 and 4.4: primary use, dual use, and non-literal use.

As discussed in Section 4.4, the first step in the analysis was the automatic disambiguation of cases in which *literally* is used with its primary meaning from all other uses (see Appendix C for a description of the procedure). The results of this analysis are plotted in Figure 6.11, which shows the proportion of primary uses of

literally as opposed to all other uses. The two graphs in the figure show that the use of *literally* in its primary meaning has remained fairly stable over time. The figure also

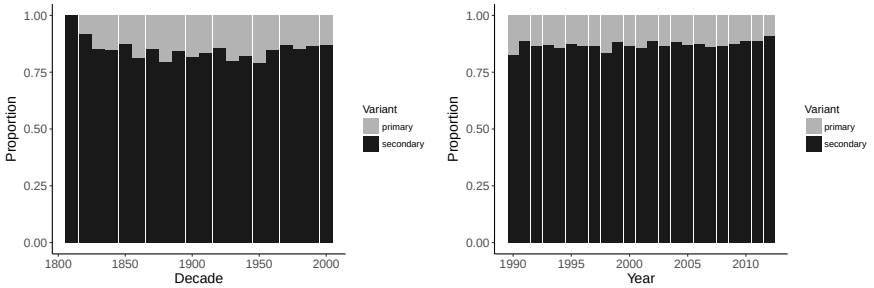

Figure 6.11: Proportion of primary uses of *literally* (COHA: n = 1,079; COCA: n = 1,937) across time compared to all other uses (COHA: n = 6,848; COCA: n = 14,946)

shows that the primary uses of *literally* are not the majority of the occurrences; rather, the reverse is the case. Note that this kind of automatic disambiguation, which is carried out using Python scripts, and relies on the part-of-speech tags in the corpus data, is bound to contain some degree of error in its precision and recall. In order to obtain a better picture of the rest of the uses of *literally*, as well as to supplement the automatic disambiguation, additional manual disambiguation was conducted on a sample of the total number of occurrences of *literally*, as described in Section 4.4. In this manual analysis I distinguished between the three uses of *literally*, *viz.* its primary, dual, and non-literal uses.

The results from the manual analysis are given in Figure 6.12, which plots the proportions of the three uses of *literally* across decades in COHA and years in COCA. A number of observations can be made on the basis of these trends. First, the graphs show that the number of primary uses of *literally* has decreased slightly over time. This is certainly the case if the distributions of *literally* in COHA and COCA are compared. It is worth comparing this figure with Figure 6.11, which shows the proportion of primary uses of *literally* against all other uses. The comparison shows that the difference between these two is in degree, but not in quality. This difference is not surprising, given that the automatic disambiguation is not as precise as manual analysis. However, it is reassuring that the patterns of distribution follow the same trend, which means that the automatic disambiguation is to a large extent reliable in tracking patterns of use. Secondly, it can also be observed that non-literal uses of *literally* are fairly rare, and that there has been little change in this respect in the course of the twentieth century.

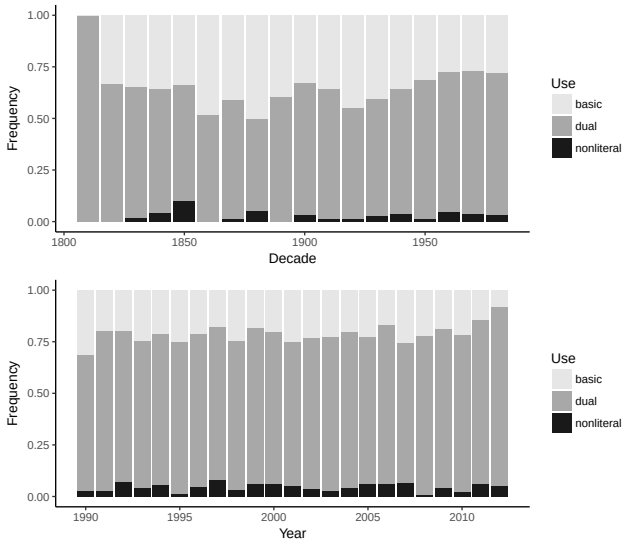

Figure 6.12: Proportion of the three different uses of *literally* across time, based on a sample of all occurrences of *literally* (COHA: n = 1,141; COCA: n = 2,864)

Finally, the dual uses of *literally* seem to be the most common, and the results of this analysis indicate that this was the case throughout the nineteenth and twentieth centuries. An interesting question is when in the history of the English language this use started to increase. I have not explored this question further, as the period before the nineteenth century is beyond the scope of this thesis, and previous studies on *literally* provide few corpus-based insights into its use in earlier periods. Some evidence on when the dual and non-literal uses of *literally* were first recorded can be found in the entry on *literally* in the *Oxford English Dictionary*;[2] on the basis of the instances recorded there, it can perhaps be hypothesised that these dual and non-literal uses of *literally* started to develop and to increase in frequency during the seventeenth century.

The frequency of occurrence of *literally* in all its uses across the genre sections in the corpora is given in Figure 6.13. The figure shows that the sections in COHA do not differ greatly in terms of frequency per million words. In COCA, the spoken section contains more occurrences of *literally* than any other sections. In both COCA and COHA, the fiction and the newspaper sections have the lowest frequency of occurrence of *literally*.

[2]See entry on *literally* in *OED Online*, available at www.oed.com.

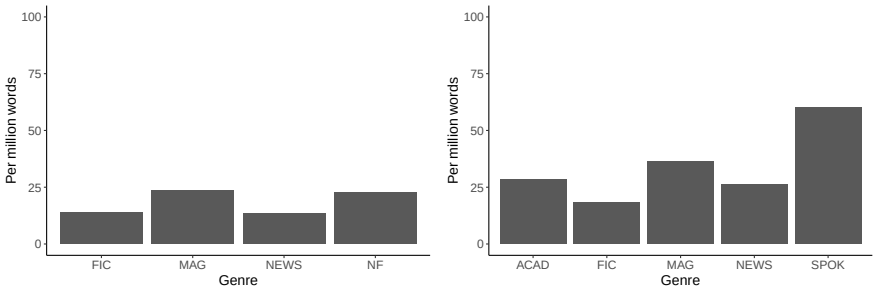

Figure 6.13: Text-linguistic frequencies of all occurrences of *literally* across corpus sections (COHA: n = 6,848; COCA: n = 14,946)

The analysis of the primary use of *literally*, as opposed to all other uses, shows that primary uses are highest in non-fiction texts in COHA and in academic texts in COCA (Figure 6.14). This distribution is expected, given that the primary meanings of *literally* are its oldest and the unproblematic uses. The manual disambiguation of a sample of these uses, the results of which are presented in Figure 6.15, shows a similar distribution pattern to that observed on the basis of the automatic disambiguation of the uses of *literally* plotted in Figure 6.14. A comparison between Figures 6.14 and 6.15 shows that the difference between these two is one of degree, rather than quality. For instance, for COHA, the non-fiction section has the highest proportion of primary uses of *literally*, followed by magazine, fiction, and newspaper; the differences are the same in both Figures 6.14 and 6.15, even though in Figure 6.14 the differences across corpus sections are less pronounced. This is likely the result of the fact that some relevant cases of the primary use of *literally* have not been identified using automatic disambiguation, based on part-of-speech tags. The difference between COHA and COCA which can be established on the basis of Figure 6.15 confirms that primary uses of *literally* are higher in frequency in COHA than in COCA, which might suggest a slow pace of change over time in the distribution of the uses of *literally*. Finally, Figure 6.15 also shows that non-literal uses of *literally* are very rare across all corpus sections.

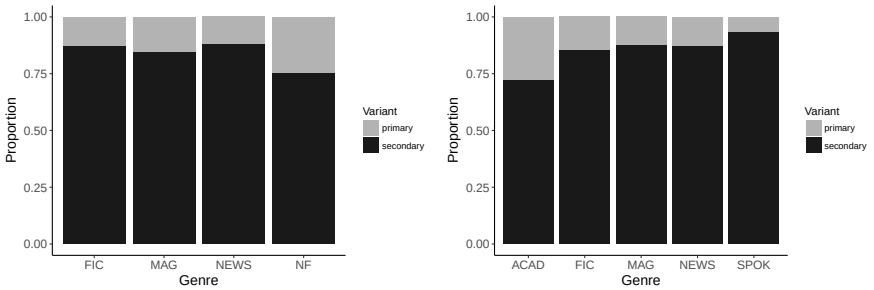

Figure 6.14: Proportion of primary uses of *literally* (COHA: n = 1,079; COCA: n = 1,937) across corpus sections out of all other uses (COHA: n = 6,848; COCA: n = 14,946)

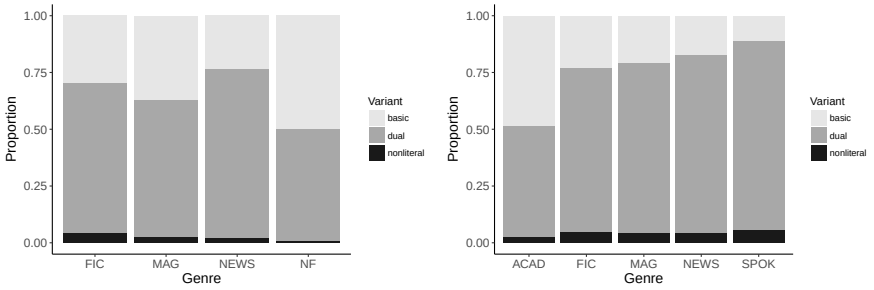

Figure 6.15: Proportion of uses of *literally* out of a sample of occurrences across corpus sections (COHA: n = 1,141; COCA: n = 2,864)

From these results it can be concluded that the frequency of the word *literally*, in all its uses, has not increased strikingly in the last 200 years, and that non-literal uses of *literally* are very rare. The primary uses of *literally* seem to have decreased somewhat in frequency in favour of its dual uses, although this change does not seem to be progressing rapidly. Comparing these results with the treatment of *literally* in usage guides leads to a number of observations. First, *literally* is a salient case of variation, and the extension of its meaning is considered problematic mostly due to the perceived opposition between its primary and its secondary uses (but see Powell 1992, who argues that there is a continuity of metalinguistic meaning underlying all uses of *literally*). Second, due to the salience of this process of variation and change, and perhaps in part due to the characteristic case of non-literal *literally*, this process has been interpreted by usage guide writers in a way which is not entirely supported by evidence from language use. The usage guide treatment of *literally* tends to distinguish

between the 'strict' use of the word, which corresponds to its primary use, and all other uses, where *literally* is used either to express the opposite of its literal meaning, or simply to intensify an expression. However, the biggest problem with this kind of division is that the majority of the uses of *literally* are dual uses: these are cases in which it has both a literal meaning and an intensifying function. For example, in cases such as *There were literally millions of people*, the function of *literally* is both to express that there were more than one million people and to intensify the fact that this piece of information is surprising, and therefore worth emphasising. As a result of the lack of this kind of distinction, dual uses of *literally* may often be perceived as intensifying and superfluous, even though in principle they do satisfy the condition for the "proper" use of *literally*, in that in dual uses *literally* does not violate a literal reading.

In summary, what the case of *literally* shows is that salient cases of language variation and change may rise above the level of consciousness, and provoke metalinguistic discussions. It also shows that it takes time for a certain new language variant to rise above the level of consciousness before prescriptivists start noticing it (cf. Laitinen 2009). The same argument could be made for the case of the discourse particle *like*. Another aspect of the case of *literally* and the relationship between its status as a usage problem and its treatment in usage guides is that usage guide writers are in general mistaken in their overall characterisation of the use of *literally*. First, as I mentioned above, observations about an increase in frequency of the 'overuse' of the word are not supported by the data, which show a fairly stable and low increase in the frequency of use of *literally*. Second, the statements that *literally* has increasingly been used to mean precisely the opposite of its primary meaning are not supported by the data either: it is fairly clear that the incidence of non-literal uses of *literally* is very low, and has remained so for around two centuries. Finally, since *literally* is undergoing a slow process of change, which is at present perceived as an increase of variation in its meanings, what usage guide writers might be reacting to is the high number of dual uses of *literally*. In these uses, *literally* not only retains its literal meaning, but it also performs an intensifying function within an utterance. It may be these uses which contribute to the high salience of this feature, resulting in metalinguistic awareness and proscriptive commentary.

6.5 Negative concord

Negative concord is a non-standard vernacular feature, and this seems to be reflected in the very low frequency with which it is found in both COHA and COCA. The analysis of this feature was carried out on the basis of cases of negative concord with the three indefinites *no one*, *nobody*, and *nothing*. The use of Python scripts to identify and extract such occurrences in the corpora (see Appendix C for a description of the procedure) resulted in a dataset on the basis of which the frequency distribution of this feature is plotted in Figure 6.16. The figure shows that the normalised frequency of negative concord constructions with *no one*, *nobody*, and *nothing* is somewhat higher in COHA than in COCA; for COCA, the frequency has remained close to zero for the greater part of the last two and a half decades or so.

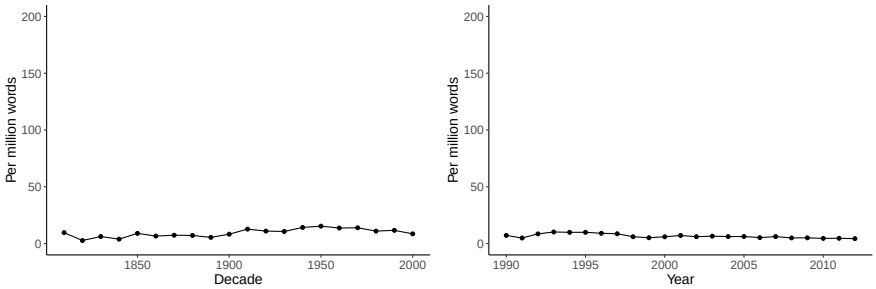

Figure 6.16: Text-linguistic frequency of negative concord across time (COHA: n = 3,912; COCA: n = 2,917)

The variationist analysis of negative concord identifies the proportion of negative concord with the three indefinites *no one*, *nobody*, and *nothing* of all potential uses of negative concord, by contrasting instances of negative concord with those of single negation with the indefinites *anyone*, *anybody*, and *anything* (cf. Nevalainen 2000). The results of this analysis are presented in Figure 6.17. The figure shows that the feature is not common in standard American English, with about 4% of all potential cases in both corpora being realised with negative concord. It is, however, worth noting that the three different indefinites exhibit slightly varying ratios of negative concord: cases of negative concord with *nobody* are found on average in 7.4% of all possible occurrences, compared to 4.2% for *nothing* and 2.4% for *no one*.

Turning to the examination of potential genre effects, Figure 6.18 shows the frequency per million words of negative concord across corpus sections. The

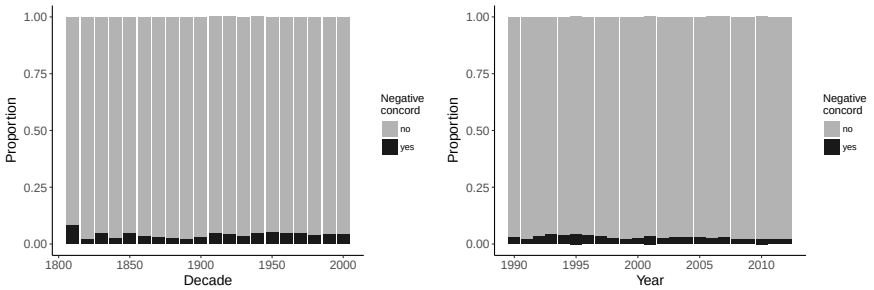

Figure 6.17: Proportion of occurrences of negative concord (COHA: n = 3,912; COCA: n = 2,917) across time out of the total number of environments for negation with the indeterminates *anything*, *anyone*, *anybody* (COHA: n = 91,165; COCA: n = 91,436)

frequency is indeed very low; while, like *ain't*, negative concord is limited to use in fiction in COHA, and fiction and spoken in COCA, the frequencies are lower than those of *ain't*. The feature is clearly not frequent in standard American English, and its uses are non-standard and limited to particular genres which are stylistically varied enough to contain higher levels of frequency of the construction.

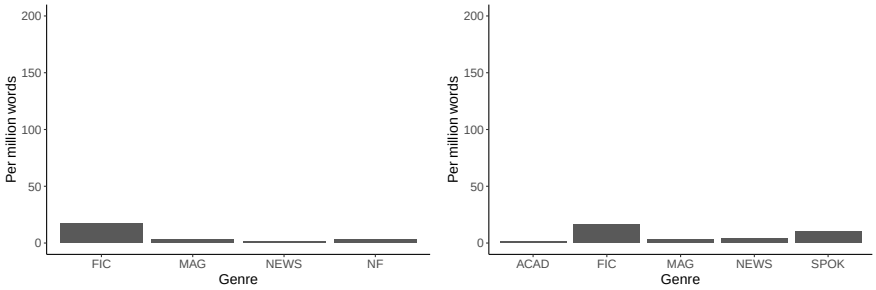

Figure 6.18: Text-linguistic frequencies of all occurrences of negative concord across corpus sections (COHA: n = 3,912; COCA: n = 2,917)

Negative concord is a very rare feature in edited standard American English. In this respect, it is fairly similar to *ain't*, with the difference that the frequency of *ain't* is higher in fiction than that of negative concord. The results are not surprising, given that the feature indeed disappeared from standard English during the seventeenth century (Nevalainen 2000; Tieken-Boon van Ostade 2008a; see also Nevalainen and Raumolin-Bunberg 2003), but remained a feature of the vernacular in both British and American English. What is interesting, however, is the significance of these results in

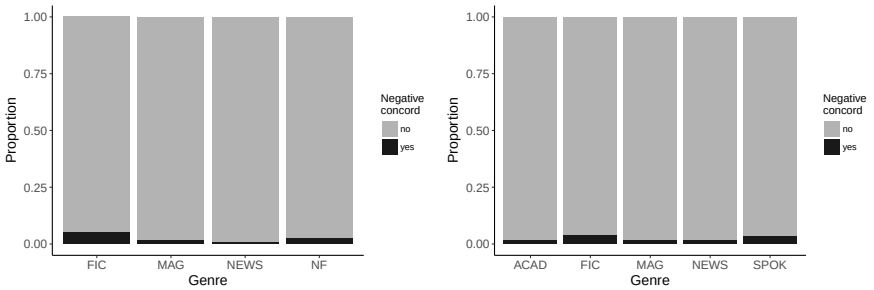

Figure 6.19: Proportion of negative concord (COHA: n = 3,912; COCA: n = 2,197) across corpus sections out of the total number of environments for negation with the indefinities *anybody*, *anyone*, *anything* (COHA: n = 91,165; COCA: n = 91,436)

the context of the treatment of this feature in usage guides. As I discussed in Section 5.2 above, negative concord is one of the features which is least frequently covered in the usage guides consulted. This may indicate that its frequency of occurrence is low in standard American English, and it is consequently not seen as a usage problem. Furthermore, the case of negative concord may provide evidence for the relationship between usage guides and frequency of use. As Ilson (1985) observed, a usage problem is usually a linguistic variant which has a high enough frequency of occurrence in order to be salient enough to be a usage problem. The reverse process might be taking place in the case of negative concord: the less the feature is used in standard American English, the less it will be treated in usage guides. On this basis, we could possibly even predict that negative concord is on its way out of the usage problem canon.

6.6 Pronouns in coordinated phrases

The proscribed forms of pronouns in coordinate phrases are also fairly low in frequency. On the basis of text-linguistic frequency, object *I* is somewhat less frequent than subject *me*, as shown in Figures 6.20 and 6.21, respectively; this difference, however, is not large, as the fluctuations in frequency for both features do not exceed 2.5 occurrences per million words.

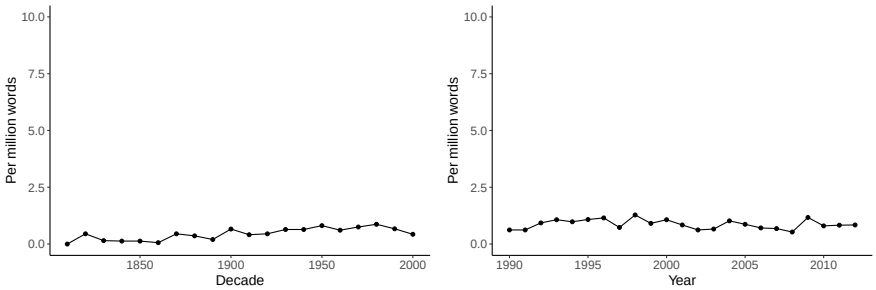

Figure 6.20: Text-linguistic frequency of object *I* across time (COHA: n = 194; COCA: n = 380)

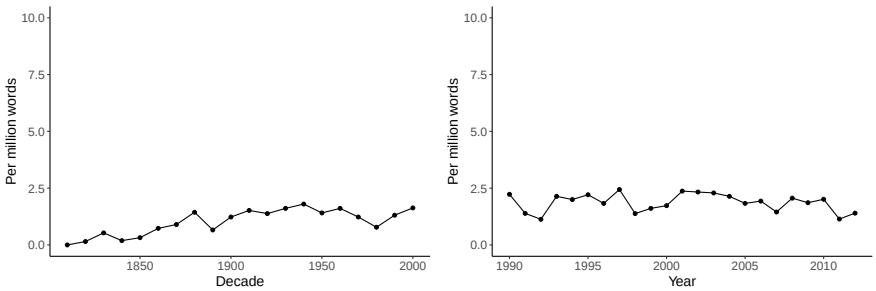

Figure 6.21: Text-linguistic frequency of subject *me* across time (COHA: n = 456; COCA: n = 819)

The results from the variationist analysis, given in Figures 6.22 and 6.23, partly support the text-linguistic frequency patterns. The proportions of both variants in relation to their standard counterparts are fairly low in the two corpora. There is one difference here with respect to the results from the text-linguistic analysis. While on the basis of the text-linguistic frequency distributions the occurrence of subject *me* is slightly higher than that of object *I*, especially in COCA, the variationist analysis shows that object *I* is more frequent than subject *me* when we take into account the total number of possible environments of each of the variants. This may indicate that while neither variant is very frequent in standard American English in terms of rate of occurrence, subject *me* is less often used in all possible environments compared to object *I* because it is seen as a more serious mistake. Object *I* is a well-known case of hypercorrection, and is considered a mark of formality. This difference between the two variants may account for the fact that the variationist analysis shows that object *I* is used more often than subject *me*.

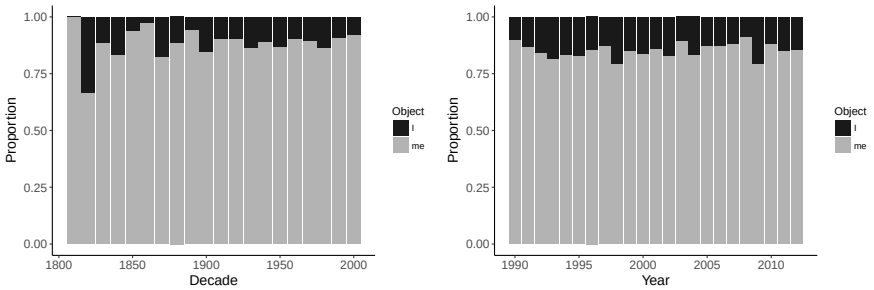

Figure 6.22: Proportion of object *I* (COHA: n = 194; COCA: n = 380) out of all possible environments across time (COHA: n = 1,808; COCA: n = 2,621)

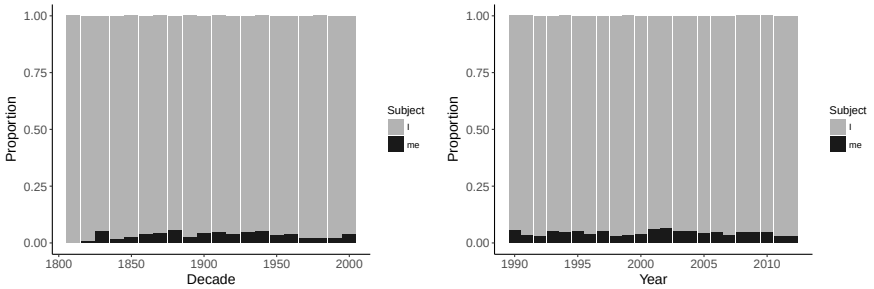

Figure 6.23: Proportion of subject *me* (COHA: n = 456; COCA: n = 819) out of all possible environments across time (COHA: n = 12,087; COCA: n = 17,546)

In addition to these analyses, I carried out further analysis on a portion of this dataset, focusing on cases of pronouns in coordinated phrases headed by *between*. The analysis consisted of manually disambiguating between cases with *between x and I* and cases with *between x and me*. It is important to note here that while the analysis based on all cases of object *I* and subject *me* were restricted to cases where the pronouns are used with a proper noun, the analysis of cases of *between x and I* and *between x and me* was carried out on the basis of all occurrences of the phrase, not only those with proper nouns. The phrase *between you and I* is the most commonly mentioned one in the entries on object *I*; consequently, the analysis considered this specific case in more detail, and explored the extent to which observations about this feature made in usage guides relate to patterns of actual use. In addition, this manual analysis was done in order to gain more reliable insights into the distribution of this feature, which is not possible to the same extent with automatic disambiguation. The dataset analysed is small enough for variants to be manually disambiguated, enabling

us to gain an insight into the frequency distribution of one particular proscribed variant, which features strongly in discussions of object *I*.

The results from this analysis are presented in Figure 6.24. The frequency distribution shows that the variant *between x and I* is very infrequent. This in turn suggests that objections to the use of *between x and I* identified in the usage guides analysed do not relate to any evidence that the phrase is used frequently. The COHA data show that the proscribed variant is barely found during the nineteenth and twentieth centuries. The proportion of uses of *between x and I* out of the total number of possible environments is slightly higher in COCA, but this is a far from striking difference. On the whole, then, the variant is very infrequently used.

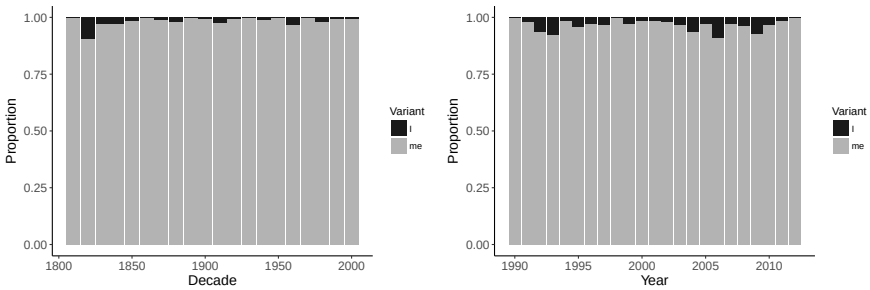

Figure 6.24: Proportion of object *I* and *me* in cases with *between* across time (COHA: n = 27; COCA: n = 44)

Turning to the distribution of object *I* and subject *me* across corpus sections, the pattern which emerges with respect to object *I* is expected. While the frequencies are overall very low, the COHA data have a slightly higher frequency of object *I* in fiction, while the spoken section in the COCA data contains most cases of object *I*, followed by fiction. This shows both that object *I* is infrequently found in general American English, and that when it is used, it is restricted to spoken registers. Of the written registers, fiction comes closer to colloquial text types, so these results are not surprising. As for subject *me*, the pattern of frequency distribution is similar to that of object *I* in the data from COHA, with fiction texts containing the highest rate of occurrence of subject *me*. In the COCA data, however, the situation is more striking. While object *I* seems to be most common in spoken texts, followed by fiction, subject *me* is most often used in fiction, while its use in the spoken sections is not higher than that in magazines or newspapers. This could perhaps be explained in part by the composition of the corpus sections in more detail. The spoken section of COCA contains spoken texts taken from television programmes, which means that

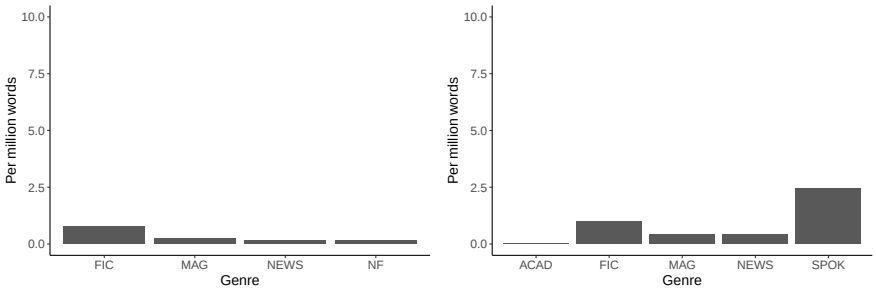

Figure 6.25: Text-linguistic frequency of object *I* across corpus section (COHA: n = 194; COCA: n = 380)

the language found in this section may not be as colloquial and informal as what one would expect to encounter in everyday colloquial settings. More specifically, when it comes to proscribed variants such as subject *me*, speakers in these contexts may have a tendency to avoid such uses altogether, which might explain the relatively low frequency of subject *me*. Fiction, on the other hand, contains a fair number of film scripts alongside novels and other fiction texts, such as short stories. The language in film scripts can be expected to be affected less by prescriptive norms than the language used by speakers in at least some television programmes. This might account for the higher frequency of use of subject *me*. While this may explain the distribution of subject *me* across corpus sections, it does not provide a satisfactory explanation for the difference in the patterns of occurrence of object *I* and subject *me*. This is, I believe, related to the difference in the features themselves. The fact that object *I* is more frequent in standard spoken data may suggest that it is indeed seen as a less serious error than subject *me*, whereas subject *me* is considered to be characteristic of very informal colloquial language use, and is consequently more frequent in fiction texts, including movie scripts.

The variationist analysis of object *I* and subject *me* reveals that while the text-frequencies of object *I* (Figure 6.25) are lower than those of subject *me* (Figure 6.26), the situation is reversed when we look at the proportion of uses of the two variants out of the total number of possible environments. While the prescriptively targeted variant object *I* appears to be most frequent in the spoken sections of COCA, it is also relatively frequent in the academic section, which might be seen as unexpected, given that academic texts are usually heavily edited, and proscribed variants would be expected to be rare (Figure 6.27). The same goes for its distribution across sections in COHA, where the magazine, newspaper, and non-fiction sections

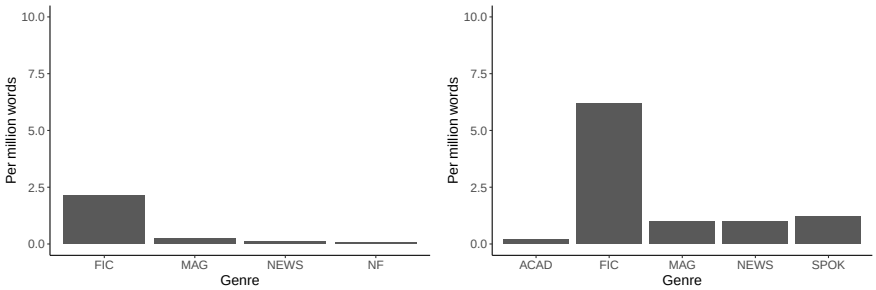

Figure 6.26: Text-linguistic frequency of subject *me* across corpus section (COHA: n = 456; COCA: n = 819)

contain the highest rates of object *I*.

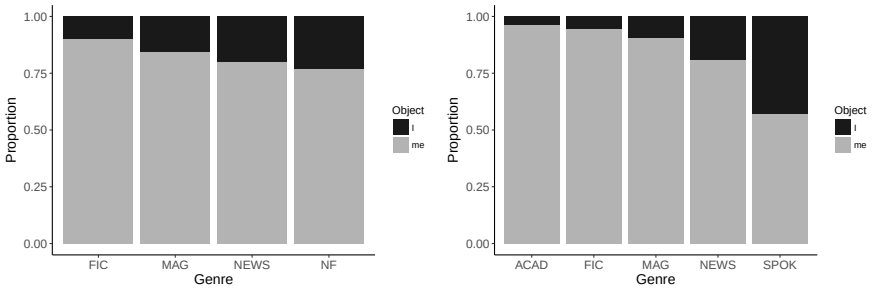

Figure 6.27: Proportion of object *I* (COHA: n = 194; COCA: n = 380) out of all possible environments across time (COHA: n = 1,808; COCA: n = 2,621)

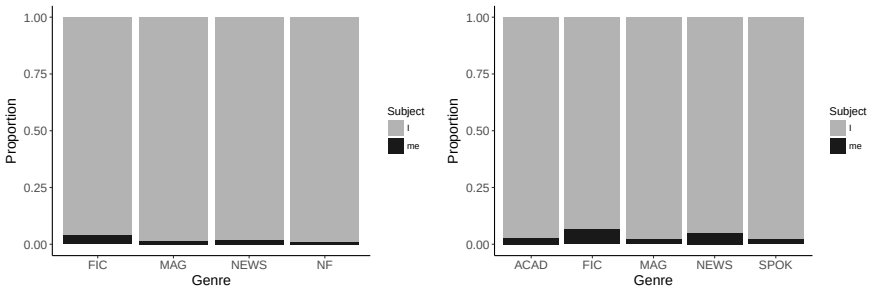

Figure 6.28: Proportion of subject *me* (COHA: n = 456; COCA: n = 819) out of all possible environments across corpus section (COHA: n = 12,087; COCA: n = 17,546)

Finally, the results from the variationist analysis of *between x and I* across sections of the corpora, shown in Figure 6.29, indicate that *between x and I* is most commonly

found in the spoken and fiction sections. However, the proportion of the uses is still relatively low.

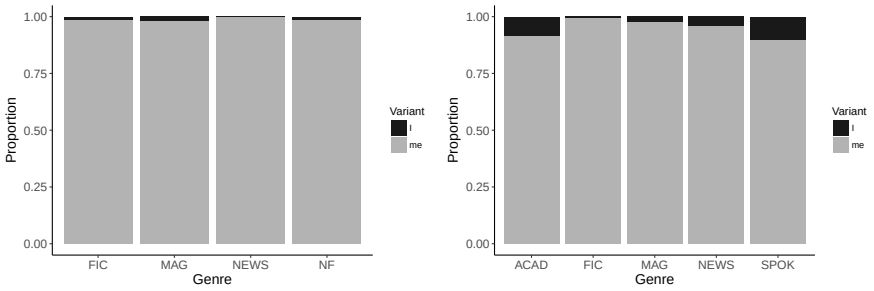

Figure 6.29: Proportion of *between x and I* (COHA: n = 27; COCA: n = 44) out of the total possible environments (COHA: n = 2,027; COCA: n = 1,362) across corpus section

In summary, both object *I* and subject *me* are very low in frequency in both COHA and COCA. In the data presented, we do not see any clear evidence of change in usage over time. This is also the case with the special case of this feature, *between x and I/me*. While both variants are very infrequent across time as well as across corpus sections, there is an important difference between the patterns of occurrence across time and across corpus sections. With respect to the former, the frequencies of both object *I* and subject *me* are very low, and there are no discernible patterns of change across time. It is important to note that the stability of the frequencies over time also indicates that the variants are not disappearing from the language. Furthermore, if the low frequencies of the features are in part a consequence of the fact that the corpora represent relatively standard language, it can be assumed that both variants are more frequent features of spoken language. With respect to the patterns of occurrence of the variants across sections of the corpora, the evidence suggests that object *I* is more often used in more standard or more formal colloquial registers, while subject *me* is more often used in more informal colloquial registers.

In the context of the coverage and treatment of these variants in usage guides, it seems that object *I* and subject *me* are rather straightforward cases of 'old chestnuts'. The stability of their frequencies suggests that they are rare, but possible variants, and are mostly used in informal colloquial speech. The fact that they continue to be included in usage guides indicates that they are still considered problematic, which explains the high number of RESTRICTED entries for object *I* and UNACCEPTABLE entries for subject *me*. In other words, usage guides are not reacting to an increase in

frequency of usage of these variants. Rather, they may be reacting to register variation in the use of these variants, or they may simply be perpetuating prescriptions on the basis of an established tradition, in a way which does not consider contemporary evidence from actual language use. In part, this could be considered evidence that usage guide writers indeed do not always distinguish between spoken and written levels of usage.

6.7 The split infinitive

In this section I present the analysis of the split infinitive, the final feature investigated on the basis of text-linguistic and variationist frequencies. There is a difference in the way these two frequencies were calculated. The text-linguistic frequencies were calculated on the basis of the identification of all infinitives split by one word, including *-ly* adverbs, other types of adverbs, and the negator *not*. The variationist frequencies were calculated on the basis of identifying the variable MODIFIED INFINITIVE, which is defined for the purposes of this analysis as any full infinitive modified by a single *-ly* adverb.

The text-linguistic frequency of split infinitives across time is given in Figure 6.30. The data show clearly that the rate of occurrence of the split infinitive has indeed been undergoing an increase; this is especially clear for the COCA data. There is a sharp drop in the trend for the last decade in the COHA data, which is surprising, and rather difficult to explain, because there is not a similar drop in the same decade in COCA. This might be in part a result of the fact that COCA and COHA have a different make-up (COHA is composed of about 50% fiction texts). What is important

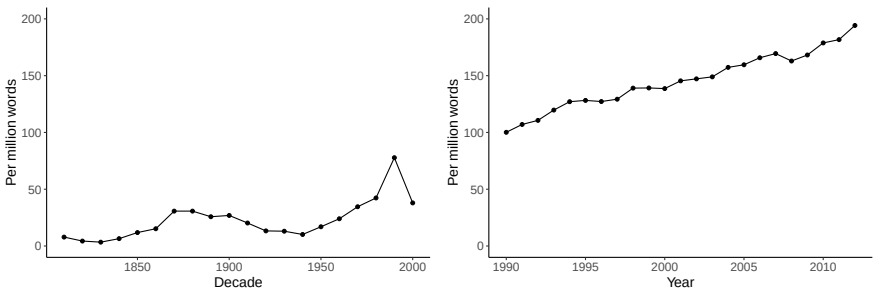

Figure 6.30: Text-linguistic frequency of split infinitives across time (COHA: n = 10,062; COCA: n = 63,079)

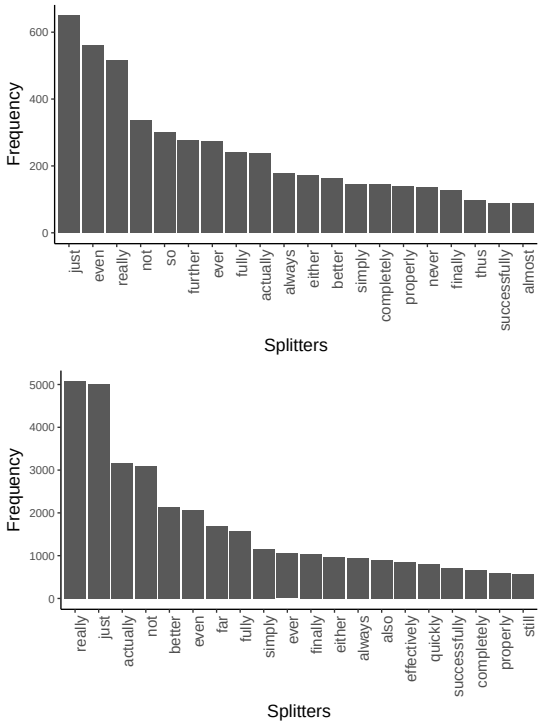

Figure 6.31: Most common splitters in COHA and COCA

though, as I will show below, is that this drop in the data disappears when we take a variationist approach to this feature (see Figure 6.32). This example nicely illustrates the point made by Biber et al. (2016) that normalised text-linguistic frequencies and variationist frequencies often produce differing accounts of the use of a particular variant. On the basis of this, I think it is not unreasonable to assume that the low rate of occurrence of the split infinitive in the last decade of the COHA data may be the result of the types of materials included in the corpus.

In addition to plotting the text-linguistic frequencies of occurrence of the split infinitive, we can perform an analysis on the items which most commonly split infinitives, i.e. the so-called 'splitters'. Figure 6.31 shows that while lexical *-ly* adverbs are the most common splitters, other types of adverbs and the negator *not* are also very common. While it would certainly be of interest to explore all the potential constraints on the occurrence of the split infinitive, including the variation in the use of all splitters, for the present study I limited myself to analysing the proportion of split infinitives (out of the total number of modified infinitives) only in contexts where

the modifier is a lexical *-ly* adverb, as explained in more detail in Section 4.4.

The results from the analysis of the occurrence of split infinitives from the variationist analysis corroborate the increase observed in the text-linguistic frequency of the split infinitive. As Figure 6.32 shows, there is a definite increase in split infinitives over time, though the increase is only small during the second half of the nineteenth century, and is matched by a similar increase in the text-linguistic frequencies. After the middle of the nineteenth century, the trend decreases, and it picks up again after the 1940s. Since then, there has been a steady increase in the use of this feature.

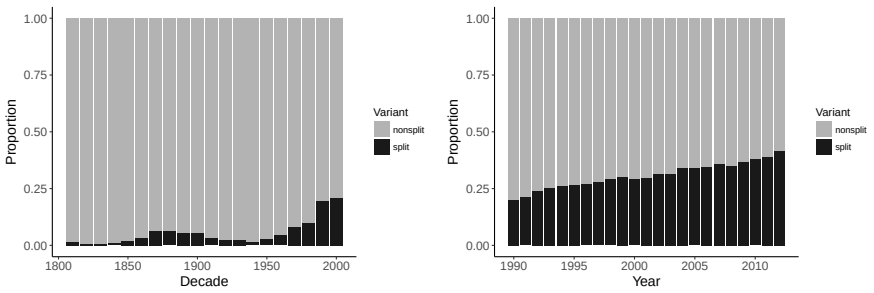

Figure 6.32: Proportion of split infinitives with lexical *-ly* adverbs across time (COHA: n = 6,037; COCA: n = 40,053) out of the total number of modified infinitives (COHA: n = 108,399; COCA: n = 130,855)

The use of the split infinitive across genre sections of the corpora (Figures 6.33 and 6.34) reveals that in the COHA data the rate is much lower; the newspapers section seems to have a slightly higher frequency of split infinitives, but on the whole the frequencies of occurrence of split infinitives in all sections in COHA are low compared to those in COCA. In the data from COCA, the rate of use of the split infinitive varies across sections, with spoken texts containing the highest rate of split infinitives, followed by academic. Magazines and newspapers have more or less equal number of occurrences of split infinitives per million words, while fiction has the lowest frequency of all sections.

The variationist analysis of the proportion of infinitives split by a single *-ly* adverb as opposed to non-split infinitives across corpus sections in COHA and COCA is plotted in Figure 6.34. The proportions in the figure exhibit similar patterns to those based on text-linguistic frequencies, which confirms the text-type distribution of split

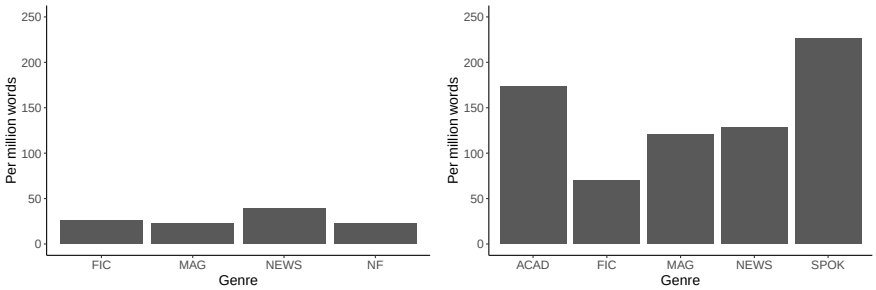

Figure 6.33: Text-linguistic frequency of split infinitives across corpus section (COHA: n = 10,062; COCA: n = 63,079)

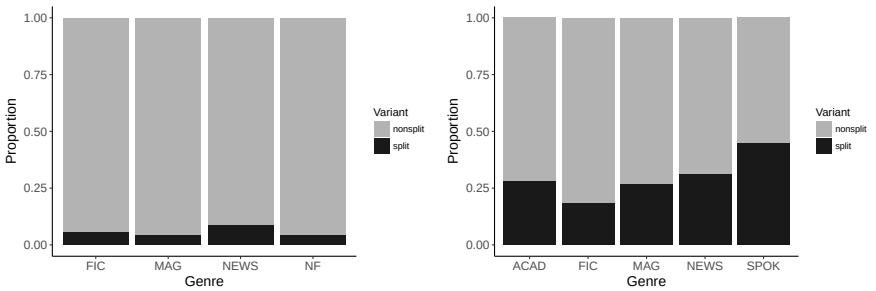

Figure 6.34: Proportion of split infinitives with lexical *-ly* adverbs across corpus sections (COHA: n = 6,037; COCA: n = 40,053) out of the total number of modified infinitives (COHA: n = 108,399; COCA: n = 130,855)

infinitives, i.e. the fact that they are most commonly found in speech. Academic texts show a different pattern, however. While the text-linguistic frequency of split infinitives is higher in academic texts than in magazines and newspapers, the proportional frequencies shown in Figure 6.34 are more or less the same for all three sections.

Turning to the importance of these results for the question of how these trends relate to the usage guide treatment of the split infinitive, as well as the changes observed in that treatment, this case presents us with two possible scenarios. First, on the basis of the increase in the use of the split infinitive after the 1950s, it might be argued that the split infinitive has increased despite prescriptive pressures against its use. This is an observation which has been made in previous studies (e.g. Calle-Martín and Miranda-García 2009; Leech et al. 2009). However, the analysis of treatment of the split infinitive discussed in Section 5.3 suggests that the treatment itself has

started to change, and is becoming more accepting of the split infinitive. In fact, Albakry (2007) has also shown, on the basis of a smaller set of usage data and style guides, that, compared to the other usage features he looked at (sentence-initial coordinating conjunctions, stranded prepositions, functional shift, and modifying absolute adjectives), the split infinitive is not a strongly dispreferred feature. This brings me to the second scenario, in which we might consider the increase in the use of the split infinitive to be a consequence of the loosening of the stricture against its use.

The problem remains, however, of the impossibility of explaining this kind of increase in the use of split infinitives in terms of a weakening of prescriptive influence only. What we can observe here are two separate trends: one, in prescriptive literature, of loosening the prescription against the split infinitive, and the other, in the actual usage observed here, of increasing patterns of use of the feature. Again, as in other cases, this can be interpreted in three ways: first, prescriptivism influences usage; second, prescriptivism is influenced by usage; and third, there is no connection between these two whatsoever, and the observed change is coincidental. In addition to these three possibilities, it is important to consider a fourth one, which is a combination of the three possibilities, which are not mutually exclusive. However, in order to investigate this case further, and in the hope of gaining a better understanding of the level at which prescriptivism might affect language use, I conducted a multifactorial analysis to explore the extent to which the use of other proscribed features in a text may predict the use of one proscribed feature.

6.8 Identifying prescriptive influence at the textual level

Having explored the evidence for potential prescriptive influence, and having applied the traditional approaches in interpreting the trends observed, I now turn to a different approach to investigating prescriptive influence.[3] In the preceding sections, I explored the patterns of use of the six linguistic features investigated in the study, in order to gain insights into how they are used, with the ultimate goal of shedding light on the relationship between usage guides and actual language use. I applied both text-linguistic and variationist metrics in order to obtain more robust evidence for the patterns of use of the linguistic features investigated. While this approach revealed interesting and relevant aspects of the relationship between usage guide

[3] A version of this section also appears in Kostadinova (forthcoming).

treatment of usage and patterns of actual use, it still presents us with the challenge of ascertaining prescriptive influence in a manner that goes beyond the difficulty of equating correlation with causation. We also saw that prescriptive influence, even if it exists, is crucially conditioned by a number of other aspects of language use, both linguistic and stylistic, as the innovative study by Hinrichs et al. (2015) has shown. Inspired by their approach to the analysis of prescriptive influence in the use of restrictive relativiser *that*, I adopt and expand this approach by applying it to the analysis of the potential influence of prescriptivism on the use of the split infinitive.

The logic of this approach, as outlined in Section 4.4, is that many details of the variation patterns of a particular variant are lost when we look at corpus sections in terms of time periods or types of texts. Often, choices in usage which may be affected by prescriptivism are made by individual speakers or writers. So, while corpus-based frequency patterns might not on the whole contain any indication of potential influence of prescriptivism, this influence may be more readily identified at the level of individual texts. The level of specific texts thus provides a higher level of resolution at which prescriptive influence can be investigated.[4] The first assumption of this approach, then, is that prescriptive influence can be more meaningfully explored at the level of individual texts. The second assumption is that, if individual texts are influenced by prescriptive concerns for norms and correctness, this will be manifested in the use of many prescriptively targeted features simultaneously, not just one. Applied to the case of the split infinitive, I formulated the following hypothesis: if the split infinitive is influenced by prescriptivism in individual texts, the likelihood that a modified infinitive will be split will be higher in the presence of other prescriptively targeted variants.

There are a number of motivations for choosing to apply this approach to the split infinitive as a case study. The main motivation is the difficulty of arriving at a more decisive understanding of how prescriptivism has or has not affected the use of the split infinitive on the basis of the comparison between precept trends and actual use data discussed at the end of Section 6.7 above. The split infinitive is one of the features which seems to be losing its usage problem status, raising the question of the extent to which this has or has not influenced its use. Pragmatic motivations for focusing on the case of the split infinitive included the nature of the variable, as well as the size

[4] I use "texts" here to refer specifically to segments of language use included in the Corpus of Contemporary American English. In one sense, this is a specific use of the term, because it refers to corpus texts; in another sense, I use the term broadly, to refer both to more traditional types of texts, such as magazine articles, and to language segments which are not traditionally thought of as texts, such as television shows.

of the dataset. With respect to the former, the variants of the split infinitive are fairly straightforward to determine, and are both used widely in all kinds of texts. *Literally*, for instance, was not considered a good candidate for this kind of analysis, because of the difficulty of applying a variationist approach when analysing this feature. *Ain't*, on the other hand, was found to be restricted to specific text types. Similar issues were present for the other features included in this study.

The dataset for the analysis included all cases in which a *to* infinitive is modified by a single *-ly* adverb (see Section 4.4 for an explanation of how the dataset was produced). Each occurrence of a MODIFIED INFINITIVE was classified as either SPLIT, if the *-ly* adverb is placed between *to* and the verb, or NON-SPLIT, if the adverb is placed either before *to* or after the verb. Thus, the realisation of the variant SPLIT as opposed to NON-SPLIT modified infinitives was modelled as a binary choice in a binomial logistic regression model, the selection of which is explained in the next paragraph. Each case of a modified infinitive was classified as either SPLIT or NON-SPLIT; this was the dependent variable. A number of predictors were used in the model, as explained in Section 4.4, including internal predictors, ADVERB TYPE and ADVERB LENGTH; external predictors, YEAR and GENRE, and a number of prescriptivism-related predictors (see Section 4.4 for a more detailed explanation and examples). These predictors are other prescriptively targeted features, whose frequencies of occurrence in each individual text in the corpus were included as predictors in the model. The following language features were used in the model as prescriptivism related predictors: *ain't*, sentence-initial *and/but*, singular *data*, *hopefully*, *these kind/sort of*, plural *less*, the discourse particle *like*, *literally*, negative concord, plural *none*, passives, *shall*, *try and*, and *whom*. For each text in the corpus, I calculated the normalised frequency of occurrence per 1000 words for each of these features (see Appendix C on the extraction of these features from COCA).

The statistical model used to explore the relationship between the occurrence of split infinitives in a text in relation to other prescriptively targeted features was a binomial logistic regression model. The analysis was conducted on the basis of a procedure outlined in Levshina (2015). The best model was selected using backward stepwise selection on the basis of the lowest AIC (Aikake Information Criterion) value. In addition, the function `drop1` was used to check which of the predictors contribute significantly to explaining the variance in the dependent variable. On the basis of both the backward stepwise selection process, and the results on the predictors which significantly contributed to explaining the variance in the data, the model given in Table 6.1 was selected. As the final model shows, a number of

the prescriptivism-related predictors did not survive the model-fitting stage: *these kind/sort of*, plural *less*, *literally*, negative concord, plural *none*, and *try and*. Only the predictors which are significant in explaining the variance in the dependent variable were thus included in the final model. Even though this model did not show significant improvement in the concordance index C compared to a model with all predictors included, the simpler model was selected, and the value for C was considered acceptable (see Levshina 2015: 259). Following the procedure outlined in Levshina (2015: 274), bootstrap validation was applied to the model to check for over-fitting, using the function `validate()` in the package `rms` (Harrell 2018). The model was refitted 200 times, and the optimism scores were low for all the goodness-of-fit statistics, indicating that the model is satisfactory in accounting for the relationship between the variables.

I now turn to an examination of the results for each predictor in the model. Starting from the internal predictors, the model shows that both ADVERB TYPE and ADVERB LENGTH are significant predictors, indicated in Table 6.1 by asterisks. For ADVERB TYPE, the reference level is ADDITIVE-RESTRICTIVE adverbs. This means that the results displayed in Table 6.1 show how the likelihood that a modified infinitive is split differs in cases in which an adverb is, for instance, a DEGREE adverb, as opposed to cases in which it is ADDITIVE-RESTRICTIVE. The results thus show that the likelihood that a modified infinitive is split is higher if the adverb belongs to one of the following four levels: DEGREE, MANNER, STANCE, or TIME, compared to cases in which the adverb is ADDITIVE-RESTRICTIVE. Cases with LINKING adverbs do not significantly predict the likelihood of a modified infinitive being SPLIT.

The second linguistic predictor, ADVERB LENGTH (measured in syllables), is also significant. As already explained in Section 4.4, ADVERB LENGTH is operationalised as the difference in number of syllables between the adverb and the verb in each case of a modified infinitive in the dataset; the variable has three levels: LONGER, if the adverb is longer than the verb; SHORTER, if the adverb is shorter than the verb; and EQUAL, if the adverb has the same number of syllables as the verb. The reference level here is EQUAL. Compared to cases in which the length of the adverb is the same as that of the verb, the odds of an infinitive being split decrease by 0.70 when the adverb is longer than the verb ($p < 0.01$). In other words, if an adverb is longer than the verb, it tends to come after the verb, rather than before. There was no such significant difference for shorter adverbs.

From the external predictors, I analysed YEAR and GENRE. The external predictor YEAR, which is operationalised as a continuous variable, and is associated with the

year of publication of the corpus text in which each case of a modified infinitive was identified, is also significant, and shows that the odds of an infinitive being SPLIT

predictor:level	*b*	OR	*p*	
(Intercept)	−122.00	0.00	< 0.01	***
INTERNAL PREDICTORS				
adverb class:degree	2.07	7.88	< 0.01	***
adverb class:linking	1.00	2.72	0.05	
adverb class:manner	1.21	3.34	< 0.01	***
adverb class:stance	2.05	7.74	< 0.01	***
adverb class:time	1.57	4.82	< 0.01	***
adverb length:longer	−0.36	0.70	< 0.01	***
adverb length:shorter	0.14	1.15	0.33	
EXTERNAL PREDICTORS				
year	0.06	1.06	< 0.01	***
genre:fiction	−0.11	0.89	0.34	
genre:magazine	0.28	1.32	0.03	*
genre:newspaper	0.49	1.64	0.12	
genre:spoken	1.10	2.99	< 0.01	***
PRESCRIPTIVISM PREDICTORS				
ain't	0.46	1.58	0.08	
And/But	−0.03	0.97	0.02	*
data sg.	1.95	7.04	< 0.01	***
hopefully	0.89	2.43	0.09	
like	0.66	1.93	0.01	***
passives	0.04	1.04	< 0.01	***
shall	−0.80	0.45	< 0.01	***
whom	−0.80	0.45	< 0.01	***
SUMMARY STATISTICS		n = 4,925		
		LR χ^2	812.03	
		Pr(>χ^2)	<0.0001	
		df	20	
		R	0.205	
		C	0.729	
		Somer's Dxy	0.45	
observations	4925			
non-split	2873			
split	2053			

Table 6.1: Binomial logistic regression model for the alternation between SPLIT and NON-SPLIT infinitives modified by one *-ly* lexical adverb. Reference level is NON-SPLIT infinitive

increase by 0.06 for each one-unit increase in YEAR. The predictor GENRE is a categorical variable with five levels: ACADEMIC, FICTION, MAGAZINE, NEWSPAPER, and SPOKEN. The level ACADEMIC was used as the reference level in the model. The

model shows that the likelihood of an infinitive being SPLIT is significantly different for the magazine section and the spoken section, compared to academic. Compared to academic, the likelihood that an infinitive is split increases by 1.32 ($p = 0.03$) in magazine texts. The significance is stronger for spoken texts: compared to academic texts, the likelihood that an infinitive will be split in spoken texts increases by 2.99 ($p < 0.01$).

Finally, the prescriptivism-related predictors show that the significant predictors here are: sentence-initial *and/but*, singular *data*, the discourse particle *like*, passives, *shall*, and *whom*. Of these, sentence-initial *and/but*, singular *data*, the discourse particle *like*, and passives significantly increase the likelihood of an infinitive being SPLIT. In other words, in texts in which these features occur, for every one-unit increase in the frequency of these features, measured as normalised frequency of occurrence of the relevant feature per 1,000 words, the likelihood of an infinitive being SPLIT increases. The statistical significance is the weakest for sentence-initial *and/but* ($p = 0.02$), while all other features are statistically significant predictors ($p < 0.01$). The highest increase in the odds that an infinitive is split is predicted by the occurrence of singular *data*; for each one-unit increase in the normalised frequency of singular *data*, the odds of an infinitive being split increase by 7.04 ($p < 0.01$). The other two significant prescriptivism-related predictors affect the likelihood of an infinitive being split in the opposite direction. For every one-unit increase in the normalised frequency of *shall*, the odds of an infinitive being split decrease by 0.45 ($p < 0.01$); the same result was obtained for *whom*. *Hopefully* and *ain't* are not significant predictors for the use of split infinitives.

Checking the model for interactions showed that the most interesting significant interaction is between YEAR and GENRE. Figure 6.35[5] shows the change in the odds of a modified infinitive being realised as split (as opposed to non-split) per one year for each GENRE level in the corpus separately. As evident from the plots, the change in odds across YEAR is different for the different GENRE levels. The figure shows that the odds that an infinitive is split decrease over the course of the period between 1990 and 2012 in the newspaper section of the corpus, while they increase in all other sections. Most interesting here is that the increase in the odds of a modified infinitive being realised as a split infinitive seems to be greatest in academic texts. There were a number of other interactions, but they did not produce any differences in direction, just in the size of the effect. Consequently, I will not discuss them here in further detail.

[5]This plot was produced using the `visreg` package in R (Breheny and Burchett 2017).

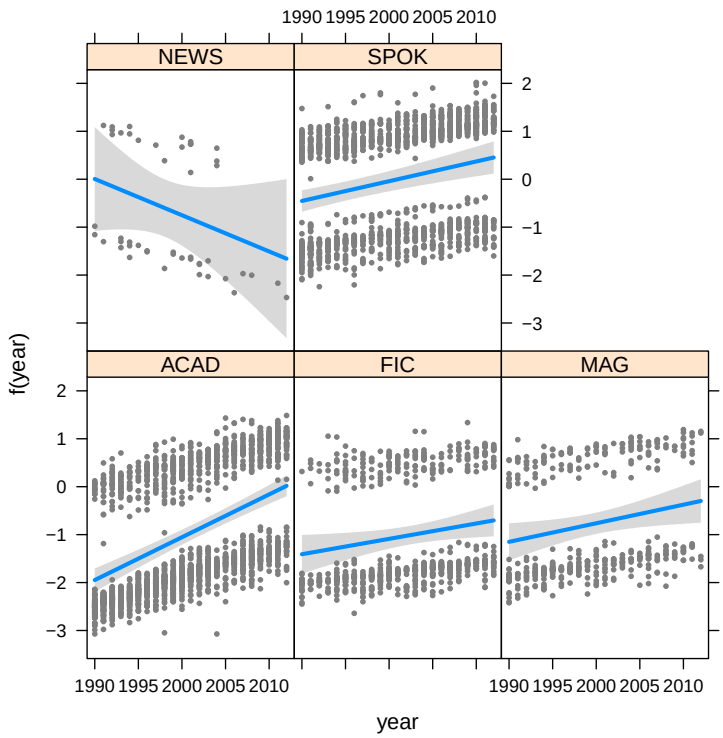

Figure 6.35: Interaction between genre and year

What do these results reveal about what constrains the use of split infinitives as opposed to non-split modified infinitives, and the role of prescriptive ideology in that context? As the binomial regression model shows, texts in which sentence-initial *and/but*, singular *data*, the discourse particle *like*, and passives are used would be less influenced by prescriptive strictures, and would consequently be more likely to contain split infinitives. In other words, writers or speakers who are not concerned about using, for instance, singular *data*, would also be unconcerned about using split infinitives. On the other hand, texts in which authors (or editors) use *shall* and *whom* would be texts in which split infinitives are less likely to be used. *Ain't* is not a significant predictor, because unlike all the other predictors, which belong to stylistic prescriptivism, *ain't* belongs to standardising prescriptivism (cf. Curzan 2014: Chapter 1). Thus, the choice of *ain't* over *be not* or *have not* forms is affected by a different set of considerations, which have to do with following a standard grammatical norm, rather than stylistic

preferences. This of course does not apply to the case of *hopefully*, which Curzan (2014: 33–34) argues is an example of a feature related to stylistic prescriptivism, so in this case the explanation used for *ain't* does not hold for *hopefully*. It can be hypothesised that one possible reason that *hopefully* does not significantly predict the likelihood of an infinitive being split could be that *hopefully* itself is not affected by prescriptive ideology. In any event, this is something which remains to be further investigated in the future.

The interaction identified in the model provides further evidence of how likely the split infinitive is to be used in different types of texts. The most striking finding here is that the increase in the likelihood of an infinitive being split is greatest in academic texts; this suggests that the change towards more split infinitives is led by its use in academic language. Since the increase of the likelihood of an infinitive being split can also be identified for fiction, magazine, and spoken texts, it is reasonable to expect that split infinitives will increasingly be used in those types of texts as well. On the other hand, in newspaper texts, the odds of an infinitive being split decrease across the time period studied, as shown in Figure 6.35. An issue with relying too much on this finding is that the confidence intervals are fairly large, and the level NEWSPAPER was not significant in the model discussed above. Any interpretation would thus have to be made tentatively. This is an indication, albeit weak, that the newspaper genre might still be influenced by stylistic prescriptivism.

6.9 Conclusion

A number of observations can be made based on the results of my analysis of the six features separately, as well as on the results taken together. First, with respect to the six features separately, perhaps the most surprising result is the decrease in the frequency of use of *ain't*. The results do not bear out our original assumption that the increased acceptability of this feature will result in an increase in use. What seems to have happened is that the public discourse on *ain't* may have affected the frequency of use much more than the discussion of this form in usage guides. This case shows that the ways in which usage guides respond, if they do so at all, to ongoing changes in language use are different for different features. The other interesting case is the use of the discourse particle *like*, which is a clear-cut case of prescriptivism responding to a highly salient language change. In this case, it is highly unlikely that we will see a strong influence of prescriptivism on the use of this feature. *Literally* does not

seem to be affected by prescriptivism either. The proportion of the non-primary uses of *literally* has increased over time. The case of *literally* also shows that the reactions to its non-literal use, which tend to exaggerate the frequency with which the feature is actually found, do not seem to be based on empirical evidence. The use of negative concord is a stable non-standard feature, and here there is little change both in how it is used and how it is treated in usage guides. My analysis of the use of pronouns in coordinated phrases shows that object *I* and subject *me* are predominantly restricted to informal contexts, as the proportion of these variants in the corpus data was not very high. The corpus data provided some interesting evidence that the notion that object *I* and subject *me* are problematic in a different way, as shown by the difference in their treatment in usage guides (see Section 5.3), may be borne out by corpus evidence. Finally, the split infinitive is a complicated case, which presents us with the difficulty of ascertaining prescriptive influence by relying solely on a comparison between precept and practice. This kind of comparison for the split infinitive confronted us with more than one possible interpretation of what may be the case in reality. The novel approach applied to the analysis of multiple possible factors constraining the use of the split infinitive showed that the split infinitive is a stylistically prescriptive feature which seems to be favoured in some cases and disfavoured in others. Academic texts, which perhaps tend to be less stringent when it comes to stylistic prescriptivism, seem to be promoting the change towards split infinitives. Other text types, however, may not follow the same trend. While more research certainly needs to be done for this finding to be corroborated for other text types, I believe the results show the complex and dialectic nature of the interplay between prescriptivism and actual use. In other words, in the long run split infinitives may certainly be expected to continue to be used (and critics of prescriptivist efforts may use this case as yet another example of the failure of prescriptivism). However, at present the use of the split infinitive may still be constrained by prescriptivist concerns, and this may be especially true in the context of specific text types.

All in all, the results show the complicated nature of the relationship between prescriptivism and actual usage, which prevents us from making generalisations based on individual features alone. It appears that for some features, such as *ain't* and negative concord, prescriptivism may have an influence over a longer period of time, but these features are non-standard and highly stigmatised. Even in these cases, the usage guide tradition alone may not have a strong influence if it is not backed up by a public discourse denouncing these features, as well as the educational system, through which non-standard features are regulated. In these cases, we see an example of what

Curzan (2014) calls standardising prescriptivism. In the case of the split infinitive, which is a typical stylistic feature, the influence of prescriptivism is of a different nature, and may be restricted to individual cases – speakers or texts – but not at the level of the language system. Finally, the case of the discourse particle *like* is perhaps the most recent example of how prescriptivism can respond negatively to changes in the language, which is one of its most distinguishing characteristics. Even though it is questionable whether over time prescriptivism will have an effect on the use of the discourse particle *like*, this remains to be seen.

CHAPTER 7

Speakers' attitudes to usage in American English

7.1 Introduction

This chapter presents the analysis of speakers' attitudes to the use of the six linguistic features selected for the present study: *ain't*, the discourse particle *like*, *literally*, negative concord, object *I*/subject *me*, and the split infinitive. I explained in Chapter 1 that the inclusion of a study of speakers' attitudes is crucial to understanding the influence of prescriptivism. The most important point in this respect is that when talking about prescriptive influence we need to consider both its influence on the language and its influence on speakers, because it is possible for prescriptivism to have no measurable influence on language use, while nevertheless influencing speakers. In addition, speakers' prescriptivism-related attitudes may not always necessarily result in changes in language use, but may be influential in terms of how they evaluating both themselves and other speakers.

Ideally, prescriptive influence would be investigated by studying the language practice of specific speakers, with a focus on the way in which prescriptively targeted features are used. Such a study could, for instance, involve a very precise definition of language attitudes, an experimental investigation of such attitudes, and the collection of actual language use data produced by the same speakers whose attitudes are studied.

In addition, since the majority of the usage problems are grammatical variables, the language output collected from each speaker would have to be relatively substantial in order to obtain enough instances of each case of a specific usage problem. Given the highly complicated nature of conducting a study of that kind, a more straightforward approach was adopted in the investigation of speakers' attitudes separately from language use data, which in this case was based on the corpus study presented in Chapter 6. While it should be borne in mind that speakers' attitudes are merely reports on speakers' ideas about language, rather than their actual attitudes (which are notoriously hard to tap into), reported attitudes can nevertheless reveal something about what speakers think about the use of specific features. In other words, it may be difficult to find out what speakers' actual attitudes are, but it is less problematic to find out the attitudes speakers think they are expected to have. In the context of attitudes influenced by prescriptive language ideology, this is important to keep in mind.

These attitudes will be analysed and then discussed, in order to arrive at answers to a number of questions. The first question is concerned with the differences in attitudes to the use of the different language features analysed here. By exploring these differences, I hope to provide insights into how attitudes to usage may differ, depending on the usage problem itself. The hypothesis here, in broad terms, is that the usage features which are fairly limited in frequency, such as non-standard *ain't* and negative concord, would be rated more negatively than usage features which are of a stylistic nature (cf. Curzan's "stylistic prescriptivism"), such as the split infinitive. The second question this chapter will address is that of the difference in the ratings across the different levels of language use. These levels, as discussed in Section 4.5, are: CORRECTNESS, ACCEPTABILITY, GOODNESS, and EDUCATEDNESS. Analysing them will serve to explore the different types of attitudes speakers might have when it comes to judging usage problems. In addition, these levels were meant to explore an alternative kind of approach to rating usage problems to the ones which have been used in previous studies of attitudes to usage, such as Leonard (1932), Mittins et al. (1970), and Ebner (2017). By including these four levels, I attempt to show that they reveal a more complex picture of the attitudes to usage among speakers than would be apparent by simply using the notion of "acceptability". The third question related to the difference in attitudes to the usage features is: how does register, understood as context of use, affect the ratings by the respondents? The final question explored here is related to the respondents' social backgrounds, and the way these may have affected the ratings. Here I discuss specifically what the potential influence of these social factors might indicate about the attitudes of speakers, as well as prescriptivism

in general.

The chapter is divided into seven sections; the first six sections cover the six features analysed. The final section discusses the results of a comparison among the linguistic features. For each of the features, I discuss the results of speakers' attitudes on the basis of the data obtained with the survey discussed in Section 4.5. As mentioned there, after each respondent completed the survey, an unstructured post-survey interview followed. Three of the features investigated here were a frequent topic in the post-survey interviews; as a result, for these three features (*ain't*, the discourse particle *like*, and non-literal *literally*) I also discuss some interesting topics which came up in the post-survey interviews. The rest of the features were not discussed in the interviews to the same extent as *ain't*, the discourse particle *like*, and non-literal *literally*, so these data have not been included here.

7.2 *Ain't*

The three sentences containing *ain't* included in the survey are given in Table 7.1 below; as explained in Section 4.5, the sentences were taken from COCA, and slightly modified where necessary to avoid overly complex stimuli sentences. Given that *ain't* is a feature more characteristic of spoken language, two of the stimuli presented were spoken, and the third one was presented in an informal context. These spoken stimuli were recordings of sentences spoken by a male speaker of American English. The context of use for each stimulus was given in the survey as part of the description of what the respondents were about to hear or see; both spoken sentences were spoken by the same male speaker.

Context	**Stimulus sentence**
Spoken informal	I ain't going to see them next month.
Spoken formal	In school they ain't pushing me, they are encouraging me.
Written formal	You won't move forward in your career if you ain't brave enough.

Table 7.1: Stimuli sentences for *ain't*

I already mentioned in the explanation of the survey procedure in Section 4.5 that respondents rated each sentence across the four different levels, i.e. ACCEPTABILITY, CORRECTNESS, GOODNESS, and EDUCATEDNESS, on a five-point semantic differential scale. Having rated each sentence along the four different levels, respondents were then asked whether their ratings were affected by specific words,

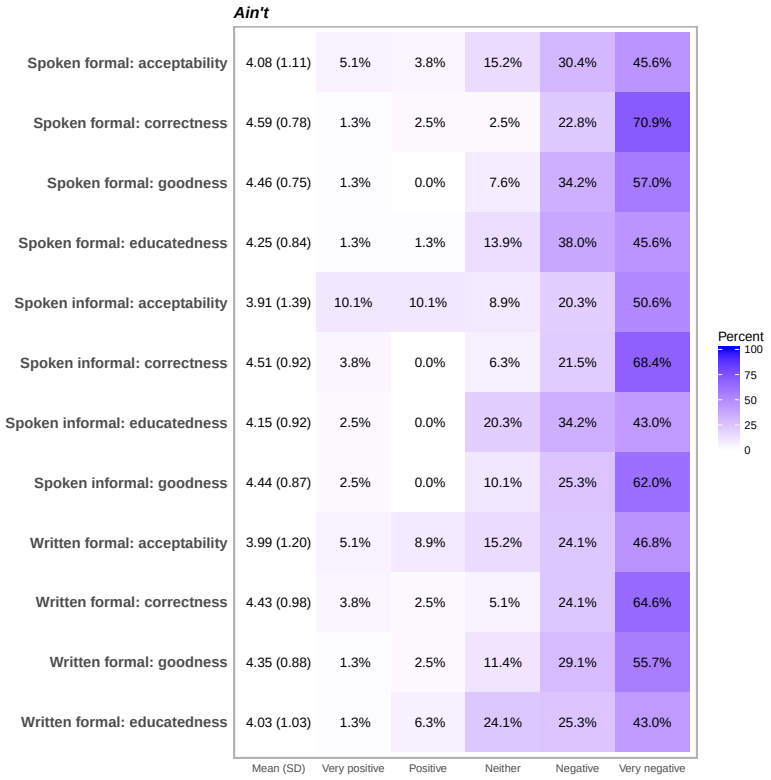

Figure 7.1: Distribution of ratings for *ain't*, n = 79

and, if so, which words. On the basis of these responses, the RECOGNITION LEVEL for each feature was calculated by counting the number of respondents who stated that their ratings were affected by the use of *ain't*. In the context of *ain't*, the RECOGNITION LEVEL differed across the three sentences, but it was fairly high for all three: between 83% and 91% of the respondents explicitly mentioned the use of *ain't* as the reason for the way they rated each of the sentences.

The distribution of the ratings of the three sentences with *ain't* across the four semantic differential scales is presented with the graph in Figure 7.1.[1] The horizontal axis shows the percentage of respondents who selected that particular point on the five-point scale; the vertical axis gives the description for the context of use of each

[1]The graphs were produced using the `Likert` package (Bryer and Speerschneider 2017) in R (R Core Team 2013).

sentence with *ain't*, and the four different levels.[2] The figure shows that the majority of the ratings belong to the 'very negative' end of the scale, across the four levels, i.e. ACCEPTABILITY, CORRECTNESS, GOODNESS, and EDUCATEDNESS, and this is true for each of the three sentences. For CORRECTNESS, more than 60% of the people considered the sentences incorrect. ACCEPTABILITY was rated the most evenly of all the levels, and the ACCEPTABILITY ratings for the 'very positive' and 'positive' points on the scale are highest for the spoken informal sentence.

To investigate these differences more robustly, multiple comparison tests for significance were carried out, in order to (a) identify differences in the ratings across the four levels and (b) identify differences in the ratings for the different contexts. For the first part, pairwise comparisons for all the levels were conducted. The data were tested for normality using the Shapiro-Wilk normality test (Baayen 2008: 73), which confirmed that the data are not normally distributed. For this reason, and because the comparisons are between paired samples, the Wilcoxon Signed Ranks test was used to compare the differences in ratings between various groupings in the data. All the tests were done with the `wilcox.test()` function in R. Since multiple comparisons were conducted, the conventional level of 0.05 was adjusted with a Bonferroni correction by dividing the significance level of 0.05 by the number of tests done for each feature (Baayen 2008: 106). The actual level at which a result was considered significant is given for each feature separately, because the number of tests done per feature differed.

For *ain't* the only significant difference was found between the ratings for CORRECTNESS and those for ACCEPTABILITY for the spoken formal stimulus (W = 4013, Z = -3.505, $p = 0.0004$, effect size = 0.394[3]). In this context, the sentence was found to be more acceptable than correct. For all other comparisons, there was no statistically significant difference between the ratings. This suggests that *ain't* is not seen as unacceptable, but as incorrect.

The effects of context of use were tested with two pairwise comparison tests: one comparing the ratings for the spoken informal sentence and those for the spoken formal one for the four different levels, and another comparing the ratings for the

[2]I have chosen to represent the five points on the scale in the graphs with 'very negative', 'negative', 'neither', 'positive', and 'very positive' for practical reasons; the actual scales used in the survey were based on the four levels, ACCEPTABILITY, CORRECTNESS, GOODNESS, and EDUCATEDNESS, as can be seen in Figure 4.2

[3]There is no consensus on the most appropriate effect size measure for the Wilcoxon Signed Ranks test; I have taken the standardised measure from Cohen (1988) as cited in Corder and Foreman (2009: 40). The conventions for the effect size, which ranges from 0 to 1, are: 0.10 as small, 0.30 as medium, and 0.50 as large.

spoken formal sentence and those for the written formal sentence, again, for the four different levels separately. In this way, the first type of comparison tests for differences in ratings between formal and informal spoken contexts, while the second tests for differences in ratings between spoken and written informal contexts. These kinds of comparisons did not result in any statistically significant differences in the ratings for the different contexts of use, which provides further evidence that the attitudes to *ain't* are fairly negative regardless of context of use.

Finally, testing for differences between the ratings of the two different age groups, gender groups, and ethnicity groups produced no statistically significant differences (see Table 4.10 for an overview of respondents' age and gender). This means that *ain't* is rated equally negatively by all respondents.

I now turn to discussing the insights provided by the post-survey interviews, as *ain't* is one of the features which was explicitly discussed in many of these interviews. The interviews revealed additional information about the attitudes speakers report to have towards this feature.[4] On the more negative end of the spectrum of attitudes, *ain't* is seen as characterising "broken English", as "not proper English" (58) or as "completely unacceptable" (59). The idea that *ain't* is not a word was expressed a few times (60, 61), and some informants also reacted quite viscerally to the word (62).

(58) The ones I rated as lower in education or correctness were the ones where they used *ain't*, because *ain't*, you know, it's not proper English. You understand what the person is saying, but just in terms of the basic structure of English it's not English. So, if I say 'this is a good film' then that's proper English. But if I say 'This ain't a good film' then that's generally not considered proper English. (A, m, 25)[5]

(59) I think I reacted a little bit more strongly to the ... more to *ain't* instead of *aren't*. I don't know why it just sounded completely unacceptable to me. (J, m, 32)

(60) I was not allowed to speak like that growing up. If I tried to say *ain't*, my parents would be like 'No, that's not a word!' (E, f, 19)

(61) Come on, you know *ain't* isn't a word, it just sounds silly when you say it. (R, m, 30)

(62) *Ain't*, um, *ain't* just, it hits me in my solar plexus ... it should not be used at all.

[4]The rest of this section is based on an analysis of attitudes towards *ain't* expressed in the interviews I conducted, published in Kostadinova (2018b).

[5]For each quote, I include the respondent's first-name initial, their gender, and their age.

(J, m, 29)

Milder reactions referred to the ACCEPTABILITY of the word depending on the context of use (63, 64), as well as its ACCEPTABILITY in particular regions of the United States, or with particular groups of people (65, 66). It was generally recognised that as long as it is used for effect in a situation in which there is a basis to do so, *ain't* is not a problematic usage, and does not leave a negative impression.

(63) *Ain't* is one of those things that it's like – as long as the context is informal, then it's fine. (D, m, 37)

(64) Well, it depends in what context, I mean… it could be understood as slang, you know in certain contexts, especially if someone is meaning to sound very casual, very colloquial. (A, f, 23)

(65) It's more acceptable for certain cultures. It doesn't mean that it's incorrect, but it's just different. (M, f, 27)

(66) *Ain't* usually people don't say unless it's in conversation or just if you're in a certain region or somewhere where that's acceptable (C, f, 28)

A number of informants associated the word with lower-class speakers (67) and lower levels of education (68). Some also associated the word with African American or Hispanic speakers (69, 70), and associations with the South were also common (70). The majority of the informants saw no clear or straightforward relationship between the race or ethnic background of the speakers and their use of *ain't*. The region and the economic status of the speakers seemed more of a determinant than race or ethnicity. It is also important to point out that the use of *ain't* does not seem to be perceived as a marker of a particular social class if the person displays the right context-sensitivity about when to use the form (see Section 5.2.1). In other words, *ain't* as a variant may occur in any social variety of American English, but is not believed to do so with high frequency. When it does, it becomes a marker of a specific stigmatised dialect (cf. Wolfram 2004: 65). It is important to note here that this observation agrees with a similar line of argumentation found in a number of usage guides discussing *ain't*. In these guides, it is argued that *ain't* is acceptable in the language of speakers whose status as "educated speakers" is established, and when they use *ain't*, it is evident from the context that they do so for specific reasons (e.g. to be funny, to make a point, etc.). However, the argument goes, speakers who

use *ain't* all the time are likely to be seen as uneducated.

(67) It's a real kind of style thing. It shows where somebody is from or it might show somebody's class; probably like a lesser tier class. Higher-tiered people do not touch that word! (A, m, 33)

(68) I feel like, I have friends who'd say like 'I ain't got time for that' but they're joking cause they're being ridiculous. But if someone actually said that non-ironically, like if someone said that seriously like I would think they were an idiot or like not educated. (A, f, 32)

(69) [...] both African American and Hispanic [use *ain't*] because they pattern off of what they hear. If you hear it you use it. (L, f, 33)

(70) Those things that are aberrant to standard English are I feel like dialects that come from rural areas, like – and that seems really obvious, but words like *ain't* or the double negative tend to come from – or like, I think people associate them with places like the South which also tend to have – and also get associated with I guess African American population as well and that's such a – that is so fraught with the potential for judgements in a way that I don't think needs to be sometimes. (E, f, 34)

A number of speakers also commented on their own use of the word. Some of them gave a more positive account of using the word as something that makes them feel casual and colloquial and in line with the speech norms of their community, and as something about which they have a sense of when to use it and when not to (71, 72). Others commented on avoiding the word *ain't* (73, 74). What is interesting to note here is that all of the examples here come from African American speakers. These attitudes are a clear indication of the covert prestige of this feature.

(71) I'll be honest with you I use ain't sometimes, yeah. I know it's not grammatically correct but sometimes like, you know 'I ain't going there', you know, it's like casual. I wouldn't use it in like a classroom setting, but I use it with like most of my friends. A lot of African Americans you know black people say it so... Me, I say it all the time, I hear it at my job, well I work in a mostly cultural area in LA, so I hear it all the time. (K, m, 22)

(72) Ain't was definitely something that I had to figure out. If I'm in a professional setting those are not words I would use. I hear it a lot, now a lot more, and I find myself as well correcting the children when they say it. (L, f, 40)

(73) I wouldn't really use *ain't* anymore because I was corrected as a kid. ... It

doesn't bother me, so I've come to know what it means. (T, m, 21)

(74) I say it every now and then, but not as consistently as I used to. (P, f, 19)

As a final note with reference to the responses to *ain't* given in the interviews, it may be said that some of them reflect the observation of Bloomfield (1944) about secondary and tertiary responses to language. The first interesting case in this respect is a couple of exchanges I had with speakers on the issue of *ain't* not being a word. The statement that *ain't* is not a word was usually made after I had asked these respondents about their thoughts on *ain't*. The observation that *ain't* is not a real word is thus a secondary response, and it is usually made in an authoritative, confident fashion. A tertiary response is usually a response to the interviewer pointing out that *ain't* actually is a word. Bloomfield notes that "the tertiary response is hostile; the speaker grows contemptuous or angry" (Bloomfield 1944: 49). Although this was not always the case in the interviews I conducted (with the exception of one respondent who said that "if you use *ain't* you are a moron"), the informants did become passionate about proving their point of view about the use or status of a particular form in the language.

7.3 The discourse particle *like*

Attitudes to the discourse particle *like* were investigated on the basis of two sentences included in the survey. Because the discourse particle *like* is almost exclusively used in informal spoken language, the two stimuli were presented as spoken segments from a conversation between friends. The sentences and their contexts are given in Table 7.2. As mentioned in Section 3.4, the discourse particle *like* is often associated with female speakers, hence the decision to include two spoken stimuli which differ in the gender of the speaker.[6]

context	**stimulus sentence**
Spoken informal, M	Didn't you, like, all like go to, erm..., like a boot camp?
Spoken informal, F	I've like done a couple of like summer camps in like languages and accounting.

Table 7.2: Stimuli sentences used for the discourse particle *like*

[6]It should be noted that although a case can be made for the increasing use of the discourse particle *like* in informal online communication, the survey did not include a sentence in this context (but cf. Ebner 2017).

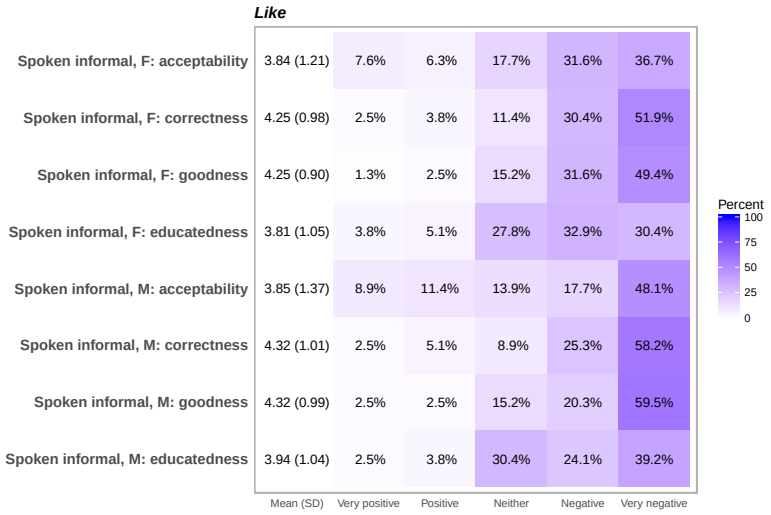

Figure 7.2: Distribution of ratings for *like*, n = 79

The RECOGNITION LEVEL for *like* was relatively high: 77% of the respondents said that their rating of the sentence spoken by the male speaker was affected by the use of the word *like*. In the case of the sentence spoken by a female speaker, the RECOGNITION LEVEL was 93%. The distribution of the ratings for the two sentences, on the basis of the entire set of responses, is given in the heat graph in Figure 7.2. The graph shows that the sentences with *like* are very negatively evaluated. The ratings for ACCEPTABILITY seem to be somewhat more evenly distributed than those for CORRECTNESS and GOODNESS. What is important to note here is that EDUCATEDNESS is rated most neutrally, with about 30% of the respondents rating both stimuli as 'neither educated nor uneducated'.

The Wilcoxon test was used to conduct similar pairwise comparisons to those carried out for *ain't*. The conventional level of significance, 0.05, was Bonferroni corrected by dividing 0.05 by the number of comparisons done for the discourse particle *like*, i.e. 56. None of the pairwise comparisons between different levels was statistically significant.

Testing for the effects of context was not technically possible in this case, because both sentences were presented in the same context, i.e. spoken informal. The only difference tested here was between the ratings for the sentence spoken by a male speaker and those for the sentence spoken by a female speaker. Four such comparisons were done, for each of the four levels: ACCEPTABILITY, CORRECTNESS, GOODNESS,

and EDUCATEDNESS. These comparisons also resulted in no statistically significant differences between the two stimuli.

Finally, with respect to the social variables of the respondents, i.e. age, gender, and ethnicity (see Section 4.5 and Table 4.10), the most significant result for the discourse particle *like*, both statistically and in general, was the difference in ratings between the two age groups, 29 OR BELOW and 30 OR ABOVE. The first group of respondents rated the sentence containing *like* spoken by the male speaker less negatively for CORRECTNESS than the second group (W = 300, Z = −3.512, $p < 0.001$, effect size = 0.611). A similar statistically significant difference was identified between the two age groups for the CORRECTNESS ratings of the sentence with the discourse particle *like* spoken by the female speaker (W = 283, Z = −3.626, $p < 0.001$, effect size = 0.631).

I now turn to the discussion of the discourse particle *like* in the post-survey interviews. The discourse particle *like* was most unequivocally noted as being very frequent, especially in the region where the interviews took place (75). Although respondents stated that they know that *like* may be frowned upon, and were aware of the stereotypes associated with its use, they still noted its high frequency of use and the fact that the word would probably not be affected by the commonly encountered negative attitudes about its use (76, 77, 78).

(75) *Like* is becoming pretty widespread that we add everywhere. (A, f, 23)

(76) *Like* is never going anywhere. Yep. *Like* is here to stay. Especially in California. (J, m, 29)

(77) *Like* is definitely a huge word, we say it all the time, we don't even really think about it until you see it or hear someone else saying it. (A, f, 24)

(78) It's common. I use it sometimes. Historically it's considered like Valley speak, like LA, kind of like ditsy, like you know, so it's – I think it's sort of widespread now. (D, m, 37)

On the more negative side, the typical attitudes expressed towards the use of *like* were that it signals weak language and the inability to speak grammatically correct English (79). However, one informant also noted that among the younger generations of speakers, with whom *like* is most readily associated by the majority of informants, not using *like* may sometimes come across as old-fashioned, and that *like* is becoming acceptable in informal or professional spoken communication (80). This may be

indicative of a growing covert prestige of *like* among the group of speakers that seem to use it most, and with whom it is most often associated.

(79) Coming from the Valley, the people who invented *like* and whatever and *oh my god*, again that falls into weak language – you are trying to delay your point and I don't like it... I just don't like it. It's weak language. (J, m, 19)

(80) I believe it's a little more old-school to not use the word *like*, because professors over 30 would question it, but younger professors I've had, for example my debate professor, he would always denounce the use of *like* during speech, like when it was professional, but throughout regular conversation he would use it, we would all use it and it wasn't stigmatised in that sense. (A, f, 22)

Discourse particle *like* is clearly associated with a set of personal qualities related to absent-mindedness and low intelligence (81), which is probably in turn related to the stereotype that the people who use *like* are "Valley girls", i.e. rich young women who have too much money and time on their hands (82). However, some informants also recognised that such stereotypes do not really hold if they think about their own experience with the word *like*. Thus, the two main associations of the use of *like* were with the region of Los Angeles, or the West Coast more generally (76, 78), and with younger speakers (83), regardless of race or ethnicity. In some cases, however, the discourse particle *like* was associated with white people more than any other ethnicity. A possibly positive association with the word *like* can be found in the observation that *like* represents a fashionable way of speaking, popularised by celebrities (84). Finally, *like* is seen as becoming so widespread that it crosses the boundaries of age and it is becoming ever more prevalent across all age groups (85).

(81) I think for me it conjures up an association with empty-headedness essentially. (E, f, 34)

(82) For example girls who use *like* a lot, most people's stereotype is that they're very air-headed or dumb girls, 'like I don't know like' ... like, the idea you get is that it must be some really dumb-headed girl. (A, m, 25)

(83) *Like*? Um, younger – I don't know why younger. It just makes me think younger, or someone that doesn't have that much to say even though I just said *like*. It's *like* really hard to erase it. (A, f, 24)

(84) I feel like *like* is a fad. Celebrities and people will be like 'like, you know' and they do it on purpose – and it just became a thing and I say it a lot. (B, f, 20)

(85) I think that it's just infiltrated every – almost every age, every community, both genders that, it's become so accepted to the point where anyone says Valley girl anymore anyway . . . it is something you hear in every group of people. (A, f, 32)

There is a clear tendency to stigmatise *like* and to associate it with a set of negative personal characteristics, as well as with a particular social status of white middle class (female) speakers. It is interesting to note in this context that, while Wolfram (2004: 59) observes that "the speech of low-status groups in American society tends to be much more socially marked than that of high-status groups", this is certainly not what we can observe in the context of the attitudes expressed towards *like*. The most interesting finding that came out of the interviews is that *like* may be developing some degree of covert prestige among certain groups of speakers, as evidenced by more positive comments on its use (e.g. 80).

7.4 Non-literal *literally*

Attitudes to non-literal *literally* were explored using three different sentences. Given that non-literal *literally* would be expected to be found more often in spoken language or informal contexts, two of the sentences were spoken informal, and one was written informal. The three sentences used in the survey are given in Table 7.3. Two of the sentences were presented in a spoken informal context, one spoken by a male speaker and the other spoken by a female speaker. I was also interested in exploring the extent to which non-literal *literally* would be rated differently when used by men than when used by women.

context	**stimulus sentence**
Spoken informal, M	I literally died from boredom on my date last night!
Spoken informal, F	There is a story in this book that literally blew my mind!
Written informal	This book literally blew my mind.

Table 7.3: Stimuli sentences for *literally*

The RECOGNITION LEVEL for *literally* was fairly high, at about 70%, but lower than that for both *ain't*, which was between 83% and 91%, and the discourse particle *like*, which was between 77% and 93%, which might suggest that fewer respondents considered *literally* problematic in this context. Further evidence for this observation comes from an analysis of the distribution patterns of the ratings for the sentences with

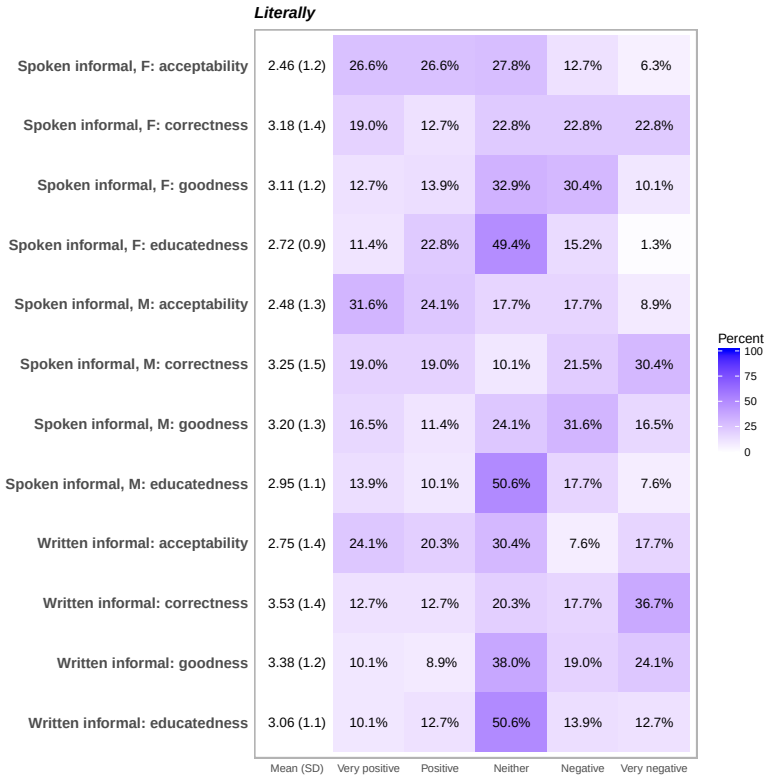

Figure 7.3: Distribution of ratings for *literally*, n = 79

literally. The distribution of the ratings for the four different levels for these sentences is plotted in Figure 7.3. The plot shows that the ratings are not predominantly negative, but are rather differently distributed across the four levels. The ratings for ACCEPTABILITY seem to be predominantly on the positive side of the scale, while those for CORRECTNESS tend to be more on the negative side. GOODNESS and EDUCATEDNESS are predominantly neutrally rated. These patterns for the ratings clearly point to a difference between these ratings and the ratings for *ain't* and *like*, which were more uniformly negative; I will discuss this question in more detail in the final section of this chapter, where I will compare between the ratings for the different language features included in this study.

Looking at the results from the statistical significance testing, I found that only one statistically significant result was obtained, and that was in the difference between the ratings for ACCEPTABILITY and those for GOODNESS for the spoken informal

sentence uttered by a female speaker (W = 2143.5, Z = −3.492, $p < 0.0005$, effect size = 0.392). This difference shows that while non-literal *literally* tends to be rated more positively for ACCEPTABILITY, it is at the same time rated more negatively for GOODNESS. The rest of the ratings were not significantly different. There were no differences between the ratings for the four different levels for the sentence spoken by a male speaker, nor for the written informal sentence.

To determine if differences in the ratings influenced by context of use, two sets of pairwise comparisons were carried out. The first set consisted of a comparison between the ratings for the sentence spoken by a male speaker and the one spoken by a female speaker, for each of the four levels separately. The second set of comparisons was carried out between one of the spoken informal sentences and the written informal sentence. These two were compared across each of the four different levels. There were no significant differences between the different stimuli across the four levels.

Finally, in terms of social variables, differences in the ratings across age, gender, and ethnicity groups were tested. These resulted in no statistically significant differences, which suggests that the social background of informants may not play a role in how non-literal *literally* is rated (see Section 7.9 for further discussion of this issue).

Literally was often mentioned in interviews. The attitudes expressed towards the non-literal use of *literally* range from stronger and more negative reactions to reactions that display a more moderate account of the use. What is striking in almost all of the opinions expressed, however, is the observation that this use of *literally* is quite prevalent, and tends to be associated with a younger generation of speakers, as well as with white Americans. In terms of meaning, people seem to be predominantly aware of its INCORRECTNESS, but at the same time quite attuned to the frequent use of the word as an intensifier. This use is folk-linguistically explained in terms of someone trying to be dramatic or funny when using it. The negative reactions came from people who stated that they are personally bothered by the word, as exemplified in (86, 87) below, and often tend to associate this usage with stupidity (87), immaturity (88), or lack of knowledge about what the word means (89).

(86) Yeah, there it kept saying *literally* – 'it literally blew my mind'. *Blew my mind* didn't bother me. It was the word *literally* that bothered me because if you say 'it literally blew my mind' it means it actually blew your mind and your mind exploded. (M, m, 42)

(87) I hate the misuse of the word *literally* – that just, to me – but I was trying to

think why I hate it and I can't come up with an argument, 'cause I do think it just sounds stupid to use it so incorrectly and it's so prevalent. That is a pet peeve of mine. (E, f, 34)

(88) I think at an unconscious level it just means that they are less of an adult. Because my mother would not say that. There are just certain people who wouldn't say literally. Yeah, it just seems immature. (A, f, 32)

(89) *Literally* is used a lot; I associate it with not knowing what the word means. (K, m, 60)

Most informants, however, gave a more moderate account of the non-literal use of *literally*, and the main argument for why they thought the feature was accepted was that they noticed it was becoming prevalent and more acceptable (90). A number of respondents even said that, despite their awareness of non-literal *literally* being "technically incorrect", they would still use it because the word is so often used, and that generally they are not bothered by it (91, 92).

(90) You know, I don't use that one myself very much and I think people use *literally* so much that I'm probably deaf to it unless it's, once again, egregious misuse. I mean, I remember reading an article not too long ago on the law blog that it's invaded legal script too. But everyone does it now so I think that may be one that's getting more accepted and more normalised. (R, f, 32)

(91) I used to use that all the time but then I felt like it was putting a big emphasis on something. (L, f, 40)

(92) I do that. I know it's not correct English, but I would say *literally* – 'literally this, literally that' – even though it's not technically always correct. (E, f, 19)

Finally, in terms of associations with particular groups of speakers, the majority of the informants stated that the strongest association of non-literal *literally* is with younger educated speakers (93). A number of people also related it to white American speakers (94). Finally, one informant, belonging, interestingly, to the category of young, highly educated white American speakers interviewed, identified this use of *literally* as something that is part of how they talk and as something that shows knowledge of language and ability to use language creatively (95).

(93) *Literally* has been totally abused lately and I don't think people who use it sound as uneducated as people who use *like* just because it's more current, I guess, and

it's a more complicated word than *like*. . . . It's definitely I younger person thing. Yeah, like, like, pre-teens to twenty-four. (C, f, 21)

(94) There were clearly, like, white people mistakes, grammatically, like *literally* and *like* and this and that, and then there were African American mistakes of *ain't* and *got* and. . . To break it out of race, maybe it's more socioeconomic, and what sort of neighborhood you're from, but yeah. (E, m, 28)

(95) The use of the word *literally*, I think of it as somewhat acceptable. 'Cause it's more hyperbole in what you're saying, so I think of it as more of a descriptive term. Yes, it's being used incorrectly, but it's being used in almost a funny way, and to use *literally* is not so erroneous that it's a problem, but it's definitely wrong. Sometimes I think the use of the word *literally* is just funny. I have done it before. A lot of us have. 'I literally wanted to kill myself!' The majority people that I know, especially those I interact with, they use *literally*; they use it a lot. It's almost funny when they use it and it shows more breadth of languages because *literally* is, I mean, to know what *literally* means – It's actually a word that I don't think a lot of the general population, especially people who do not have interesting grammar, would actually know what it means. (J, m, 26)

In summary, the accounts of and opinions about non-literal *literally* which the informants give show a high degree of complexity and awareness of the usage of *literally*, as well as its sociolinguistic variation. Strong opinions are present in some speakers, but generally, despite those strong opinions, speakers seem to be aware of its increasing use. This awareness seems to be the reason for the ACCEPTABILITY of the word, especially in context of its emphatic, dramatic, or humorous use. Its main association is with the language of younger speakers, and among this age group the use of non-literal *literally* does not seem to be related to education or social class as much as with a particular type of mainstream youth culture. Although it is too early to predict future trends for certain, it can be hypothesised that the positive interpretation of intensifier *literally* as exemplified in (95) above, may signal a tendency towards greater general acceptance of the word, as well as towards a potential change in the norms of usage. What this shows seems to be in line with the findings of Ebner (2017) on the attitudes to *literally* among British English speakers, where it is also associated with younger speakers.

7.5 Negative concord

The attitudes towards negative concord were investigated on the basis of the collection of ratings data for three sentences. These sentences were presented in both spoken and written, as well as formal and informal contexts. The sentences containing negative concord, and their contexts, are given in Table 7.4 below.

Context	**Sentence**
Spoken formal	I'm strong minded and I'm not going to let nobody lead me off in the wrong direction.
Written informal	I'm sorry. But I'm not going to argue with nobody.
Written formal	I thanked the good lord that I had not killed nobody.

Table 7.4: Stimuli sentences for negative concord

The RECOGNITION LEVEL of negative concord was about 75%, suggesting that the majority of the ratings for these three sentences were influenced by the presence of negative concord. The distribution of the ratings for each of the three sentences, as well as the four different levels, is given in Figure 7.4.

As can be seen from the graph, negative concord is in the same category as *ain't* and the discourse particle *like*. The ratings are all predominantly negative for each sentence, across the four different levels. The distribution patterns show that CORRECTNESS is most strongly negatively rated, with more than 60% of the respondents rating all sentences as 'very incorrect'. Ratings seem to be somewhat less negative for ACCEPTABILITY and EDUCATEDNESS, but they still remain on the negative side of the scale.

Looking at statistically significant differences between the four different levels for each of the sentences separately produced only one significant result. A statistically significant difference was identified between the ratings for ACCEPTABILITY and those for CORRECTNESS for the written informal sentence ($W = 4100.5$, $Z = -3.680$, $p < 0.0004$, effect size = 0.414). The rest of the pairwise comparisons between different levels did not result in any statistically significant differences.

Context of use and the social variables included in the survey were found not to result in statistically significant differences in ratings either. In other words, sentences with negative concord were rated negatively across the four different scales, and these ratings were not affected by context of use or the social background of the respondents.

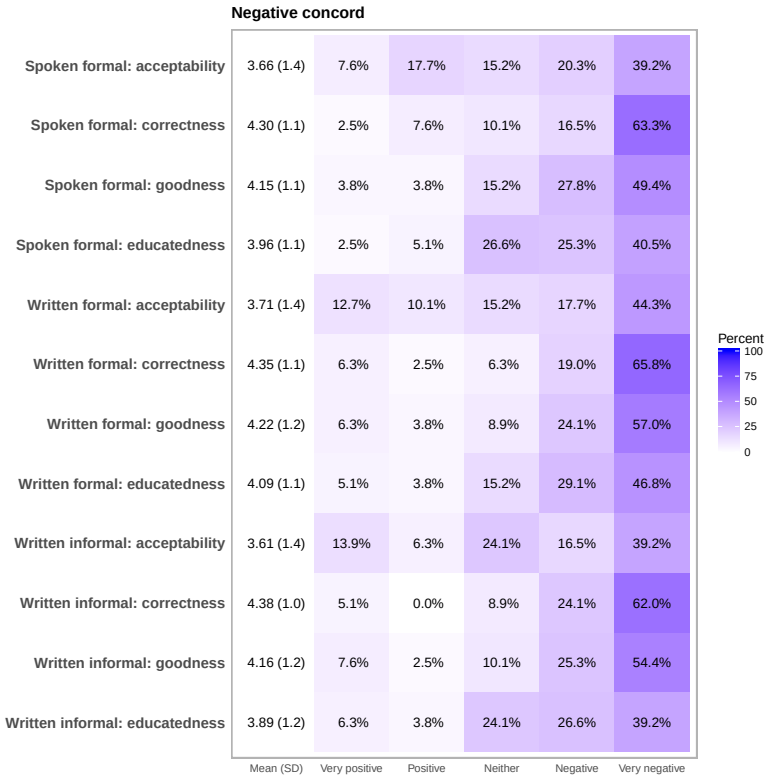

Figure 7.4: Distribution of ratings for negative concord, n = 79

7.6 Pronouns in coordinated phrases

The attitudes to pronouns in coordinated phrases were tested with more sentences than those included for the other features. On the assumption that object *I* and subject *me* are more commonly used in informal language in standard English, for each of these features an additional sentence was included in the survey in which the 'correct' variant is used (see Section 4.5). In this way, an additional analysis was done to test for any potential differences in the ratings between object *I*/subject *me* and their respective 'correct' variants. The sentences for object *I* and subject *me* are given in Table 7.5, where the sentence with the 'correct' variant is marked with 'C'.

The RECOGNITION LEVEL for sentences with object *I* was lower than for the features discussed so far, with a little over 50% of the respondents explicitly stating that their ratings of the sentences containing object *I* were affected by the use of this

Context	Sentence
Object *I*	
Spoken informal	I think this has been the trouble between you and I.
Written informal	This trip has been a great adventure for my parents and I.
Written formal	The collaboration with your company has been a great pleasure for my workers and I.
Written formal, C	These findings have been very important for my colleagues and me.
Subject *me*	
Spoken informal	Me and my husband went to a party with several other young couples.
Written informal	Me and dad are on our way home!
Written formal	My team and me are working to resolve your problem as soon as possible.
Written formal, C	My colleagues and I will look into this and get back to you as soon as possible.

Table 7.5: Stimuli sentences for object *I* and subject *me*

variant. The distribution of ratings for the four sentences with object *I*, across the four different levels, is shown in Figure 7.5. The figure shows that there is a fairly positive to neutral distribution of the ratings. This is also the first feature among those discussed so far for which some of the ratings are on the 'very positive' side of the scale. Some variation is nevertheless noticeable. First of all, the sentence which contained the 'correct' variant, object *me*, is decidedly positively rated. The rest of the sentences are more varied in their ratings. The spoken informal sentence, as well as the written informal one, are positively rated for ACCEPTABILITY. The ratings for CORRECTNESS are more evenly distributed between the two extremes, and this is especially the case for the written informal sentence. For EDUCATEDNESS, the three sentences with object *I* were all rated neutrally, while the rest of the ratings were distributed evenly across the two extremes of the scale.

Pairwise comparison tests were applied to explore differences between the ratings for the four levels for each of the sentences separately. No significant differences were found in the ratings across the four different levels, which may suggest that ACCEPTABILITY, CORRECTNESS, GOODNESS, and EDUCATEDNESS were not seen as different dimensions by the respondents.

In terms of differences in ratings affected by the opposition between spoken and written contexts of use, Wilcoxon tests were used to compare the ratings for the

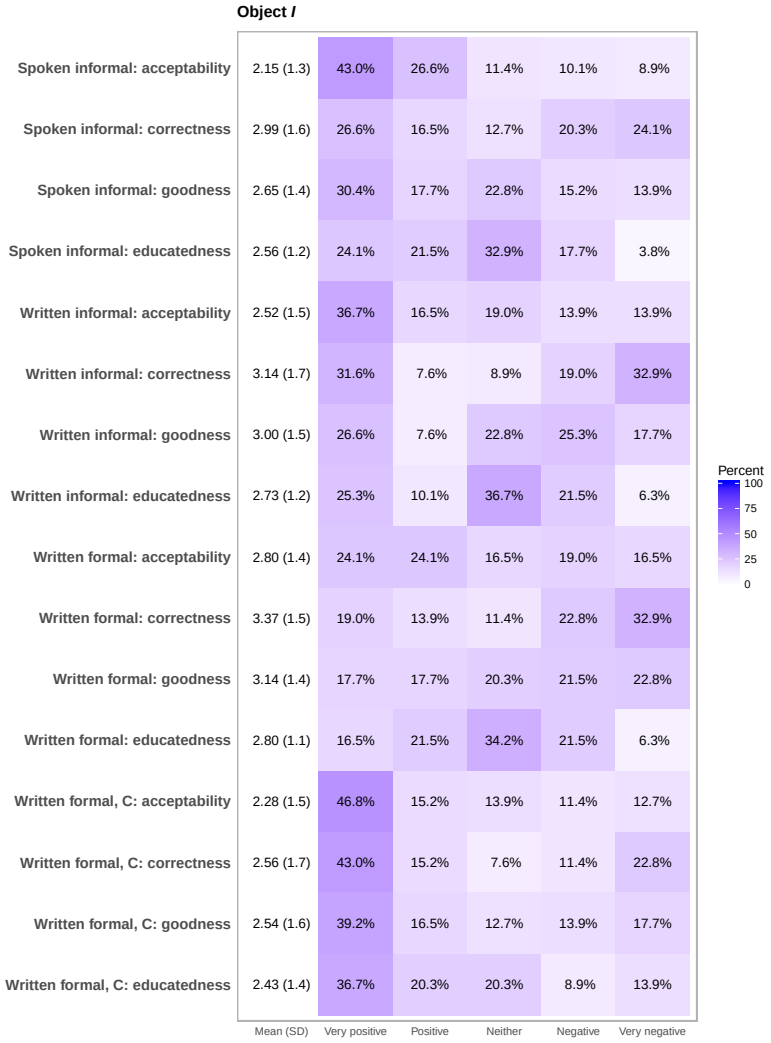

Figure 7.5: Distribution of ratings for object *I*, n = 79

spoken informal sentence with those for the written informal one. The ratings were compared for each of the four levels separately. These tests did not result in statistically significant differences between these two sentences. A similar comparison was done between the ratings for the written informal sentence and those for the written formal one, for each level separately. Here too, no statistically significant differences were identified at the Bonferroni corrected significance level. Finally, a comparison between

the written formal sentence with object *I* and the written formal sentence with object *me* also showed that there are no significant differences in the ratings for these two sentences. No statistically significant differences were identified in relation to the social factors included in the analysis.

The ratings for subject *me* are distributed rather differently than those for object *I*. The first difference was identified in the RECOGNITION LEVEL for the two features. While for object *I* only about 50% of the respondents pointed to the use of this variant as the factor affecting their ratings, in the case of subject *me* the recognition level was between 85% and 90%. This indicates that subject *me* is more salient as a problematic usage than object *I*. The ratings for the sentences with subject *me* are plotted in in Figure 7.6. The plot shows that the majority of the sentences with subject *me* are rated negatively, with a few being rated neutrally. This is especially the case with the ratings for CORRECTNESS: almost 50% of the respondents rated the spoken informal sentence as 'very incorrect', with the figures reaching 54% for the written informal sentence and 67% for the written formal one. The ratings for ACCEPTABILITY were distributed fairly evenly across the five points on the scale, which suggests that CORRECTNESS and ACCEPTABILITY may be perceived differently in the case of subject *me*; I return to this question in the next paragraph, where I discuss the results from the statistical tests. Before considering this, two more observations should be made on the basis of Figure 7.6. First, in terms of EDUCATEDNESS, the ratings tend to be predominantly neutral for the spoken informal and the written informal sentences. In the case of the written formal sentence, the EDUCATEDNESS ratings are somewhat more negative. Finally, the sentence with subject *I* is very positively rated across all four levels.

In order to obtain more robust evidence for these differences, pairwise tests were done to compare the ratings for the four different levels for each of the sentences. These tests showed that subject *me* in spoken informal contexts is rated more positively for ACCEPTABILITY than for CORRECTNESS; the difference is statistically significant ($W = 4400$, $Z = -4.581$, $p < 0.0004$, effect size = 0.515). The same sentence was rated more negatively for CORRECTNESS than for EDUCATEDNESS; for the latter level, the ratings are neutral ($W = 4263.5$, $Z = -4.110$, $p < 0.0004$, effect size = 0.462). In written informal contexts, ACCEPTABILITY and CORRECTNESS are rated differently ($W = 4455$, $Z = -4.810$, $p < 0.0004$, effect size = 0.541): while the use of subject *me* is considered incorrect, it is also considered acceptable. CORRECTNESS and EDUCATEDNESS were also rated differently ($W = 4294$, $Z = -4.254$, $p < 0.0004$, effect size = 0.478). In the written formal context, CORRECTNESS and ACCEPTABILITY are not rated differently; however, there is a statistically significant

difference between the ratings for CORRECTNESS, which are negative, and those for EDUCATEDNESS, which are neutral (W = 4205, Z = −4.107, $p < 0.0004$, effect size = 0.462). Finally, for the control stimulus, which contained the subject *I* variant, all ratings were positive, and there is no statistically significant difference between the four levels.

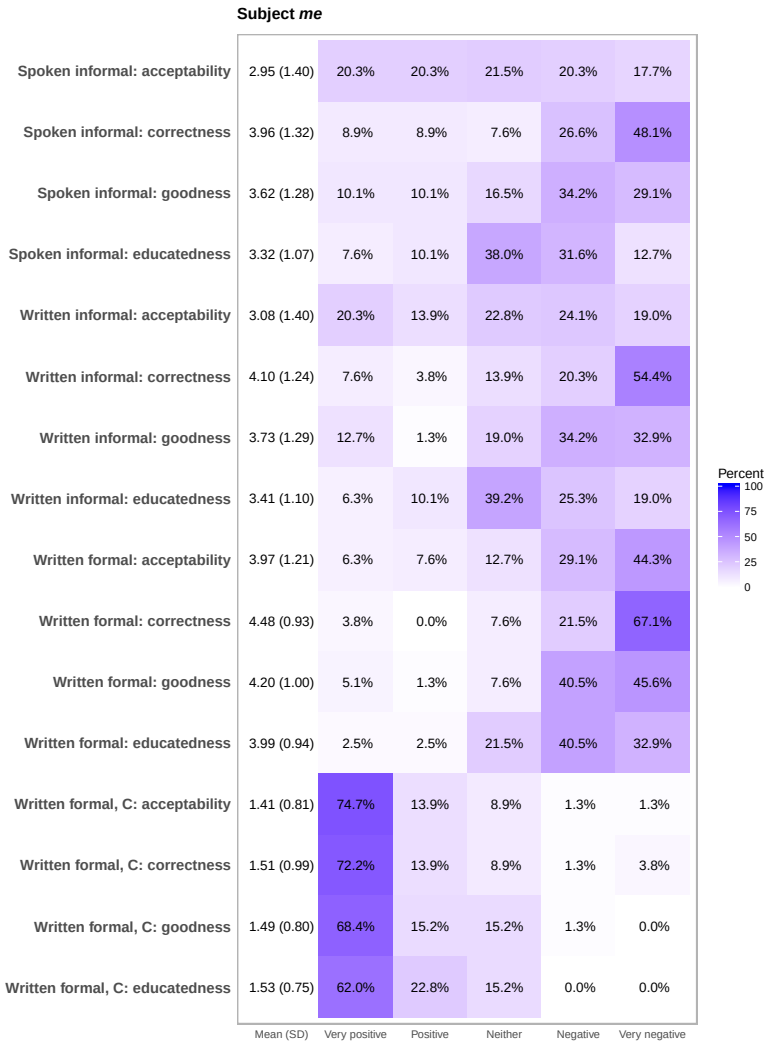

Figure 7.6: Distribution of ratings for subject *me*, n = 79

Comparing the ratings for the spoken informal and the written informal sentences

resulted in no statistically significant differences for any of the four levels. Statistically significant differences were identified between the ratings for ACCEPTABILITY for the written informal and the written formal sentence, with the latter being rated more negatively than the former (W = 4285, Z = −4.174, $p < 0.0004$, effect size = 0.469). The ratings for EDUCATEDNESS were also statistically significantly different (W = 4094.5, Z = −3.532, $p < 0.00045$, effect size = 0.397). Finally, the comparison between the written formal sentence and the control sentence containing subject *I* were significantly different for ACCEPTABILITY (W = 5799.5, Z = −9.733, $p < 2.2e\text{-}16$, effect size = 1.095), CORRECTNESS (W = 5905.5, Z = −10.216, $p < 0.00045$, effect size = 1.149), GOODNESS (W = 5956, Z = −10.236, $p < 0.00045$, effect size = 1.151), and EDUCATEDNESS (W = 5988, Z = −10.253, $p < 0.00045$, effect size = 1.153). The last result is especially interesting, in the light of the parallel comparison done for sentences with object *I* and object *me*, which were not rated significantly differently. Comparing the ratings by different gender, age, and ethnicity groups did not produce any statistically significant results.

7.7 The split infinitive

The final feature investigated is the split infinitive. The sentences containing a split infinitive which were included in the survey are given in Table 7.6. In the context of this feature, sentences with a split infinitive were presented in spoken formal, written

context	**stimulus sentence**
Spoken formal	So, I would encourage young men and women to seriously consider a career in law enforcement.
Written formal	This therapy has been shown to significantly reduce the risks of heart attacks and strokes.
Written informal	Trying to decide if there is anything interesting to further explore in my new town.
Written informal, C	Trying to find out if there is anything interesting to explore further in my new town.

Table 7.6: Stimuli sentences for the split infinitive

formal, and written informal contexts. In addition to these three, a sentence with a modified non-split infinitive was also included in a written informal context; this sentence is marked ‘C’ in Table 7.6. This sentence allows for a comparison to be done between the ratings for the sentence containing a split infinitive and the sentence with

a non-split infinitive in written informal contexts.

The split infinitive was the feature with the lowest RECOGNITION LEVEL: only about 4% of the respondents explicitly mentioned the split infinitive as the words which affected their ratings for the sentences. The rest of the respondents commented on other aspects of the sentence, but not on the split infinitive. This suggests that the split infinitive is not indexical of incorrectness. Figure 7.7 shows that, of all the features analysed, the ratings for the split infinitive are the most positive overall. The figure shows that for both the spoken formal and the written formal sentences, the majority of the ratings were on the 'very positive' side of the scale, and here the ratings are the highest for ACCEPTABILITY, followed by CORRECTNESS, GOODNESS, and EDUCATEDNESS. In the case of EDUCATEDNESS and GOODNESS for the spoken formal sentence, no respondents rated this sentence at the 'very negative' end of the scale. The peculiar thing about the ratings for the sentences with a split infinitive is that the sentence in the written informal context was not rated as positively as the other two sentences with a split infinitive. In addition, there are no noticeable differences in the ratings between the written informal sentence with a split infinitive and the one with a non-split infinitive, even though they are fairly similar, as shown in Table 7.6. The reason for this is in part the result of a limitation in the sentence presented to respondents; the sentence contained no explicitly realised subject, which resulted in most of the respondents criticising this aspect of the sentence structure. I will discuss this in further detail below, after providing an overview of the results from the statistical tests.

The four different levels were not rated significantly differently for any of the stimuli. In terms of register, the comparison between spoken formal and written formal contexts is not statistically significant across the four levels. The ratings for written informal contexts are, however, statistically significantly more negative than those for the written formal stimulus for ACCEPTABILITY ($W = 4198.5$, $Z = -4.003$, $p < 0.00055$, effect size = 0.450), CORRECTNESS ($W = 4647.5$, $Z = -5.522$, $p < 0.00055$, effect size = 0.621), GOODNESS ($W = 4791$, $Z = -6.061$, $p < 0.00055$, effect size = 0.682), and EDUCATEDNESS ($W = 4654$, $Z = -5.561$, $p < 0.00055$, effect size = 0.625).

Similar statistically significant differences were identified between the ratings for the spoken formal sentence and the written informal one. Finally, a comparison between two written informal sentences, one with a split infinitive and the other with a post-modified infinitive, shows that these two sentences were rated the same: no statistically significant differences were identified. In terms of social variables, no

statistically significant differences in the ratings were identified.

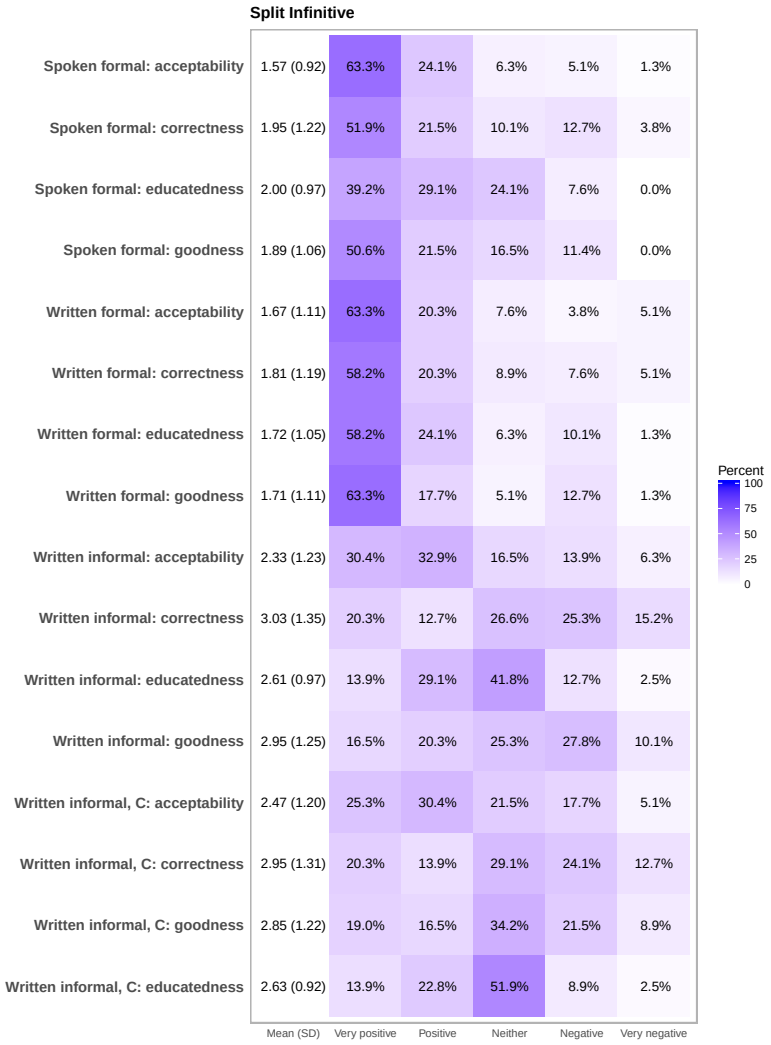

Figure 7.7: Distribution of ratings for the split infinitive, n = 79

These results show that the majority of the respondents are unaware of the split infinitive. This is shown first of all by the positive ratings for the spoken and the written formal sentences. While at first glance the negative ratings for the written informal sentence might appear to be surprising, an examination of the comments given by the respondents on the motivation for their ratings showed that the negative

ratings for this sentence are not due to the split infinitive, but rather to the fact that the sentence does not contain a subject. Even though some respondents noted that leaving out the subject is acceptable on social media (which is the context in which this sentence was provided), the overall ratings were still negative. Thus, what might seem at first a conflicting result may actually suggest that it is the lack of awareness of the split infinitive as a mistake that resulted in the ratings. Additional support for this interpretation is the comparison between the two sentences presented in written informal contexts, one of which contains a non-split infinitive ("Written informal, C" in Table 7.6 above). This sentence was rated in a similar way as the one containing a split infinitive. While this may be considered a limitation in this context, and a limitation in general when working with sentence stimuli, it also shows that the split infinitive is a fairly neutral feature.

7.8 The ratings compared

In the preceding sections I discussed each feature separately, and how the ratings differed across the four levels ACCEPTABILITY, CORRECTNESS, EDUCATEDNESS, and GOODNESS, as well as across contexts of use and the social background of the respondents to the survey I carried out. I now turn to the final set of analyses conducted, which move beyond looking at individual features, and compare the ratings between features. These comparisons across features produced more significant results than comparisons of ratings across age, gender, and ethnicity groups. The main concern here is exploring the similarities and differences among the ratings for the various sentence stimuli. In addition to examining further the connection between the stimuli, this section will discuss the hypothesis that certain usage problems, such as *ain't* and negative concord, tend to be rated similarly. These comparisons were carried out on the basis of similar contexts. For example, the ratings for *ain't* in the spoken informal context were compared with the ratings for *like* in the same context.

A comparison between the ratings for *ain't* and those for *like* showed that the sentences containing these two features are not rated differently, as was the case across the four levels. These results suggest that these two features are seen as similarly problematic by all respondents. Another feature which was rated comparably to *ain't* is negative concord. For all the other features, there were statistically significant differences, but the extent to which the ratings for *ain't* differed from those of other features varied. In the context of three other features, *ain't* is rated significantly more

negatively across the four levels compared to those three features, i.e. *literally*, object *I*, and the split infinitive. *Ain't* is rated more negatively than *literally* across all four levels, and this difference is statistically significant for ACCEPTABILITY (W = 4793, Z = −5.964, $p < 0.00058$, effect size = 0.671), CORRECTNESS (W = 4601, Z = −5.527, $p < 0.00058$, effect size = 0.621), GOODNESS (W = 4887.5, Z = −6.437, $p < 0.00058$, effect size = 0.724), and EDUCATEDNESS (W = 5001, Z = −6.811, $p < 0.00058$, effect size = 0.766). The ratings for *ain't* are significantly more negative than those for object *I*, on the basis of a comparison between the spoken informal stimuli for the two features. This difference was significant for the four levels: ACCEPTABILITY (W = 5016.5, Z = −6.783, $p < 0.00045$, effect size = 0.763), CORRECTNESS (W = 4863.5, Z = −6.432, $p < 0.00045$, effect size = 0.723), GOODNESS (W = 5241.5, Z = −7.657, $p < 0.00045$, effect size = 0.861), and EDUCATEDNESS (W = 5301, Z = −7.788, $p < 0.00045$, effect size = 0.876). Finally, *ain't* is also rated more negatively than the split infinitive in spoken formal contexts across the four levels: ACCEPTABILITY (W = 5786.5, Z = −9.571, $p < 0.00055$, effect size = 1.076), *correctness* (W = 5861.5, Z = −9.926, $p < 0.00055$, effect size = 1.116), GOODNESS (W = 5967.5, Z = −10.184, $p < 0.00055$, effect size = 1.145), and EDUCATEDNESS (W = 5880.5, Z = −9.805, $p < 0.00055$, effect size = 1.103).

The comparison between the ratings for *ain't* and for subject *me* reveals interesting insights into the sensitivity with which respondents rated these sentences. The comparison was done on the basis of the spoken informal stimuli. The sentences were not rated differently for CORRECTNESS, i.e. being both rated negatively. However, the sentences are rated differently for ACCEPTABILITY, with *ain't* being considered more unacceptable than subject *me* (W = 4331.5, Z = −4.338, $p < 0.00045$, effect size = 0.488). The ratings for the ACCEPTABILITY of subject *me* in spoken informal usage are fairly balanced, and tend towards the positive end of the scale. There was also a difference between *ain't* and subject *me* in the ratings for GOODNESS (W = 4345.5, Z = −4.548, $p < 0.00045$, effect size = 0.511) and EDUCATEDNESS (W = 4511.5, Z = −5.047, $p < 0.00045$, effect size = 0.567). This shows that while subject *me* is considered acceptable, it is nevertheless viewed as incorrect and 'bad English'. On the level of EDUCATEDNESS, the ratings were predominantly neutral.

In the case of *like*, I mentioned above that there were no differences in the ratings for the discourse particle *like* and those for *ain't* for any of the four levels. The ratings for the discourse particle *like* and for negative concord were not compared, because the survey did not include sentences in which these features were used in the same context, but it might be expected that the two would not be rated differently, on the

basis of the similarity in ratings between *ain't* and *like*, on the one hand, and *ain't* and negative concord, on the other. Comparison with the rest of the features yielded the following statistically significant differences. *Like* and *literally* were rated differently in spoken cases (both male and female speakers), and the ratings were different across all four levels. The stimuli sentences with *like* and *literally* were rated more negatively for ACCEPTABILITY (W = 1335, Z = −6.342, $p < 0.00055$, effect size = 0.713), CORRECTNESS (W = 1742, Z = −4.989, $p < 0.00055$, effect size = 0.561), GOODNESS (W = 1397, Z = −6.216, $p < 0.00055$, effect size = 0.699), and EDUCATEDNESS (W = 1376.5, Z = −6.316, $p < 0.00055$, effect size = 0.710). Similar results were obtained when the tests were repeated on the stimuli spoken by a male speaker. *Like* and object *I* were rated differently across the four scales: ACCEPTABILITY (W = 1194, Z = −6.852422, $p < 0.00045$, effect size = 0.770), CORRECTNESS (W = 1689, Z = −5.179, $p < 0.00045$, effect size = 0.582), GOODNESS (W = 1186.5, Z = −6.921, $p < 0.00045$, effect size = 0.778), and EDUCATEDNESS (W = 1370.5, Z = −6.255, $p < 0.00045$, effect size = 0.703). In all these cases the sentences with *like* were rated more negatively than those with object *I*. *Like* and subject *me*, on the other hand, were rated differently only for ACCEPTABILITY (W = 1993, Z = −4.021, $p < 0.00045$, effect size = 0.452); the differences in the ratings for CORRECTNESS, GOODNESS, and EDUCATEDNESS for *like* and subject *me* are not significant. This pattern follows the one observed between *ain't* and subject *me*, as well as between *ain't* and object *I*. Finally, sentences with *like* and the split infinitive were not compared, because they did not occur in the same context in the survey.

The difference between the ratings for *literally* and negative concord is also statistically significant. This was tested by comparing the ratings for the two features in the written informal context. The ratings were different for ACCEPTABILITY (W = 2070.5, Z = −3.752, $p < 0.00055$, effect size = 0.422), CORRECTNESS (W = 2047.5, Z = −4.006, $p < 0.00055$, effect size = 0.450), GOODNESS (W = 1915.5, Z = −4.381, $p < 0.00055$, effect size = 0.492), and EDUCATEDNESS (W = 1832.5, Z = −4.665, $p < 0.00055$, effect size = 0.524).

As already discussed in the context of comparisons between *ain't* and *literally*, and between *like* and *literally*, *literally* is rated less negatively than the other two features. A comparison between the ratings of sentences with *literally* and sentences with object *I* did not result in any statistically significant differences. When compared to subject *me*, *literally* was rated differently for CORRECTNESS (W = 4148.5, Z = −3.699, $p < 0.00045$, effect size = 0.416) and EDUCATEDNESS (W = 4151, Z = −3.779, $p < 0.00045$, effect size = 0.425), but not for ACCEPTABILITY and

GOODNESS. Finally, a comparison between the ratings for sentences with *literally* and with the split infinitive showed no statistically significant differences between the two in written informal contexts. However, given the complications which arise from the nature of the written informal stimulus for the split infinitive discussed above, I compared the ratings for *literally* in the spoken informal context with those for the split infinitive in the spoken formal one. These proved statistically significant across the four levels: ACCEPTABILITY (W = 4370, Z = −4.641, $p < 0.00045$, effect size = 0.522), CORRECTNESS (W = 4613, Z = −5.362, $p < 0.00045$, effect size = 0.603), GOODNESS (W = 4792.5, Z = −5.998, $p < 0.00045$, effect size = 0.674), and EDUCATEDNESS (W = 4588.5, Z = −5.321, $p < 0.00045$, effect size = 0.598).

Negative concord was compared with *ain't*, *literally*, object *I*, subject *me*, and the split infinitive. The comparison between the ratings for negative concord and for *ain't* showed that there are no statistically significant differences at any of the four levels. The ratings for negative concord and *literally* are different across all four levels; sentences with negative concord are rated more negatively than those with non-literal *literally*.

Negative concord and object *I* were compared on the basis of the written informal sentences. This comparison showed that the ratings differed significantly across the four levels: ACCEPTABILITY (W = 4368.5, Z = −4.450, $p < 0.00045$, effect size = 0.500), CORRECTNESS (W = 4390, Z = −4.714, $p < 0.00045$, effect size = 0.530), GOODNESS (W = 4595, Z = −4.714, $p < 0.00045$, effect size = 0.530), and EDUCATEDNESS (W = 4676.5, Z = −5.573, $p < 0.00045$, effect size = 0.627). A similar comparison between the ratings for the written informal sentence with negative concord and those for the written informal sentence with subject *me* produced no statistically significant results. Finally, comparing the ratings for negative concord and those for the split infinitive showed that in the written informal context the two sentences were rated differently for ACCEPTABILITY (W = 4652.5, Z = −5.441, $p < 0.00045$, effect size = 0.612), CORRECTNESS (W = 4947.5, Z = −6.623, $p < 0.00045$, effect size = 0.745), GOODNESS (W = 4818, Z = −6.090, $p < 0.00045$, effect size = 0.685), and EDUCATEDNESS (W = 4984.5, Z = −6.675, $p < 0.00045$, effect size = 0.751).

As for the sentences with pronouns in coordinated phrases, object *I* and subject *me* display different patterns in the ratings. In the case of object *I*, I mentioned above that this form is rated significantly more positively than *ain't* for the four levels investigated. The same holds for *like*. Object *I* and negative concord were also rated differently, with object *I* being the most positively rated feature on the basis

of the sentences used in the survey. Comparing the differences between object *I* and subject *me* in the spoken informal sentences provides further interesting differences in how the two variants are rated. In spoken informal contexts the two variants are rated differently for CORRECTNESS (W = 4233, Z = −4.006, $p < 0.00045$, effect size = 0.450), ACCEPTABILITY (W = 4140.5, Z = −3.646, $p < 0.00045$, effect size = 0.410), GOODNESS (W = 4319.5, Z = −4.261, $p < 0.00045$, effect size = 0.479), and EDUCATEDNESS (W = 4247, Z = −4.055, $p < 0.00045$, effect size = 0.456). Finally, object *I* and split infinitive are significantly different across the four levels: ACCEPTABILITY (W = 4577.5, Z = −5.334, $p < 0.00045$, effect size = 0.600), CORRECTNESS (W = 4826.5, Z = −6.162, $p < 0.00045$, effect size = 0.693), GOODNESS (W = 4858, Z = −6.300, $p < 0.00045$, effect size = 0.708), and EDUCATEDNESS (W = 4715.5, Z = −5.768, $p < 0.00045$, effect size = 0.649).

The ratings for subject *me* are different from those for *ain't* only for ACCEPTABILITY, as I mentioned above in the discussion of *ain't*. For CORRECTNESS, both features are rated negatively, and no statistically significant differences were identified. However, for ACCEPTABILITY, subject *me* was seen as more acceptable than *ain't* on the basis of the ratings. Similar results were obtained from a comparison between the ratings for the sentence with the discourse particle *like* and those for the spoken informal sentence with subject *me*. There were no differences in ACCEPTABILITY, but there were differences for the other three levels analysed. I also discussed the difference between subject *me* and *literally*; the sentences with these two features were rated differently for CORRECTNESS and EDUCATEDNESS, but not for ACCEPTABILITY and GOODNESS. Subject *me* was also rated more positively than negative concord, with statistically significant differences across all four levels. Finally, comparing the ratings for sentences with subject *me* and with a split infinitive shows that subject *me* is rated more negatively; this difference is statistically significant across the four levels: ACCEPTABILITY (W = 5548.5, Z = −8.731, $p < 0.00045$, effect size = 0.982), CORRECTNESS (W = 5767, Z = −9.600, $p < 0.00045$, effect size = 1.080), GOODNESS (W = 5715, Z = −9.367, $p < 0.00045$, effect size = 1.053), and EDUCATEDNESS (W = 5713.5, Z = −9.270, $p < 0.00045$, effect size = 1.043).

Finally, the sentences with split infinitives are the most positively rated out of all the sentences analysed. The pairwise comparisons between the ratings for sentences with a split infinitive and those with other features showed that the sentences with a split infinitive are rated statistically significantly more positively. This is shown in a comparison with *ain't*, *literally*, negative concord, object *I*, and subject *me*, all

described above.

On the basis of the ratings obtained, we can conclude that there is variation in the perception of these features. Interesting patterns emerge which suggest the existence of levels of acceptability in relation to the different features. For instance, *ain't*, *like*, and negative concord seem to be associated with similar patterns of negative evaluation, while *literally*, object *I*, and the split infinitive seem to be more neutrally to positively evaluated. Positive evaluation is particularly associated with the split infinitive. Sentences with subject *me* are somewhere in between, and exhibit the greatest differences between the ratings across the four different levels. While this is of course hardly surprising, the important questions to address here are what this variation reveals, and how it relates to the empirical study of usage guides and patterns of actual language use. In the final section of this chapter, I turn to a discussion of the importance of the results presented in this chapter so far, in order to try to answer these questions.

7.9 Discussion and conclusion

In this chapter I have presented the results of an analysis of attitudes to the use of the six linguistic features investigated in this study. I have focused predominantly on the ratings of sentences in a survey collected from 79 speakers of American English, as explained in Section 4.5. The ratings were used to explore the attitudes of speakers to the use of the six linguistic features across different levels (ACCEPTABILITY, CORRECTNESS, GOODNESS, and EDUCATEDNESS), as well as in different contexts of use. The effects of social variables, such as age and gender, were also tested. In addition, the ratings for the different features were compared to each other, in order to ascertain the degrees of general acceptability of each of the six features.

First of all, with respect to the four different levels (ACCEPTABILITY, CORRECTNESS, GOODNESS, and EDUCATEDNESS), the predominant tendency I observed was that they were rated similarly. While this was indeed found most frequently, in the case of some features there were differences in the ratings across the four levels. First, there was a statistically significant difference between the ratings for ACCEPTABILITY and those for CORRECTNESS for the spoken formal sentence with *ain't* (see Section 7.2), with ACCEPTABILITY being rated more positively than CORRECTNESS. This result may suggest that these two levels are considered to represent two different notions, but the evidence for this is insufficient. If this were the

case, the more intuitive result would be for this kind of split to be found in informal, rather than formal contexts, because in informal contexts grammatically 'incorrect' forms may be expected to be more acceptable than in formal ones, where acceptability and correctness can be considered two sides of the same coin (i.e. what is acceptable has to be grammatically 'correct'). A difference in the ratings between two of the four levels was also found in the case of one of the sentences with *literally*. The informal sentence spoken by a female speaker was rated more positively for ACCEPTABILITY and more negatively for GOODNESS (see Section 7.4). This suggests that in the case of *literally*, the use of non-literal *literally* may be considered acceptable, even when it is not necessarily seen as 'good English'. Moreover, there was a statistically significant difference in the ratings for ACCEPTABILITY and CORRECTNESS for the written informal sentence with negative concord (see Section 7.5), with the former being more evenly distributed along the five-point scale, and the latter being predominantly rated 'very negative' (see Figure 7.4). This is similar to the difference identified in the ratings of sentences with *ain't*, and may indeed suggest that in some cases these two levels reflect different notions. The sentences with pronouns in coordinated phrases were different in this respect. While no statistically significant differences between the four levels were identified on the basis of the ratings for object *I*, subject *me* was the feature for which most differences were identified. In the case of sentences with subject *me*, ACCEPTABILITY and CORRECTNESS were rated differently in the spoken informal and written informal contexts, while they were not rated differently in the written formal context. This provides further evidence for the fact that these two levels signal different notions: in informal contexts, respondents rate subject *me* as incorrect, but also as acceptable, which shows that they are aware of both the standard norm for pronominal usage and the acceptability of non-standard pronominal forms in informal contexts. In the written formal context however, both ACCEPTABILITY and CORRECTNESS are rated equally negatively, which suggests sensitivity to contexts of use: subject *me* is considered acceptable in informal, but not necessarily in formal contexts. In addition to the difference between ACCEPTABILITY and CORRECTNESS, CORRECTNESS and EDUCATEDNESS are also rated differently, but in this context statistically significant differences were identified in the ratings for all three sentences with subject *me*. This suggests that while subject *me* tends to be considered incorrect yet acceptable in informal contexts but not in formal contexts, when it comes to EDUCATEDNESS, the respondents tend to see subject *me* as neutral. Finally, the fact that there were no statistically significant differences in the ratings for the sentence with subject *I*, which was rated predominantly positively across all four levels,

provides further confirmation that subject *me* is seen as a problematic usage feature which is seen as incorrect, but acceptable and unrelated to the education status of its users.

Secondly, in terms of contexts of use, statistically significant differences were identified only in the case of subject *me*, where evidence for the effect of formality was identified, with the written informal sentence being rated more positively on the ACCEPTABILITY scale than the written formal sentence. The written formal sentence was rated more negatively for EDUCATEDNESS, while the written informal one with subject *me* was rated more neutrally. This suggests that this feature may be considered incorrect and unacceptable in written formal contexts, and that failure to observe these norms may be perceived as uneducated.

Thirdly, in terms of social variables, there were almost no differences in the ratings between genders and age groups. The only statistically significant difference was found in the ratings for the CORRECTNESS of the discourse particle *like*, with the age group 30 OR ABOVE rating the sentence more negatively for CORRECTNESS than the age group 29 OR BELOW. This suggests that the community norms for the general acceptability of this feature are changing, but it is worth noting that while both age groups consider this sentence 'incorrect', there is a difference in the evaluation. For the other features no statistically significant differences were found on the basis of the pairwise comparison tests. These results suggest multiple possible scenarios. The first is that there are no differences in the ratings because there are no differences in the attitudes between different genders, ages, and ethnicities. This may be the result of the fact that most respondents live in a multicultural urban area, where these different groups of people are not isolated from each other. The second scenario is that the nature of the study itself may have been an influencing factor. A study investigating attitudes to usage related to prescriptivism, which is often an approach to language teaching that is part and parcel of the educational system in the United States, may result in the majority of language users expressing the same kinds of attitudes with respect to these features. Precisely because these features are very salient and overtly discussed in public, most respondents may have expressed similar attitudes because these attitudes are shared by the community and are the result of having been imposed top-down. The third scenario is that there are potential shortcomings resulting from the sample not being representative of the general population. While there was indeed a more or less equal proportion of men and women in the sample, the sample was not properly balanced for the other variables, so this may have affected the results. In addition, despite the variation in gender and age, the majority of the respondents were

highly educated, which might also explain why they express similar attitudes to the use of these features.

Finally, comparing the ratings across the six features confirms both that different prescriptively targeted features are indeed evaluated differently, and that certain prescriptively targeted features tend to cluster together in terms of how acceptable they are. The pairwise comparisons between ratings for the linguistic features shows that *ain't*, *like*, and negative concord are the most negatively evaluated features, regardless of the level of evaluation (i.e. ACCEPTABILITY, CORRECTNESS, EDUCATEDNESS, GOODNESS). In addition, context of use has very little to no effect on the evaluation of these three features. The second group of features comprises those with a less negative evaluation than that for the first group of features; *literally* and object *I*/subject *me* belong here. These features are clearly different from the first group, in that they are rated more positively in general, and a comparison between the ratings for these features and those for the features in the first group showed statistically significant differences. However, it is worth noting that even though non-literal *literally*, object *I*, and subject *me* are in this group, there are still important differences between the three. Of special significance here is the difference in ratings for sentences with object *I* and subject *me*. Sentences with object *I* were rated more positively than those with subject *me*, and apart from the results from the multiple pairwise comparison tests discussed in Section 7.6 above, this is also confirmed by the comparison between object *I* and object *me*, as opposed to subject *me* and subject *I*. The first comparison did not result in statistically significant differences in the ratings, while the second did. This suggests that object *I* is not considered to be different from object *me*, while this is definitely not the case when sentences with subject *me* are compared with sentences with subject *I*. This shows that these two features are problematic in a different way, which is an especially relevant issue to address in the context of a comparison with the results from the analysis of the features' usage guide treatment, a point I will discuss in the concluding chapter of this study. Finally, the split infinitive does not appear to be salient as a usage problem, as is evident from the fact that most respondents rated sentences with split infinitives positively across the four different levels. This puts the split infinitive into a group of its own, which represents usage problems which have become increasingly more accepted as part of the standard language usage norms.

This kind of grouping is based on the five-point scale ratings, and does not take into account RECOGNITION LEVEL, which might also be considered as an indicator of how salient and problematic a feature is. In order to arrive at a better understanding of the attitudes to the use of the six features studied, Figure 7.8 presents a visual summary

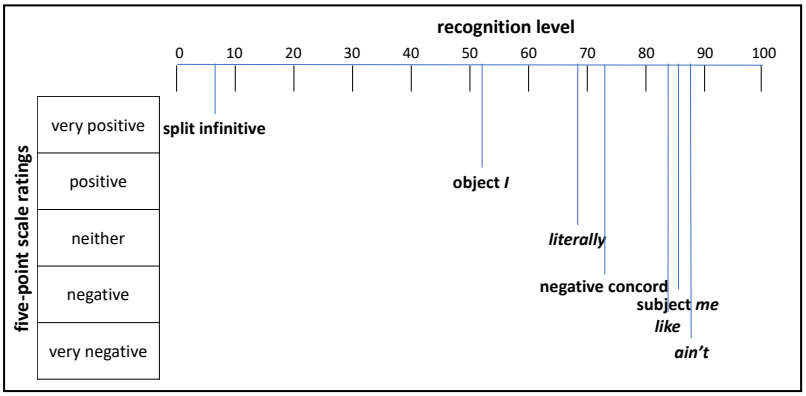

Figure 7.8: Schematic representation of the acceptability of usage problems on the basis of median ratings and average RECOGNITION LEVEL

of how the features are positioned with respect to both RECOGNITION LEVEL and the five-point scale ratings, from 'very positive' to 'very negative'. The horizontal axis represents the mean RECOGNITION LEVEL across all sentences included in the survey for each of the features: a low RECOGNITION LEVEL means that the feature in question was not stated by respondents as a reason for their ratings; a high RECOGNITION LEVEL means that the feature was generally cited as the reason for the specific ratings of the sentence containing that feature. On this scale, then, the split infinitive has a fairly low RECOGNITION LEVEL, which may indicate that most respondents do not recognise it as a problematic usage. On the other hand, subject *me*, the discourse particle *like*, and *ain't* have high RECOGNITION LEVELS, which may be indicative of their salience as problematic usages. The vertical axis represents where each feature stands on the five-point scale; the position for each feature is determined on the basis of the median ratings for all sentences and all four levels (i.e. ACCEPTABILITY, CORRECTNESS, EDUCATEDNESS, GOODNESS), for each of the features. While this is undoubtedly a rough representation, it is nevertheless effective for my purpose here, which is to provide a conceptual mapping of the attitudes to the use of the six features investigated, on the basis of the survey data. The figure thus shows that the language features display a continuum of problematicity or acceptability. This kind of visualisation allows us to compare the results from the analysis of speakers' attitudes to those from the analysis of usage guide treatment, a question discussed in the concluding chapter of this study.

CHAPTER 8

Conclusion

This study set out to explore empirically the influence of prescriptivism on language use and on language users in American English. Given that the English prescriptive tradition targets a relatively small set of language features,[1] the study focused on six of them: *ain't*, the discourse particle *like*, *literally*, negative concord, object *I*/subject *me*, and the split infinitive. A number of assumptions about the nature of the prescriptive influence were established at the outset, in Chapter 1, which provided the starting points for developing the methodology as outlined in Chapter 4, and the analyses presented in Chapters 5, 6, and 7. The first assumption was that a better understanding of the potential influence of prescriptivism requires a careful analysis of prescriptive ideology, understood as the attitudes to language usage propagated by top-down means, such as the educational system and publications on usage. In order to analyse prescriptive attitudes and how they have changed in the course of the nineteenth and twentieth centuries, I analysed 70 usage guides published in the United States between 1847 and 2015. The second assumption was that the influence

[1] The HUGE database contains 123 usage problems, which is of course not an exhaustive list. Chapman (2017: 246) identifies over 10,000 prescriptive rules on the basis of a survey of 30 handbooks; however, it is not clear whether all these rules refer to different usage problems. While there are many usage features addressed in the usage guide tradition, it is also true that the canon of traditionally discussed features is fairly small.

of prescriptivism on language use can reasonably be considered (a) to be potentially relevant only in relation to a small set of prescriptively targeted features, and (b) to be identified in patterns of variation, rather than in patterns of language change. The latter essentially means that prescriptive influence is in general temporally limited. The influence of prescriptivism on patterns of language variation and change was then explored by examining six linguistic features in this context. This was done by using both text-linguistic and variationist approaches to the analysis of most of these features, in order to obtain a reliable picture of the possible influence of prescriptivism. In addition to these analyses, a multifactorial analysis was conducted for one of the features, the split infinitive, in order to explore the extent to which the likelihood that a modified infinitive occurring in a specific text is split is associated with the occurrence in the same text of other prescriptively targeted features. Finally, the third assumption was that, in addition to the potential influence of prescriptivism on language variation and change, prescriptive attitudes also affect language speakers. Here, the attitudes of American English speakers towards the set of six linguistic features were explored by means of a survey and post-survey interviews with 79 speakers. In what follows, I will summarise the findings of this analysis for each of the language features separately, and discuss their importance in revealing the potential influence of prescriptivism.

In the case of *ain't*, the analysis of usage guides showed that the feature has become somewhat acceptable over time, but that this acceptability is restricted to specific contexts of use. These contexts include the use of *ain't* in specific fixed phrases or catchy expressions, in fiction, and in song lyrics. Despite this acceptability, the form is still considered indexical of uneducated speech. The results from the analysis of patterns of variation and change showed that the word is infrequently found in the corpus data. The analysis provided no evidence of change; rather, it suggested that there is register variation in the use of *ain't*. In part, this confirms the observations found in usage guides concerning the restricted acceptability of *ain't*. The analysis of speakers' attitudes showed that *ain't* is the most negatively rated feature of the six linguistic features investigated. All these results together suggest that *ain't* is an 'old chestnut' which continues to be covered in usage guides. Its use is restricted, but it has not entirely been ousted from the language, and, on the basis of a number of studies of non-standard varieties of English, it can be expected that *ain't* will remain part of vernacular usage. In standard language data, however, its use is mostly restricted to very specific contexts. In addition to this, the speakers' attitude data showed that it is a highly stigmatised feature. All in all, it is fairly difficult to ascertain prescriptive influence on the use of *ain't* on the basis of these results, in the context of decreasing

frequencies of use of the form, specifically for the period investigated. However, what can be said with perhaps a greater degree of certainty is that the treatment of *ain't* found in usage guides reflects the main findings on the register sensitivity of *ain't*, as well as its negative social indexicality, as evidenced by the predominantly negative ratings this feature received from the respondents. In other words, in the case of *ain't*, usage guide treatment reflects the sociolinguistic reality of the use of *ain't* in standard American English.

Discourse particle *like* provides evidence for how usage guides react to highly salient language changes. On the basis of the treatment of *like* in usage guides (as shown in Sections 5.2.2, 5.3.2, 5.4.2, and 5.5.2), it can be concluded that language changes which become salient for speakers and are negatively socially evaluated become part of the usage problem canon. A parallel can be made here with the process by which *you was* became a usage problem during the eighteenth century (Laitinen 2009). The speakers' attitudes data also confirmed the negative social indexicality of the use of the discourse particle *like* (see Section 7.3). Despite this, however, the analysis of the actual use of *like* showed an increase in the frequency of occurrence of the word, especially in the COCA data (see Section 6.3). This shows that prescriptivism, as well as the negative social evaluations of the discourse particle *like*, seem not to have exerted any noticeable influence on its use, and little can be said with respect to predicting future trends. However, it has been shown with some level of certainty that usage guides, and perhaps similar metalinguistic works, react to robust and highly salient language changes; this confirms that usage guides and similar works can be used as additional evidence of language changes, especially in past periods, for which there is a paucity of spoken language data.

The case of *literally* is in some respects similar to *like*, in that here too usage guides seem to be responding to an ongoing language change. While the patterns of change in the uses of *literally* observed in the corpus data (Section 6.4) do not suggest a rapidly progressing change, additional factors which may have influenced the salience of the variation in the uses of *literally* are the nature of the feature and negative connotations associated with its use. By the nature of the feature I mean the development of the non-literal use of *literally*, which is perceived as meaning the opposite of the compositional meaning of the lexeme. As a result of the striking opposition between the non-literal use of *literally* to modify figurative expressions and its compositional meaning of 'word for word', the use of non-literal *literally* has acquired an indexical meaning related to lack of logic, absurdity, and ignorance on the part of the language user. Where *like* and *literally* differ, however, is in the extent to

which they are negatively evaluated by speakers. On the basis of a comparison between these two features (see Section 7.9, specifically Figure 7.8), *literally* is less negatively evaluated than *like*. This, together with evidence from the post-survey interviews, suggests that, for *literally*, covert sociolinguistic prestige may be developing, where the non-literal use of *literally* is associated with a specific group of speakers. What is more, among this group of speakers, the word is not seen as negative, but rather as a group-identity marker, as shown in Section 7.4. This may explain the use of the word in the spoken language, despite the negative treatment found in usage guides. Ultimately, what this suggests is that the influence of prescriptive attitudes may be modified by other sociolinguistic processes related to group identity and non-standard usages which function as markers of group identity.

Negative concord is a feature very similar to *ain't* in that it is very limited in actual use data, but it is different in that it is not a frequently discussed feature in the usage guides analysed. The frequency of negative concord started to diminish in the course of the seventeenth century (Nevalainen 2000), and by the eighteenth century the feature had largely been ousted from the standard variety in England, but not from non-standard usage (Tieken-Boon van Ostade 2008a). In twentieth-century American English, the feature is fairly infrequent in the standard language corpus data used in this study, which may explain in part its relatively infrequent coverage in usage guides. In other words, this is a fairly stable vernacular feature, whose discussion in usage guides may be a relic from the prescriptive tradition.

Pronouns in coordinated phrases provided interesting insights into the relationship between usage guide attitudes and speakers' attitudes. Usage guides tend to be more positive about the use of object *I* than about subject *me*, and this kind of distinction is also reflected in the ratings for these features obtained by respondents (see Section 7.6). However, when it comes to any effects of prescriptivism on actual language use, non-standard pronominal forms are relatively rarely used, so in this case prescriptive influence was not identified. In this case, too, I believe, it can be said that both object *I* and subject *me* are used relatively infrequently in standard American English. What is more, the difference between attitudes towards object *I* and subject *me* does not seem to be reflected in patterns of variation in the use of the two features.

Finally, the split infinitive provides the most interesting case of the potential influence of prescriptivism on language use. The case of the split infinitive suggests that, in relation to the increased usage of a stylistic prescriptively targeted feature, usage guides treatment can develop towards higher levels of acceptability. The fact that the split infinitive is not a salient usage problem for speakers may suggest

that the change in usage guides treatment is more likely for features which are not characterised by negative social evaluations. Thus, the example of the split infinitive shows that, in the long run, prescriptive attitudes to usage tend to change under the influence of both the increasing use of a feature and the lack of negative social evaluations of that feature. However, the split infinitive also shows the complexity and dialectic nature of prescriptive influence. The multifactorial analysis of the use of split infinitives as opposed to non-split modified infinitives on the level of individual texts, discussed in Section 6.8, showed that even though the general long-term trends show that prescriptivism does not exert any influence on the use of split infinitives, synchronic patterns of register variation, and the co-occurrence of a number of prescriptively targeted features show that prescriptive influence can be identified in individual texts, as well as, to some extent, in certain registers or text types. This is, I believe, important evidence suggesting that the way in which prescriptive influence is conceptualised should be more nuanced, and should take into account multiple processes which interact with each other. This leads me to a number of final conclusions, which can be made on the basis of the results discussed above.

The first conclusion is that prescriptive attitudes as instantiated in the American usage guide tradition are not, as has often been suggested, entirely divorced from the facts of actual language use. Usage guides are not produced in a vacuum, even though it is certainly possible to come across pronouncements on language usage which seem to be at odds with how language is actually used. While such examples can indeed be found in individual usage guides, taken on the whole, my study of treatment has shown that the usage guide genre is varied, and continues to develop in ways which are intricately linked with developments in language use. These developments are to a certain extent idiosyncratic for each specific feature, but some similar trends may be observed. One of these trends is that usage features which are 'old chestnuts', and whose actual use is limited to specific contexts, seem to be accepted in these particular contexts. Such features are *ain't*, pronouns in coordinated phrases, and, to a lesser extent, negative concord. For other features whose frequency in use grew in the course of the twentieth century, the treatment has not changed noticeably, and their treatment is predominantly negative; non-literal *literally* and the discourse particle *like* are examples of this trend. Finally, with respect to the split infinitive, the treatment appears to have changed most significantly compared to all the other features analysed.

The second conclusion is that this relationship between prescriptive attitudes and actual use is not a one-way street. It appears to be counterproductive in a study of the relationship between prescriptive attitudes and actual language practice to attempt

to prove either that prescriptive attitudes have absolutely no influence on language practice, rather than the other way around, or that language use affects prescriptive attitudes, and not the other way around. What appears to be the case is that prescriptive attitudes and actual language practice are mutually influenced by each other, and the extent and direction of that influence depend on very many factors, ranging from the language features concerned to the sociolinguistic context in which specific language practices take place. In this context, it is important to understand that it is almost always the case that prescriptive attitudes will affect language practice some of the time, and that the reverse will also be the case. While this observation does not originate with me, it is important to note that the empirical study I have presented here provides confirmation for it.

The third conclusion is that prescriptive attitudes and speakers' attitudes are often similar in their evaluations of prescriptively targeted language features. One of the goals of this study, and indeed, the Bridging the Unbridgeable project in general, was to investigate the extent to which speakers' attitudes reflect prescriptive attitudes to usage found in usage guides. In the concluding section of Chapter 7, I summarised the results from the survey data by way of a visual representation of the ratings and recognition level for the features investigated (see Figure 7.8). Here, I will compare this visual representation with one based on the analysis of treatment and attitude expressions in usage guides. These two visual representations are given in Figure 8.1. The figure provides a schematic representation of the way in which the linguistic features investigated are ordered in terms of acceptability, or positive evaluation on the basis of the analysis of speakers' attitudes (top figure, also discussed in Section 7.9) and on the basis of the analysis of the usage guides treatment and attitudes (bottom figure; see also Chapter 5). The two figures show that while usage guides tend to be more negative in their attitudes than the respondents, the general pattern is very similar. The split infinitive, and to some extent object *I*, are at the positive end of the continuum, while the discourse particle *like*, negative concord, and subject *me* are at the negative end. The two features for which there seems to be some difference are *ain't*, which is more negatively rated by speakers than in usage guides, and *literally*, which is more positively rated by speakers than in usage guides.

Finally, the fourth concluding point I wish to make is that this study has shown that the influence of prescriptivism on actual language use is almost always non-existent in the long term. Where this influence is felt is usually in a short time frame, and it is reflected in the register-constrained variation in the use of specific features. Furthermore, as a potential constraint on language variation, prescriptive influence

is weaker than other constraints, whether intralinguistic or extralinguistic. If we take the model for the hierarchy of variables (i.e. linguistic, stylistic, and social) proposed by Preston (1991) as a useful interpretive framework, it can be argued that when prescriptivism is an identifiable constraint on variation, it is mostly a relatively weak constraint, the influence of which is generally (though there are exceptions) superseded by other intralinguistic and extralinguistic constraints.

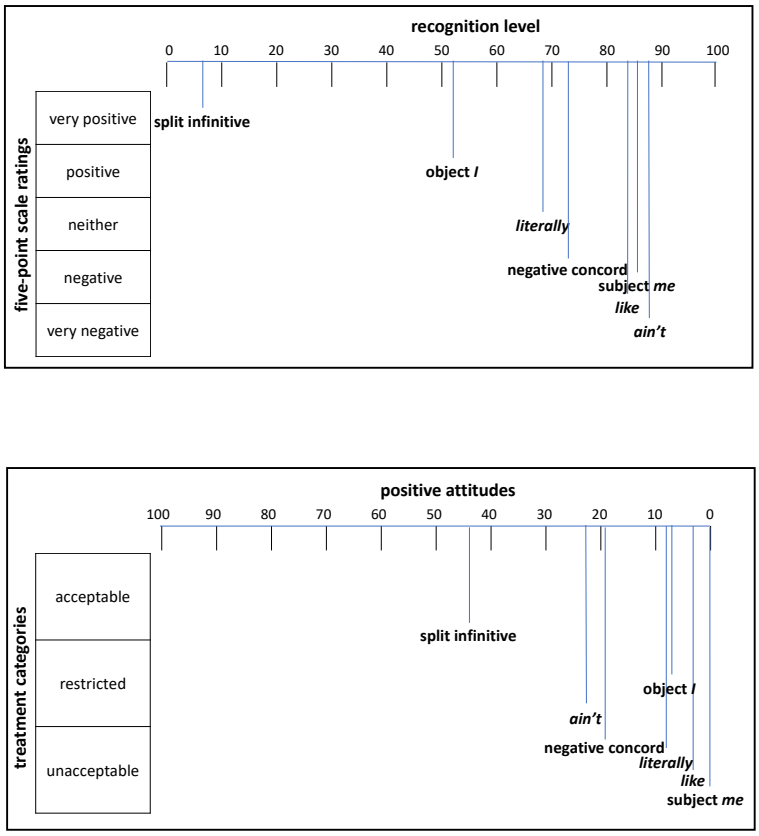

Figure 8.1: Top: schematic representation of the acceptability of usage problems on the basis of speakers' median ratings and average recognition level
Bottom: schematic representation of the acceptability of usage problems on the basis of usage guide treatment and percentage of POSITIVE attitudes

Although this study has shown that the influence of prescriptivism on language

practice may not be robust, it is nevertheless still present, and accounting for this influence in a nuanced way is necessary in studies of variation in the context of prescriptively targeted features. The study has also shown that the twentieth century is indeed the period of the prescription stage in American English (Milroy and Milroy 2012), and usage guides certainly perform a function in regulating language practice specifically in the context of standard written registers. However, there are strong indicators that usage guides are changing over time, as a result of a greater sensitivity to changing norms in language use, as well as changing attitudes. A case in point here is the acceptance of split infinitives. Finally, the application of the three-pronged approach used in this study has proved suitable for uncovering the complex nature of the interplay between prescriptive attitudes on the one hand, and language use and language users on the other.

Appendix A: Usage problems in the HUGE database

List of usage problems in the HUGE database

The following is a list of all usage problems included in the HUGE database, which I mention in Chapters 1, 3, and 4. The table contains the usage feature, the definition as included in the database, and the number of guides which treat that particular feature. The definitions have not been changed in any way, which accounts for some inconsistencies in punctuation. A shorter version of this table, excluding the definitions, can be found in Chapter 3.

usage feature	**definition**	**no. guides**
shall / *will*	use of *will* / *would* or *shall* / *should* to indicate futurity / intention / promise / threat	65
different to / *than* / *from*	variability in the choice of the particle after *different*(*ly*) (*to* or *than vs. from*)	63
who / *whom*	use of interrogative *who* or *whom* in initial position	63
lay / *lie*	use of the verbs *to lay* and *to lie*	63
only	the occurrence of only elsewhere than immediately next to the word/words it modifies (cf. Mittins et al. 1970:58)	62
split infinitive	anything inserted between the infinitive-marker *to* and the verb-form itself (Mittins et al. 1970:69)	62
I for *me*	use of subject pronouns where grammar is thought to demand the objective forms (Mittins et al. 1970:89)	61

singular *they*	use of *they* or *their* as a common-sex singular pronoun	59
less / *fewer*	use of *less* or *fewer* when referring to a number	58
none in plural context	the occurrence of *none* with plural verb where a singular is thought to be more appropriate	55
data is / *are*	use of etymologically plural subject form with a singular or plural verb. Forms: *data*, *media*, *phenomena*, *stigmata*, etc.	54
disinterested / *uninterested*	use of *disinterested* or *uninterested* to mean not interested	53
neither ... nor ... are / *is*	use of singular or plural verb with coordinated singular subjects (in negative context)	53
try and / *to*	use of *try and* or *try to*	53
like / *as*	use of *like* or *as* to introduce a clause of comparison	52
nouns of multitude	interpretation of nouns of multitude as singular or plural	52
very unique	use of *very* as an intensifier with an adjective of absolute rather than gradable meaning	52
apostrophe	apostrophes improperly used for plurals *vs.* possessives	52
a / *an*	choice of indefinite article form before words starting with *u*- or *h*-	52
both ... and	proper use of correlative conjunctions	52
between / *among*	use of *between* or *among* when referring to more than two parties	51
slow / *slowly*	use of *slow* or *slowly* as the adverbial form of the adjective *slow*; the use of flat adverbs	51
who(m) / *which* / *that*	choice of relative pronoun *who(m)* or *that*/*which* referring to a human antecedent	51
preposition at end of sentence	the occurrence of a preposition at the end of the sentence rather than before the noun phrase it modifies	50
aggravate	use in sense of to annoy, to irritate	50
snuck and *dove*	choice between weak or strong forms for the past tense or the past participle	50
dangling participle	placement of participle away from the subject of its root verb, or elision of that subject	49
was / *were*	use of indicative form *was* rather than subjunctive form *were* in subjunctive contexts	49
me for *I*	use of objective pronouns in subject position	49

foreign plurals	Anglicised and native singular and plural forms of foreign words	49
due to / *owing to*	use of *due to* or *owing to* as a preposition / adverb in contexts of causality	48
effect / *affect*	use of the verbs *effect* and *affect*	48
infer / *imply*	use of *infer* with meaning *imply* and vice versa	47
literally	use of *literally* as an intensifier in non-literal context	47
alright / *all right*	use of *alright* or *all right* as one word or two	46
this / *these sort of*	use of a plural demonstrative with a singular noun phrase that has a collective plural sense	46
compound subject	choice of compound subjects joined by a coordinator as singular or plural	45
double negatives	use of more than one negative particle to negate the same clause	44
that / *which*	choice between relative pronouns *that* or *which* in relative clauses	44
mutual	use of *mutual* (= reciprocal) and *common* (= shared)	43
can / *may*	choice between *can* and *may* as deontic auxiliary in requests	43
farther / *further*	choice between *further* and *farther* as the comparative of *far*	43
-ic / *-ical*	use of *-ic* or *-ical* to form adjectives from nouns	42
lend / *loan*	use of *loan* as a verb with the meaning of *lend*	42
me / *myself*	use of a reflexive pronoun in a non-reflexive context	42
each other / one another	use of *each other* or *one another* when referring to more than two parties	41
it is I / *it is me*	use of objective or subjective pronoun after *be* in copulative clause	41
reason is because	use of *the reason is because* for *the reason is that*	41
if / *whether*	choice between *if* and *whether* to express a condition or an alternative	41
your / *you're*	confusion of possessive pronoun with personal pronoun-verb contraction	41
one of those who	agreement with subject *one of those who*	40
one ... one / *he*	use of specific pronoun *he* to refer to antecedent unspecific pronoun *one*	39

them / *their* + V-*ing*	use of objective or possessive pronoun in gerundive construction	39
ain't	use of *ain't* to mean *isn't* or *aren't*	39
compare with	choice between the prepositions *with* or *to* with the verb *compare*	39
hopefully	use of *hopefully* as sentence modifier with the sense of *it is to be hoped that ...*	38
than I / *me*	use of *than* as a conjunction – combining with *I* – or as a preposition – combining with *me*	38
former / *latter*	use of *former* and *latter* in the sense *first and last* of three or more things rather than *first and second* of two things	38
equally as	use of *equally as* + Adj for *equally* + Adj	38
decimate	use in the sense of to destroy large proportion of or even obliterate, rather than destroy one tenth of	36
alternative	use of *alternative* to indicate more than two options	36
flaunt / *flout*	use of the verbs *to flaunt* and *to flout*	35
off of	use of *off of* for *of*	35
false attraction	subject-verb agreement when there is an intervening NP.	35
on to / *onto*	use of *on to* or *onto* as two words or one	34
either is / *are*	single or plural verbs with *either*	34
most perfect	comparative or superlative with absolute adjective	34
whose / *of which*	choice between *whose* or *of which* as possessive pronoun with non-human antecedents	34
(not) as / *so far as*	variation between *so ... as* and *as ... as* depending on whether or not it is preceded by a negative particle	33
may / *might*	choice between *may* and *might* as epistemic auxiliary	33
from thence	use of deictic preposition with deictic adverbs *thence* / *hence* / *whence*	32
like / *as if*	use of *like* with the sense of *as if*	31
either of them / *each of them*	choice between *either* or *each* referring to two or more; *either* meaning each of two or both	31
But / *And*	begin sentence with *But* or *And*	31
have to / *have got to*	use of *get* in a sense apart from *obtain* or *acquire*, in expressing necessity or obligation	30
comma splice	using a comma rather than semicolon to connect two main clauses	30

subject-complement	subject-subject complement agreement (in a copular construction)	30
averse to / from	use of *to* or *from* with *averse* and *adverse*	29
in / *into*	choice of preposition for position / movement	29
spoonsful	plural formation of the type N-ful	29
either ... or ... (*or ...*)	use of the construction *either ... or (... or)* to refer to more than two entities.	28
providing / *provided*	use of *providing* and *provided* with the same sense	28
family is / *are*	use of plural or singular verb with nouns that can be interpreted as both mass or count nouns	27
very / *much amused*	use of *very* or *much* to qualify an adjective or past participle. For the purist, *very* is an intensifier; only qualities – and not actions – may be intensified, and therefore *very* can qualify an adjective but not a past participle	26
superlative comparison	use of the superlative to compare two things	26
-lily adverbs	formation of adverbs from adjectives ending in *-ly*	26
hoi polloi	inclusion of definite article in hoi polloi (= 'the many')	25
contemporary	use of *contemporary* in the sense of present day, up-to-date, or referring to a previously established time frame	24
likely	adverbial use of *likely*	24
could of	use of *could / would / should / must of* for *could / would / should / must have*	24
dare	*dare* as marginal modal verb	23
more warmer	the use of double comparatives or double superlatives	23
in / *under circumstances*	use of the preposition *in* or *under* with the word *circumstances*	21
'd rather	choice of auxiliary *have* or *will*, and meaning of contracted form	21
there's	using a plural subject with a singular verb form in case of is a dummy subject with contracted verb form, there's	21
corporeal / *corporal*	use of *corporeal* and *corporal*	20
learn / *teach*	use of *learn* and *teach*	20

as well (as) ... or better than	elision of *as* in coordination of the phrase *as well as* with *better than*	19
pretty	use of *pretty* as a degree adverb	19
the two first	choice between *the two first / last* and *the first / last two*	19
upon	choice between *upon* and *on*	16
double passive	use of constructions with multiple passives	16
thusly	use of *thusly* instead of *thus*	16
like / the way	use of *(in) the way* or *like* in expressions of analogy	15
have went	use of the past participle form for simple past tense or vice versa	15
split auxiliaries	insertion of element between auxiliaries or auxiliary and main verb (split compound verbs)	15
off / from	use of *off* or *from*	14
quicker / more quickly than	use of *quicker* or *more quickly* as the comparative form of *quick*	14
omission of relative pronoun	omission of the relative pronouns *who*, *that*, *which* in non-restrictive relative clauses	14
gay	use of *gay* to mean 'homosexual'	14
thankfully	use in the sense 'in a thankful way' not 'let us be thankful that'	14
demonstrative *them*	use of personal pronoun *them* as a demonstrative pronoun	14
meet with / meet up with	inclusion of preposition *up* in the phrasal verb *meet with*	13
all that / so easy	use of *all that* or *so* to modify an adjective / adverb	12
hisself	use of *hisself* for *himself*	12
at / in	choice of *at* rather than *in* or *on* with *university*, *school* and similar nouns	11
less / least	choice of *less* and *least* when referring to two or more things	9
when	*when* or *where* used in defining as in *X is when/where Y*	8
get thither	use of deictic verb that also indicate location with deictic adverbs that do the same	6
evenings and Sundays	coordination of prepositional phrases with different elided prepositional heads	4
at (the) university	omission of *the* before *university* and some similar nouns	4

momentarily	the use of *momentarily* to mean 'for a moment or short time'/ 'in an instant' / 'from moment to moment' / 'at any moment'	1

Appendix B: Usage guides in American English

Anderson, Charles. 1931. *Common Errors in English Corrected*. New York: Van Nostrand Co.

Anonymous. 1886. *Many Mistakes Mended: Containing Two Thousand Five Hundred Corrections in Speaking, Pronouncing, and Writing the English language*. New York: N. Tibbals & Sons.

Anonymous. 1953. *The Associated Press Stylebook*. New York: Associated Press.

Anonymous. 1996. *The American Heritage Book of English Usage*. Boston: Houghton Mifflin Harcourt.

Bailie, John and Moyna Kitchin. 1979. *The Newnes Guide to English Usage*. Feltham, Middlesex: Newnes Books.

Baker, Josephine Turck. 1909. *Correct English in the Home*. Chicago: Correct English Publishing Company.

Baker, Sheridan. 1962. *The Practical Stylist*. New York: Thomas Y. Crowell Co.

Baker, Sheridan. 1984. *The Complete Stylist and Handbook*. New York: Harper & Row.

Ball, Francis K. 1923. *Constructive English*. Boston: Ginn.

Barnard, Ellsworth. 1979. *English for Everybody*. Amherst, Mass.: Dinosaur Press.

Barnes, Duane C. 1948. *Wordlore*. New York: E.P. Dutton & Co.

Barzun, Jacques. 1985. *Simple & Direct*. New York: Harper & Row.

Barzun, Jacques. 1986. *A Word or Two Before You Go*. Middletown, Conn.: Wesleyan University Press.

Bawer, Bruce. 1987. *The Contemporary Stylist*. New York: Harcourt Brace Jovanovich.

Belanoff, Pat, Betsy Rorschach, and Mia Oberlink. 1986. *The Right Handbook*. Upper Montclair, N.J.: Boynton/Cook.

Bell, James K. and Adrian A. Cohn. 1972. *Bell & Cohn's Handbook of Grammar, Style, and Usage*. Beverly Hills: Glencoe Press.

Bell, Janis. 2008. *Clean, Well-lighted Sentences: A Guide to Avoiding the Most Common Errors of Grammar and Punctuation*. New York: W. W. Norton & Co.

Bell, Ralcy H. 1902. *The Worth of Words*. New York: Grafton Press.

Bernstein, Theodore M. 1958. *Watch Your Language*. Manhasset, N.Y.: Channel Press.

Bernstein, Theodore M. 1962. *More Language That Needs Watching*. Manhasset, N.Y.: Channel Press.

Bernstein, Theodore M. 1965. *The Careful Writer*. New York: Atheneum.

Bernstein, Theodore M. 1971. *Miss Thistlebottom's Hobgoblins: The Careful Writer's Guide to the Taboos, Bugbears, and Outmoded Rules of English Usage*. New York: Simon & Schuster, Inc.

Bernstein, Theodore M. 1977. *Dos, Don'ts and Maybes of English Usage*. New York: Time Books.

Berry, Thomas E. 1971. *The Most Common Mistakes in English Usage*. New York: McGraw-Hill Book Co.

Blount, Alma and Clark S. Northup. 1931. *Grammar and Usage*. New York: Prentice-Hall, Inc.

Boyle, Joe. 1982. *The Federal Way with Words*. Washington, D.C.: Twain Pub.

Boyle, Toni and K. D. Sullivan. 2005. *The Gremlins of Grammar: A Guide to Conquering the Mischievous Myths that Plague American English*. Chicago: McGraw-Hill.

Bremner, John B. 1980. *Words on Words*. New York: Columbia University Press.

Brown, Ivor. 1945. *A Word in Your Ear and Just Another Word*. New York: E.P. Dutton & Co.

Brown, Ivor. 1948. *I Give You My Word and Say the Word*. New York: E.P. Dutton & Co.

Brown, Ivor. 1951. *No Idle Words and Having the Last Word*. New York: E.P. Dutton & Co.

Bryant, Margaret M. 1962. *Current American Usage*. New York: Funk & Wagnalls Co.

Bryant, William C. 1877. *Index Expurgatorius*. Originally unpublished, but reprinted in Bernstein (1971).

Buckley, William F. 1996. *Buckley: The Right Word*. New York: Random House.

Butt, Margot and Linda Lane. 1973. *Collins Gem Dictionary of American Usage*. London: Collins.

Campbell, Cathy. 2008. *The Giggly Guide to Grammar*. Shoreham, Vt.: Discover writing press.

Carew, Paul T. 1923. *The Art of Phrasing in Composition*. Boston: Stratford Co.

Carter, Ronald and Michael McCarthy. 2006. *Cambridge Grammar of English: A Comprehensive Guide*. Cambridge and New York: Cambridge University Press.

Casagrande, June. 2006. *Grammar Snobs are Great Big Meanies: A Guide to Language for Fun and Spite*. New York: Penguin Books.

Cazort, Douglas. 1997. *Under the Grammar Hammer*. Los Angeles: Lowell House.

Cioffi, Frank L. 2015. *One Day in the Life of the English Language: A Microcosmic Usage Handbook*. Princeton: Princeton University Press.

Claiborne, Robert. 1986. *Saying What You Mean: A Commonsense Guide to American Usage*. New York: Ballantine Books.

Clark, Stewart and Graham E. Pointon. 2016. *The Routledge Student Guide to English Usage*. London and New York: Routledge.

Cody, Sherwin. 1903. *The New Art of Writing & Speaking the English Language*. New York: SunDial Press, Inc.

Cody, Sherwin. 1905. *Dictionary of Errors*. New York: Funk & Wagnalls.

Colby, Frank. 1944. *The Practical Handbook of Better English*. New York: Grosset & Dunlap.

Cook, Claire K. 1985. *Line by Line: How to Improve Your Own Writing*. Boston: Houghton Mifflin Co.

Copperud, Roy H. 1960. *Words on Paper: A Manual of Prose Style for Professional Writers, Reporters, Authors, Editors, Publishers, and Teachers*. New York: Hawthorn Books, Inc.

Copperud, Roy H. 1964. *Webster's Dictionary of Usage and Style*. New York: E.P. Dutton, Inc.

Copperud, Roy H. 1970. *American Usage: The Consensus*. New York: Van Nostrand Reinhold.

Corder, Jim W. 1981. *Handbook of Current English*. Glenview, Ill.: Scott, Foresman.

Dahlskog, Helen, ed. 1972. *A Dictionary of Contemporary and Colloquial Usage.* New York: Avenel Books.

Davidson, Mark. 2006. *Right, Wrong and Risky: A Dictionary of Today's American English Usage*. New York and London: W.W. Norton & Co.

Davis, C. Rexford. 1936. *Toward Correct English*. New York: Appleton Century Crofts, Inc.

Dorey, J. Milnor. 1949. *Good English Made Easy*. New York: Permabooks.

Dumond, Val. 1990. *The Elements of Nonsexist Usage: A Guide to Inclusive Spoken and Written English*. New York: Prentice Hall Press.

Eggenschwiler, Jean. 1997. *Writing: Grammar, Usage, and Style*. Lincoln, Nebr.: Cliffs Notes.

Ehrlich, Eugene and Daniel Murphy. 1991. *The HarperCollins Concise Dictionary of English Usage*. New York: HarperCollins.

Eichler, Lillian. 1928. *Well-Bred English*. Garden City, N.Y.: Doubleday, Doran & Co.

Einstein, Charles. 1985. *How to Communicate: The Manning Selvage & Lee Guide to Clear Writing and Speech*. New York: McGraw-Hill.

Elsbree, Langdon and Frederick Bracher. 1968. *Brief Handbook of Usage*. Lexington, Mass.: D.C. Heath & Co.

Erskine, John and Helen Erskine. 1910. *Written English: A Guide to the Rules of Composition*. New York: Century Co.

Evans, Bergen. 1962. *Comfortable Words*. New York: Random House.

Fennell, Francis L. 1980. *Writing Now: A College Handbook*. Chicago: Science Research Associates, Inc.

Ferguson, Don K. 1995. *Grammar Gremlins*. Lakewood, Colo.: Glenbridge Publishing Ltd.

Fernald, James. 1904. *Connectives of English Speech: The Correct Usage of*

Conjunctions, Relative Pronouns and Adverbs Explained and Illustrated. New York: Funk & Wagnalls Co.

Fernald, James C. 1911. *Helpful Hints in English.* New York: Funk & Wagnalls Co.

Fischer, James A. 1999. *Talking Correctly for Success: A Practical Guide to Sounding 'Right' for Business, Professional and Social Success*. Blue Ridge Summit, Pa.: Avant Publishing Co., revised edn.

Flesch, Rudolf. 1962. *How to Be Brief: An Index to Simple Writing*. New York: Harper & Row.

Flesch, Rudolf. 1964. *The ABC of Style: A Guide to Plain English*. New York: Harper & Row.

Flesch, Rudolf. 1983. *Lite English: Popular Words That Are OK to Use*. New York: Crown Pubs., Inc.

Foulke, Adrienne. 1964. *English for Everyone*. New York: Pocket Books.

Freeman, Morton S. 1983. *A Treasury for Word Lovers*. Philadelphia: ISI Press.

Freeman, Morton S. 1998. *The Wordwatchers Guide to Good Grammar & Word Usage: Authoritative Answers to Today's Grammar and Usage Questions*. Medford, N.J.: Plexus Publications.

Garner, Bryan A. 1987. *A Dictionary of Modern Legal Usage*. New York: Oxford University Press.

Garner, Bryan A. 2016. *The Chicago Guide to Grammar, Usage, and Punctuation*. Chicago: Chicago University Press.

Goodman, Roger. 1962. *A Concise Handbook of Better English*. New York: Bantam Books.

Gordon, Karen E. 1993. *The New Well-Tempered Sentence: A Punctuation Handbook for the Innocent, the Eager and the Doomed*. Boston: Houghton Mifflin Harcourt.

Gordon, Karen Elizabeth. 1997. *Torn Wings and Faux Pas*. New York: Pantheon Books.

Greenwood, Joseph A. 1969. *Find It in Fowler: An Alphabetical Index*. Princeton, N.J.: Princeton Univ. Press.

Greer, Garland and Easley S. Jones. 1924. *The Century Collegiate Handbook*. New York: Century.

Griswold, William M. 1882. *A Manual for Misused Words*. Bangor, Maine: Q.P. Index Publishers.

Grossman, Ellie. 1997. *The Grammatically Correct Handbook*. New York: Hyperion.

Gummere, John F. 1989. *Words & C.* Haverford, Pa.: Bolchazy-Carducci Publishers.

Guth, Hans. 1985. *New English Handbook*. Belmont, Calif.: Wadsworth.

Gutierrez, Lucy. 2015. *English is Not Easy: A Visual Guide to the Language*. New York: Gotham Books.

Haber, Tom B. 1942. *A Writer's Handbook of American Usage*. New York: Longmans, Green & Co.

Haberstroh, Emil F. 1931. *How to Speak Good English*. Racine, Wis: Whitman Publishing Co.

Hale, Constance and Jessie Scanlon. 1999. *Wired Style: Principles of English Usage in the Digital Age*. New York: Broadway Books.

Hennesy, James A. 1917. *The Dictionary of Grammar*. New York: Funk & Wagnalls.

Herbert, A.P. 1935. *What a Word!* Garden City, N.Y.: Doubleday, Doran & Co.

Hodgson, John C. and Mary E. Whitten. 1984. *The Harbrace College Handbook*. New York: Harcourt Brace Jovanovich.

Hook, J.N. 1962. *Hook's Guide to Good Writing*. New York: Ronald Press Co.

Hook, J.N. and E.G. Mathews. 1956. *Modern American Grammar and Usage*. New York: Ronald Press Co.

Hopper, Vincent F., Cedric Gale, and Ronald C. Foote. 1984. *A Pocket Guide to Correct Grammar*. New York: Barron's Educational Series, Inc.

Hotchkiss, George B. and Edward J. Kilduff. 1914. *Handbook of Business English*. New York: New York University Book Store.

Howard, Philip. 1977. *New Words for Old*. New York: Oxford University Press.

Hudson, Randolph H. 1966. *A Modern Handbook of American English*. Chicaco: Science Research Associates, Inc.

Hyck, Wynford. 2009. *The Basics of English Usage*. London and New York: Routledge.

Irmscher, William F. 1976. *The Holt Guide to English*. New York: Holt, Rinehart and Winston.

Ivers, Mitchell. 1991. *The Random House Guide to Good Writing*. New York: Ballantine Books.

Janis, J. Harold. 1984. *Modern Business Language and Usage in Dictionary Form*. Garden City, N.Y: Doubleday.

Jenkins, Evan. 2007. *That or Which, and Why: A Usage Guide for Thoughtful Writers and Editors*. New York: Routledge.

Jones, George A. 1967. *English for Adventurers*. New York: Pageant Press, Inc.

Kilduff, Edward J. 1941. *Words and Human Nature: How to Choose and Use Effective Words*. New York: Harper & Bros.

Kilpatrick, James J. 1984. *The Writer's Art*. Kansas City: Andrews, McMeel & Parker.

Kilpatrick, James J. 1985. *The Ear Is Human: A Handbook of Homophones and Other Confusions*. Kansas City: Andrews, McMeel & Parker.

Kirtland, Elizabeth. 1968. *Write It Right: A Handbook of Homonyms*. New York: Golden Press.

Langan, John and Janet M. Goldstein. 1997. *English Brushup*. New York: McGraw-Hill.

Larsen, Karen. 1994. *The Miss Grammar Guidebook*. Lake Oswego: Oregon State Bar.

Lederer, Richard. 1987. *Anguished English: An Anthology of Accidental Assaults upon Our Language*. Charleston, S.C.: Wyrick & Co.

Leggett, Glenn H., Carl D. Mead, and Melinda G. Kramer. 1954. *Prentice Hall Handbook for Writers*. Englewood Cliffs, N.J.: Prentice Hall.

LePan, Don. 1988. *The Broadview Book of Common Errors in English: A Guide to Righting Wrongs*. Peterborough, Ontario/Orchard Park, N.Y: Broadview.

LePan, Don. 1992. *The Broadview Book of Common Errors in English*. Peterborough, Ontario: Broadview Press.

Lewis, Norman. 1948. *Better English*. New York: Thomas Y. Crowell & Co.

Lieberman, Leo. 1962. *Dictionary of Correct English Usage*. Paterson, N.J.: Littlefield Adams & Co.

Lodwig, Richard R. and Eugene F. Barrett. 1967. *The Dictionary and the Language*. New York: Hayden Book Co.

Long, J.H. 1888. *Slips of Tongue and Pen*. New York: D. Appleton & Co.

Lurie, Charles. 1927. *How to Say It*. New York: G.P. Putnam's.

Lynch, Jack. 2008. *The English Language: A User's Guide*. Newburyport, Mass.: Focus Publishing/R Pullins Co.

MacCracken, Henry N. and Helen E. Sandison. 1917. *Manual of Good English*. New York: Macmillan & Co.

Maggio, Rosalie. 2001. *How to Say it: Style Guide*. Paramus: Prentice Hall Press.

Magio, Rosalie. 1987. *The Nonsexist Wordfinder*. Phoenix: Oryx.

Manly, John M. and John A. Powell. 1913. *A Manual for Writers*. Chicago: University of Chicago Press.

Manly, John M. and Edith Rickert. 1923. *The Writer's Index of Good Form and Good English*. New York: Henry Holt & Co.

Martin, Phyllis. 1977. *Word Watcher's Handbook*. New York: David McKay.

Mathews, William. 1876. *Words: Their Use and Abuse*. Chicago: S.C. Griggs & Co.

McElfresh, G.E. and E.C. Ingalls. 1940. *Everyday English*. New York: Thomas Nelson & Sons.

McMahan, Elizabeth and Susan Day. 1980. *The Writer's Rhetoric and Handbook*. New York: McGraw-Hill.

McMahan, Elizabeth and Susan Day. 1984. *The Writer's Rhetoric and Handbook*. New York: McGraw-Hill, 2nd edn.

Miller, Casey and Kate Swift. 1980. *The Handbook of Nonsexist Writing*. New York: Lippincott.

Montgomery, Michael and John Stratton. 1981. *The Writer's Hotline Handbook*. New York: New American Library.

Nickles, Harry G. 1974. *Dictionary of Do's and Don'ts for Writers and Speakers.* New York: Greenwich House.

Nurnberg, Maxwell. 1972. *Questions You Always Wanted to Ask About English ... but Were Afraid to Raise Your Hand*. New York: Washington Square Books.

O'Connor, Patricia and Tom Stiglich. 2007. *Woe Is I Jr: The Younger Grammarphobe's Guide to Better English in Plain English*. New York: Penguin.

O'London, John. 1925. *Is It Good English?* New York: G.P. Putnam's Sons.

Opdycke, John B. 1941. *Get It Right! A Cyclopedia of Correct English Usage*. New York: Funk & Wagnalls Co.

Opdycke, John B. 1944. *Say What You Mean: Everyman's Guide to Diction and Grammar*. New York: Funk & Wagnalls Co.

Opdycke, John B. 1949. *Mark My Words: A Guide to Modern Usage and Expression*. New York: Harper & Bros.

Opdycke, John B. 1950. *The Opdycke Lexicon of Word Selection*. New York: Funk & Wagnalls Co.

Orcutt, William D. 1912. *The Writer's Desk Book*. Norwood, Mass.: Frederick A. Stokes Co.

Orcutt, William D. 1926. *The Desk Reference Book*. New York: Frederick A. Stokes Co.

Partridge, Eric. 1942. *Usage and Abusage: A Guide to Good English*. New York: Harper & Bros.

Paxson, William C. 1990. *The New American Dictionary of Confusing Words*. New York: Signet.

Pence, Raymond W. 1944. *Style Book in English*. New York: Odyssey Press.

Phythian, B.A. 1990. *A Concise Dictionary of Confusables*. New York: John Wiley & Sons.

Pocock, Guy N. 1936. *Brush Up Your Own Language*. Philadelphia: David McKay Co.

Postman, Neil, Charles Weingartner, and Terence P. Moran, eds. 1969. *Language in America*. New York: Pegasus.

Ransom, Robert W. 1911. *Hints & "Don'ts" for Writers and Copyreaders*. Chicago: Chicago Record Herald.

Rattray, David, ed. 1983. *Success with Words: A Guide to the American Language*. Pleasantville, N.Y: Reader's Digest.

Raub, Albert N. 1889. *Hints and Helps of English Grammar*. Cleveland: J.R. Holcomb & Co.

Raub, Albert N. 1897. *Helps in the Use of Good English*. Philadelphia: Albert N. Raub & Co.

Rine, George W. 1902. *The Essentials of Our Language: A Guide to Accuracy in the Use of the English Language*. Oakland: Pacific Press Publishing Co.

Rothwell, Kenneth S. 1971. *Questions of Rhetoric and Usage*. Boston: Little, Brown & Co.

Rovin, Jeff. 1994. *What's the Difference? A Compendium of Commonly Confused and Misused Words*. New York: Ballantine Books.

Runkel, Philip and Margaret Runkel. 1984. *A Guide to Usage for Writers and Students in the Social Sciences*. Totawa, N.J.: Rowman & Allanheld.

Safire, William. 1980. *On Language*. New York: Times Books.

Safire, William. 1982. *What's the Good Word?* New York: Times Books.

Safire, William. 1984. *I Stand Corrected*. New York: Times Books.

Safire, William. 1986. *Take My Word for It*. New York: Times Books.

Safire, William. 1990. *Fumblerules: A Lighthearted Guide to Grammar and Good Usage*. New York: Doubleday.

Safire, William. 1991. *Coming to Terms*. New York: Doubleday.

Safire, William. 1994. *In Love with Norma Loquendi*. New York: Random House.

Safire, William. 1997. *Watching My Language*. New York: Random House.

Schumaker, Florence. 1948. *The Student's Guide*. Cleveland, Tenn.: Church of God Publishing House.

Semmelmeyer, Madeline and Donald O. Bolander. 1955. *The Instant English Handbook*. Mundelein, Ill.: Career Institute.

Sessions, Keith W. 2013. *A Sometimes Snarky Guide to English Usage, Grammar & Punctuation*. Charleston, S.C.: CreateSpace.

Shaw, Harry. 1962. *Errors in English and Ways to Correct Them*. New York: Barnes & Noble.

Shostak, Jerome. 1968. *Concise Dictionary of Current American Usage*. New York: Washington Square Press, Inc.

Siegal, Allan M. and William G. Connolly. 1999. *The New York Times Manual of Style and Usage*. New York: Times Books.

Spector, Stephen. 2015. *May I Quote You on That?: A Guide to Grammar and Usage*. Oxford and New York: Oxford University Press.

Stratton, Clarence. 1940. *Handbook of English*. New York: McGraw-Hill Book Co.

Terputac, Thomas J. 1989. *A Handbook of English Usage*. Lawrenceville, Va.: Brunswick Publishing Corp.

Thomas, Charles S. 1941. *Your Mastery of English*. New York: P.F. Collier & Son.

Thomas, Henry. 1954. *Better English Made Easy*. New York: Greystone Press.

Thomas, Lewis. 1990. *Et Cetera, Et Cetera: Notes of a Word-Watcher*. Boston: Little, Brown & Co.

Thompson, Newton. 1947. *A Handy Guide for Writers*. St. Louis: B. Herder Book Co.

Timmons, Christine and Frank Gibney, eds. 1980. *Britannica Book of English Usage*. Garden City, N.Y.: Doubleday/Britannica Books.

Tracz, Richard F. 1991. *Dr. Grammar's Writes from Wrongs*. New York: Vintage Books.

Trask, Robert L. 2005. *Say What you Mean!: A Troubleshooter's Guide to English Style and Usage*. Boston: David R. Godine.

Trimmer, Joseph F. and James M. McCrimmon. 1988. *Writing with a Purpose*. Boston: Houghton Mifflin.

Urdang, Laurence. 1988. *The Dictionary of Confusable Words*. New York: Ballantine Books.

Utter, Robert P. 1914. *A Guide to Good English*. New York: Harper & Bros.

Venolia, Jan. 1982. *Write Right! A Desk Drawer Digest of Punctuation, Grammar and Style*. Berkeley: Ten Speed Press.

Vizetelly, Frank H. 1933. *How to Use English: A Guide to Correct Speech and Writing*. New York: Funk & Wagnalls Co.

Wagnalls, Funk. 1917. *Faulty Diction as Corrected by the Funk & Wagnalls New Standard Dictionary of the English Language*. New York: Funk & Wagnalls Co.

Wallwork, Adrian. 2013. *English for Research Usage, Style, and Grammar*. New York: Springer.

Watt, William. 1967. *A Short Guide to English Usage*. Cleveland: World.

Weekley, Ernest. 1931. *Cruelty to Words, or First Aid for the Best-Seller*. New York: E.P. Dutton & Co.

Weseen, Maurice H. 1922. *Everyday Uses of English*. New York: Thomas Y. Crowell.

Weseen, Maurice H. 1928. *Crowell's Dictionary of English Grammar and Handbook of American Usage*. New York: Thomas Y. Crowell Co.

Weseen, Maurice H. 1932. *Words Confused and Misused*. New York: Thomas Y. Crowell Company.

Whitford, Robert C. and James R. Foster. 1955. *Concise Dictionary of American Usage and Grammar*. New York: Philosophical Library.

Williams, Deborah K. 1995. *NTC's Dictionary of Easily Confused Words*. Lincolnwood, Ill.: National Textbook Co.

Williams, Joseph M. 1981. *Style: Ten Lessons in Clarity and Grace*. Glenview: Scott, Foresman & Co.

Willis, Hulon. 1964. *Structure, Style, and Usage: A Guide to Expository Writing*. New York: Holt, Rinehart & Winston.

Wilson, Robert F., ed. 1982. *The Macmillan Handbook of English*. New York: Macmillan.

Witherspoon, Alexander M. 1943. *Common Errors in English and How to Avoid Them*. New York: New Home Library.

Wood, Clement. 1944. *More Power to Your Words!* New York: Prentice-Hall, Inc.

Appendix C: Corpus data extraction

Extraction of corpus occurrences of features

In this appendix I explain how I extracted, or identified, all occurrences of each of the features analysed from corpus data used for the analyses presented in Chapter 6. I used the full-text data from the Corpus of Contemporary American English and the Corpus of Historical American English, and searched the two corpora and extracted all relevant cases of the features with the programming language Python,[2] and the command shell Jupyter Notebooks.[3] In addition to the main features investigated in this thesis, an additional number of features were used for the analysis of prescriptive influence on the occurrence of the split infinitive. The occurrences of all of these features were searched for in the corpora, and the relevant data extracted from the corpora in a similar way.

The first part of this appendix provides the details for the extraction of the occurrences of variants of the main features analysed, i.e. *ain't*, discourse particle *like*, *literally*, negative concord, object *I*/subject *me*, and the split infinitive. The second part of the appendix provides the details for the extraction of the additional language features, i.e. *and/but* at the beginning of the sentence, singular *data*, *hopefully*, *less* with plural nouns, *these kind/sort of*, *try and*, plural *none*, passives, *shall*, and *whom*.

The general procedure for each of the features described below was first to search for and extract each relevant lemma (e.g. *data*) from the

[2] See https://www.python.org.
[3] See https://jupyter.org.

word/lemma/part-of-speech-tagged files from the full-text corpus data.[4] After all occurrences of that particular lemma were extracted from the corpus files, the concordance lines were split into individual tokens. Each token has an accompanying part-of-speech tag, on the basis of which each occurrence could be further classified, and relevant occurrences could be identified. This appendix outlines how specific features were identified and extracted from the corpus texts, i.e. how this process was conducted.

Extraction of the main features used in the analysis

Ain't – All occurrences of *ain't* were extracted from the corpus files using an appropriate regular expression (essentially the string `ai n't`, as it appears in the corpus files). The occurrences of *ain't* were extracted along with the immediate preceding and following context, creating a concordance for all the uses of *ain't*. Each of these occurrences was then classified into two groups: *ain't* for *be not*, if *ain't* was followed by an *-ing* form or a noun phrase; and *ain't* for *have not*, if *ain't* was followed by a past participle. In addition to these, all occurrences of present *be not* forms and present *have not* forms were also extracted, in order to establish the total number of potential environments in which *ain't* could have occurred. In these cases, all forms of the respective verbs were taken into consideration, i.e. full forms, copula-contracted forms, and *not*-contracted forms (see Section 3.3 for more on these distinctions).

Like – First, all occurrences of *like* were extracted from the corpus data using a regular expression; the immediate preceding and following context was also extracted. Occurrences of discourse particle *like* were identified on the basis of punctuation in the corpus, i.e. commas preceding and following *like*, and further selected from the initial dataset. This was done because the initial attempt to extract these occurrences on the basis of their part-of-speech tags revealed that tags tended to be quite messy, which would have resulted in a high level of inaccuracy in the analysed data. However, as most of the occurrences of discourse particle *like* were preceded and followed by a comma, this was then used as a more practical way of extracting these occurrences.

Literally – The initial extraction of *literally* from the corpus files was the same as that for all the other features: all occurrences of *literally* were extracted using regular expressions, and the immediate preceding and following context was also included in the concordances thus produced. Further disambiguation of these cases of *literally* was done on the basis of a tripartite distinction of uses. The first one refers to the primary

[4] See https://www.corpusdata.org for more information on the full-text data.

use of *literally*, referring to what a word or a phrase means, how something is to be understood, or interpreted, pointing out that something should be taken literally, and not figuratively, or saying that something is literally true. The second use of *literally* included cases of dual use, where *literally* is used with idiomatic or figurative expressions to force a literal meaning of a conventionally non-literal expression. Finally, the third use of *literally*, the so-called non-literal use, refers to cases where *literally* is used with non-literal expressions, hyperbolic or figurative, in which it is obvious from the context that a literal meaning is not possible. As I explain in Section 4.4, the disambiguation of these uses of *literally* consisted of two parts. For the first part of the analysis, concordance lines in which *literally* is preceded or followed by words indicating that the use of *literally* is the primary use were identified. Such words included any form of *mean*, *interpret*, *read*, *say*, *write*, *translate*, and similar forms. Some manual check-up was used to identify these forms, but the disambiguation itself was done automatically in Python (this involved using 'if-functions'). On the basis of these forms, concordance lines in which *literally* was preceded or followed by these forms were classified as primary, while all the other ones were classified as secondary, here including both dual and non-literal cases. The second part of the analysis involved manually disambiguating cases in a systematic random sample of the entire dataset for *literally*; the sample was drawn by selecting every fifth occurrence from the overall set of occurrences of *literally*. These cases were then manually classified into the three categories mentioned above.

Negative concord – For the extraction of cases of negative concord, I identified occurrences of negated verbs used with indefinites in post-verbal position. I selected the indefinites *anything/nothing*, *anyone/no one*, and *anybody/nobody*. This was done by extracting all occurrences of these six words from the corpora separately, and subsequently checking whether they were preceded by a negated verbs. The part-of-speech tags of the five preceding words were checked. If one of them was identified as *not/n't* or *never*, the instance in question was identified as a case of negation with a post-verbal indefinite. The concordance lines for *anything*, *anyone*, and *anybody* in which a negative element was identified at one of the five preceding slots were classified as single negation cases, e.g. *I haven't seen anybody*, while the concordances for *nothing*, *no one*, and *nobody* in which a negative elements was identified at one of the five preceding slots were classified as double negation cases, e.g. *I haven't seen nobody*. The remaining cases in the concordances for *nothing*, *no one*, and *nobody* in which a negative element was not identified were then checked for verbs. If a non-negated verb was identified in one of the five preceding slots, those

cases were classified as cases of single negation with *nothing*, *no one* and *nobody*, e.g. *I saw nobody*.

Pronouns: object *I* and subject *me* – In order to extract occurrences of object *I*, the first step was to identify coordinated phrases functioning as objects, in which one of the coordinated phrase-constituents is realised with *I*. Taking into account all such cases proved difficult for two reasons. The first reason was that the syntactic function of a coordinated phrase is not always straightforwardly determined automatically on the basis of a part-of-speech-tagged corpus; some kind of manual disambiguation subsequent to the automatic extraction of cases was therefore necessary. Secondly, cases in which the coordinated phrase involves two pronouns can reasonably be expected to behave differently with respect to the realisation of *I/me* than cases in which the first element is a noun. For example, in a sentence such as *He brought a present for him and me*, the realisation of *me* is likely to be affected by the realisation of *him*, while in a sentence such as *He brought a present for Anne and me*, this constraint is not present. In order to minimise such additional constraints on the realisation of *I/me* in object coordinated phrases, I limited my search to object coordinated phrases in which one of the constituents is a proper noun, which was done on the basis of identifying all cases in which *I* or *me* are coordinated with a proper noun, as tagged in the part-of-speech-tagged corpus data. This was done by first extracting occurrences of the strings 'and I' and 'and me'. The function of these strings was then determined automatically on the basis of preceding and following elements. For instance, cases where the phrase NP-proper + *and* + *I* is both immediately preceded by a preposition or a verb and immediately followed by an element other than a verb were identified as cases of coordinated phrases functioning as objects in the sentence. The analysis also did not take into account cases where the coordinated phrase is a complement to *to be*, or where they are part of subordinated clauses with *for... to...*, because such cases introduce additional constraints to the choice of form.

Split infinitive – All occurrences of the infinitive marker *to* were extracted from the corpus files and were subsequently disambiguated to distinguish between cases where *to* is immediately followed by a verb, and cases where *to* is followed by another part of speech and then by a verb, in other words, *to* + V, and *to* + [WORD] + V. This was done in order to first examine the types of elements that can come between the *to* and the verb, so that those same elements could then be checked in pre-verbal and post-verbal positions. After examining the most common 'splitters', i.e. elements that can modify an infinitive, and that can vary between pre-verbal (*really to know*), post-verbal (*to know really*), or medial position (*to really know*), only those cases which the adverb

can be placed pre-verbally or post-verbally were taken into account for the analysis. Since these elements included most lexical adverbs ending in *-ly*, these cases were used for the variationist analysis of patterns of use of the split infinitive (see Sections 4.4 and 6.7). Infinitives split by the negator *not* were not taken into account, because they follow different patterns of variation, in that *not* can only take a pre-verbal (*he decided not to join*) or a split (*he decided to not join*) position; this is different from adverbial splitters, as explained above. Finally, the variable context was established as infinitives modified by one element only; thus, infinitives split by two elements, or modified by two adverbs, were not included in the analysis.

For the multifactorial analysis discussed in Section 6.8, a subset of all occurrences of modified infinitives was selected on the basis of the length of the corpus texts in which these infinitives occurred. Modified infinitives found in texts between 5,000 and 9,999 words were selected, in order to avoid a dataset with a large number of zeros, as the features extracted are not always highly frequent, and many short texts in the corpus contain no instances of these features. A random sample of a third of these instances was selected and manually classified to exclude all false positives (i.e. cases where the pre-verbal or post-verbal adverbs do not modify the infinitive, such as *to advance entirely different arguments*, as well as cases where there are two or more pre-verbal and post-verbal adverbs). Finally, the manually checked cases of split and non-split modified infinitives were further cleaned up in order to ensure that there were no two cases from the same text, because independent observations are required for the application of the binomial logistic regression analysis. The resulting dataset contained 4,926 cases of infinitives modified by a single *-ly* adverb.

Extraction of other prescriptively targeted features

The features below were used as prescriptivism-related predictors in the analysis of prescriptive influence on the use of split infinitives presented in Section 6.8. For all of these features, only the typically proscribed variant was identified, and the corpus texts were searched only for those variants, i.e. the frequencies of occurrence per 1,000 words for each text were calculated, but not the proportion of these variants in relation to their prescribed counterparts. For instance, only the frequency of singular *data* was established for each text in the corpus, and not its proportion in relation to plural *data*. The same goes for all other features.

Sentence-initial *and/but* – Each text in the corpus was searched for sentence-initial *and/but*, using regular expressions and punctuation to identify these occurrences. The

number of occurrences of these two words taken together was calculated for each text in the corpus, and was subsequently normalised per 1,000 words.

Data – Occurrences of singular *data* were identified on the basis of other elements in the sentence overtly marked for number. Thus, instances of *data* preceded by a singular determiner (e.g. *this data*) or followed by a singular verb (e.g. *the data shows*) were identified as singular *data*, while occurrences where *data* was preceded by a plural determiner or followed by a plural verb were identified as plural *data*. Instances where no overt number marking were identified (e.g. *The data showed an increase in frequency*), as well as instances where *data* was part of a noun phrase and the main verb of the clause is later in the clause, such as *The data obtained through this analysis shows that...*, were not included in the analysis (see the note at the end of this appendix). The occurrence of the singular *data* was operationalised as the normalised frequency per 1,000 words of singular *data* for each text in the corpus data.

Hopefully – All occurrences of *hopefully* were extracted from each text of the corpus with a regular expression, and no further classification was conducted, due to the difficulty of automatically disambiguating cases where *hopefully* was used as a sentence adverbial, from those where it was used as an adverb of manner. However, the frequency of this feature was still included, in view of the fact that we know from previous studies that *hopefully* as a manner adverb occurs very infrequently (Busse and Schröder 2010: 94), so we could assume that changes in the frequency of *hopefully* are indeed indicative of changes in the usage patterns of the sentence adverb *hopefully*. Thus, for each text in the corpus, the number of times *hopefully* was used was established, and the raw frequency for each text in the corpus was normalised per 1,000 words.

***Less* + PL** – All occurrences of *less* were extracted from the corpus files, using a regular expression. For each occurrence, only cases where *less* was immediately followed by a noun were used to calculate the frequency of occurrence of this feature. As in the case with *data*, cases where *less* referred to a plural noun that did not immediately follow *less* were not included in the analysis (see note at the end of this appendix). The raw frequencies per text were normalised per 1,000 words.

***None* + PL** – All occurrences of the word *none* were extracted from the corpus data, using a regular expression. The resulting concordances were then analysed using Python. Cases in which a plural verb was identified at any of the four positions following *none* were counted as *none* + PL. This means that both cases such as *none is* and cases such as *none of these is* were taken into account.

Passives – Passive constructions were identified by extracting all occurrences of the

verb *to be* followed by a past participle. For each text in the corpus, the number of such occurrences was counted, and normalised per 1,000 words.

Shall – All occurrences of *shall*, regardless of person, were extracted from each text of the corpus, using a regular expression. The total number of occurrences of *shall* in each text was normalised per 1,000 words.

These kind of/sort of – All occurrences of *kind* and *sort* were extracted from the corpus files, followed by counting the number of occurrences of the phrases: *these/those kind of* and *these/those sort of*. The total number of occurrences of these phrases was counted for each text in the corpus, and the frequency was subsequently normalised per 1,000 words.

Try and – All occurrences of *try* were extracted, using a regular expression, and cases where *try* is followed by *and* were counted for each text in the corpus. The total number of occurrences of *try and* in each text was normalised per 1,000 words.

Whom – All occurrences of *whom* were extracted from each text of the corpus, using a regular expression. The total number of occurrences of *whom* in each text was normalised per 1,000 words.

Note – In many cases described above, certain occurrences of the features were not taken into account. This was partly due to the difficulty of automatically identifying these cases. The other reason for this exclusion is that I focus on cases where a prescription is very conspicuously violated, while making sure that these cases are all quite similar. In other words, if we were to include cases where *less* refers to a plural noun that does not immediately follow *less*, the additional constraint here would be the distance between *less* and the noun, which may be expected to influence the likelihood of using *less* with a plural noun. Given that such details are hard to include in this kind of analysis, these cases were excluded. Thus, the figures for the occurrences are on the one hand conservative – in that there may be more occurrences of the proscribed variants than are included in the analysis – while on the other hand the cases which are included are very straightforward and conspicuous violations of a prescription: their occurrence in a particular text may reasonably be interpreted as an indication of weaker prescriptive influence in the text.

Raw data for the analyses presented in Chapter 6

In Section 6.2 I discussed the register effects on the increasing frequency of use of *ain't* and, on the basis of the data, argued that these are not indicative of actual changes in use, but are the by-product of the make-up of the corpus. The table below contains the raw figures on the basis of which this conclusion was reached.

Decade	**% drama texts**	**% drama texts with *ain't***
1810	16.28	0.00
1820	10.00	0.00
1830	7.18	0.00
1840	3.63	0.00
1850	0.00	0.00
1860	4.74	4.46
1870	0.47	0.97
1880	2.62	1.50
1890	0.38	0.00
1900	20.52	17.69
1910	33.90	27.01
1920	33.10	28.48
1930	7.32	8.05
1940	14.29	15.20
1950	8.92	11.38
1960	16.34	19.48
1970	18.75	20.95
1980	24.12	23.44
1990	4.79	11.52
2000	1.38	7.26

Table C.1: Percentage of drama texts in the fiction section of COHA per decade and percentage of drama texts which contain *ain't*

Appendix D: Annotations in usage guide entries

Attitudes to usage in usage guides

The tables below contain the annotations of usage guides entries for each feature, and the classification of attitude expressions into POSITIVE and NEGATIVE. These annotations formed the basis for the discussion of the attitudes to usage in usage guides in Section 5.4. The tables are given for each feature in alphabetical order. The phrases included in the tables are expressions of attitudes annotated in the usage guide entries. Following these tables, this appendix also contains the annotations for references to dimensions of usage identified in the entries, for each of the features analysed. In the tables, semi-colons separate different annotations, and, where necessary, minor changes have been made to the original text to allow for easier understanding out of context, as well as for consistency. These modifications were made only in the formatting of the text, in order to make the tables uniform. The tables contain only the years for each of the entry, as the goal is to represent how the attitudes and references to usage have changed over time. Adding the authors' names was avoided, as for some years the tables contain annotations from multiple entries and multiple usage guides. However, the corpus of entries annotated in 'brat' is available for consultation upon request.

Expressions of attitudes to usage in entries on *ain't*

Year	POSITIVE expressions
1950	could be an economical single form for *am not*, *is not*, *are not*, *has not*, *have not*

1957	*ain't* for *am not* is a natural contraction; the language needs an expression of this sort; supplies a real want
1963	appropriate; [*ain't*] would be a useful addition to informal English
1977	appropriate
1978	a deliberate attempt to suggest informality; down-to-earth common sense
1980	accepted in speech; on its way to full acceptance
1988	many English speakers go beyond defending *ain't*; handy contraction; natural; proper
1989	a few hardy souls approve the locution; approved by some; at times you will probably find *ain't* a very useful; desired by others; grammatically sound; in widespread use but usually in particular circumscribed ways that tend to remove the stigma from its use; logical; what is the matter with *ain't I?* for *am not I?*? Nothing whatever, save that a number of minor grammarians object to it
1993	acceptable; accepted; it is a word
2002	may be suitable
2003	to convey a down-to-earth quality
2005	*ain't* continues to appear in the speech of ordinary folks, leads a vibrant life in song lyrics, should be a contraction like any other
2006	little risk of censure
2014	a crisp and euphonious substitute for the strident and bisyllabic *isn't*, *hasn't*, and *doesn't*; today the word is going strong

Year	**NEGATIVE expressions**
1847	vulgarism of discourse
1901	inelegant; it will be a blessing to the English speaking people when the descendant shall sleep with his father; misleading
1907	always inelegant; atrocious
1910	always incorrect; vulgarism
1911	can not be called a contraction; vulgarism
1916	no defense possible for the vulgar use of *ain't* for *hasn't* and *haven't*; shows no signs of coming into good use; universally condemned; usually the construction can be avoided; we must get along as best we can without it
1920	inelegant; ought not to need criticism; the safe rule respecting contractions is never to use them; ungrammatical; vulgar; vulgarism
1927	careless; wrong
1934	illiterate expressions; never to be used; not in good use; vulgarism; vulgarisms
1937	vulgar
1947	error; I blush to record it; illiterate ($\times 2$)
1948	illiterate; too vulgar

1949	vulgarism; you should avoid it
1950	vulgate (×2)
1955	a vulgate contraction; never good English; should always be avoided; wrong
1957	illiterate; is an uneducated blunder; not considered standard; serves no useful purpose; unacceptable forms; unrecognised
1975	denounced; illiterate; illiterate and ungrammatical; inelegant; regarded as substandard; stigmatised; uneducated blunder; ungrammatical
1977	chiefly to record uneducated speech; leaves the writer open to the risk of having his or her intention misunderstood
1980	best avoided; illiterate; non-standard; not acceptable; the hallmark of the uneducated
1987	has not been accepted; illiterate; is cautioned against
1988	a mark of illiteracy
1989	absolutely vulgar; bugbear; incorrect; inelegant; its present disesteem; much vilified word; stigma; stigmatised word in general use; tends to mark the speaker and writer as socially or educationally inferior; the widely disparaged status; ungrammatical; utterly intolerable; vulgar; vulgarism
1993	firm rejection of *ain't*; not accepted; shibboleth; substandard; ungrammatical; vulgar
1998	a shibboleth of poor usage; it never will be OK; it's still misbehavin'; nonword; not OK
1999	affectation; stigma attached to it; social disapproval is so strong; controversial words in current English; regarded as the clearest single token of illiteracy
2002	appearance of ignorance; beyond rehabilitation; illiterate
2003	condemned; the classic 'mistake' in English; uneducated
2005	a mark of ignorance; a vulgarism; inelegant; low-class
2006	bugbear
2014	*ain't* is frowned upon

Table D.1: Expressions of attitudes to usage in entries on *ain't*

Expressions of attitudes to usage in entries on *like*

Year	NEGATIVE expressions
1975	a throwaway word; filler; ignorance; misuse; nervousness; overuse
1991	a vague qualifier; a verbal tic; apologetic overtones it gives sentences; indecisive; infests every sentence; poor
1998	it shows arrested development; juvenile colloquialism; space-filler; verbal tic; vogue word
2000	faddish; [shows] limited grasp of [...] language; maltreatment; verbal crutch

2002	filler word; flibbertigibbet; [*like* has] no more meaning than a belch
2003	habit; irritating; meaningless verbal hiccup
2004	a big indicator of unpolished; informal speaking; habit

Table D.2: Expressions of attitudes to usage in entries on discourse particle *like*

Expressions of attitudes to usage in entries on *literally*

Year	POSITIVE expressions
1989	it is neither; neither a misuse nor a mistake
1999	a little linguistic reflection will reveal a logical rigour behind a much derided use
2006	adds a hyperbolic edge to clichés; invites readers to savour the aptness of the writer's terms of reference; lends impact to quantitative statements

Year	NEGATIVE expressions
1910	intolerable
1918	incorrectly used
1927	incorrectly
1947	colloquialism, slovenly
1957	*literally* is used to mean the exact opposite of what it properly means; such false coin makes honest traffic in words impossible; the word should be avoided; we ought to be at pains to repudiate; [used] with no regard whatever to any meaning of *literally*
1977	misuse
1978	literal-minded readers find such locutions absurd
1980	a habit of heedless writers; do not recognize it; excess baggage; authorities criticise it; misuses; the consensus, however, is heavily against this sense; the sentence is more forceful without it; unnecessary emphasis
1984	a kind of disclaimer; if you don't wish to be taken literally, don't use *literally*; the result [of using *literally*] is generally painful
1988	erroneous; exasperating; loose; superfluous
1989	improperly; mistake; misuse
1991	has no meaning at all beyond a vague and unnecessary intensification; in careless writing and speech it often has the opposite meaning; misuse; misuses; often *literally* actually weakens an expression
1993	almost always overkill; bad intensifier
1994	disclaimer; enhancer; misguidedly; overused; [using *literally*] would raise some rather unpleasant images
1998	distorted beyond recognition; often confused; slipshod extension

1999	much derided use
2003	has been so overused; in danger of losing its literal meaning; should not be used as a synonym for *actually* or *really*; vague
2004	bold bluff of an intensifier; incorrect
2005	incoherence
2006	not acceptable
2008	avoid using *literally* to add extra emphasis; there are a lot of people whose blood pressure literally rises as they imagine putting lit firecrackers in your ears to make your sentence correct
2010	I think the use of *literal* as a general intensifier has become a distraction, something that tears me away from the message and makes me doubt the messenger; word confusion
2014	can evoke ludicrous imagery; it drives careful readers crazy; it screams, "I don't think about what my words mean"; problematic; superfluous

Table D.3: Expressions of attitudes to usage in entries on *literally*

Expressions of attitudes to usage in entries on negative concord

Year	POSITIVE expressions
1917	good English; natural; natural; natural to human language; springs from the desire for emphasis
1942	such a double negative is not a backsliding from the idiom of more formal English
1947	psychologically defensible
1957	normal way of strengthening a negative
1978	not a backsliding from the current idiom of standard English
1980	[two negatives] ordinarily reinforce each other and this is clearly felt by the reader
1989	it does have its uses; normal; you certainly don't need to eradicate it
1993	simply powerful
2005	alive and well; remains an effective construction in writing dialogue or striking a folksy note
2006	helps to underscore the force and/or defiance of the utterance

Year	NEGATIVE expressions
1856	mistake
1872	errors; should not be used
1884	incorrect
1895	inelegant

1895	should be avoided
1910	incorrect
1917	long banished from polite society and from literature
1920	do not say, used for ‘any’ by the illiterate
1927	wrong (×2)
1934	wrong (×2)
1938	*nothing* should not follow a negative expression in place of *anything*; wrong
1942	vulgate way
1947	contrary to the present idiom of the educated
1949	error; ignorant; of the lowest level
1955	should be avoided, vulgate idioms
1957	a shocking vulgarism; no one who values public opinion can afford to say; put a man beyond the pale
1980	are to be avoided; avoided by all except the unlettered; conspicuous; they place on the reader the burden of sorting out the meaning
1981	avoid; poor; poor English
1984	you shouldn’t say
1988	let people who write street graffiti, like this one in Kingston, Jamaica, believe that two negatives make a forcefully negative statement: “The poor can’t take no more.”; non-standard; unsuitable as prestige constructions
1989	not a prestige form; rustic; uneducated; you are not likely to impress
1990	incorrect; redundant
1991	an immediate indication that the speaker’s or writer’s diction is substandard; errors; wrong
1992	do not use more than one negative within the same clause; incorrect (×2); no exceptions about this taboo; taboo
1993	inappropriate; incorrect; mark speakers of vulgar English; shibboleths
1998	condemn the phrase; not standard English; stay away from the most flagrant examples
1999	self evidently wrong
2000	causing hundreds of English teachers to grimace in pain; error; illiterate; improper; mistake; ungrammatical; vulgar
2005	incorrect; it is not acceptable to say; violates the double-negative rule
2006	illogical; incurs more censure; target of common criticism; very conspicuous

Table D.4: Expressions of attitudes to usage in entries on negative concord

Expressions of attitudes to usage in entries on object *I*

Year	POSITIVE expressions
1947	a sense construction
1957	cannot be classed as a mistaken attempt to speak elegant English; has such a long and honourable history; used by so many great writers
1989	treated as a polite fixed unit; you are probably safe in retaining *between you and I* in your casual speech

Year	NEGATIVE expressions
1856	as that of the vulgarian who says "Him and me are going to the play," and with less excuse; erroneous; errors of the ill-bred and those of the well-bred man; fault; faults; genteel error; gross violation of a rule which ought to be familiar to everybody; heinous; manifest improprieties; not very generous; woeful confusion
1868	error; mistake
1884	carelessly used
1911	error; gross errors; insidious errors; mistakes
1916	grossly incorrect
1920	careless; error (×2); incorrect; often confused
1927	certainly not the best English; ungrammatical; wrong
1937	fault
1947	indefensible grammatically; misused
1949	always wrong
1957	a piece of false grammar; illiterate; lapses; not sanctioned even by colloquial usage
1975	even from otherwise literate speakers
1977	error (×2); genteelism; misguided
1978	half-educated
1980	blunder; deviations; erroneously; questionable; wrong (×3)
1981	poor
1984	a grammatical error of unsurpassable grossness; always wrong; gaffe
1988	error; irritating
1989	blunder; illiterate; wrong
1990	incorrect
1991	wrong
1992	incorrect; major relapse; mistake
1993	mistake; never say or write; shibboleth
1994	incorrect; misguided; most common error; ungrammatical
1996	error (×2)); grammatical errors; mistake

1998	debilitated grammar; gross linguistic gaffes; mistake; problem
1999	hypercorrection
2000	a form of overrefinement; error (×2)); mistake
2001	never acceptable; problem
2002	no exceptions
2003	misuse
2004	incorrect
2005	blunder; sign of ignorance
2006	shibboleth
2008	wrong
2010	error (×2)
2014	avoid *between you and I*; despised; error (×3); excruciating grammatical blunder; it makes many readers bristle

Table D.5: Expressions of attitudes to usage in entries on object *I*

Expressions of attitudes to usage in entries on subject *me*

Year	NEGATIVE expressions
1868	gross error; mistake
1916	vulgar
1920	vulgar; error; one should not say
1927	wrong
1977	incorrect
1984	lapses
1988	mistake
1989	disputed; problematical; likely to be unfavourably noticed in the speech and writing of adults; characteristic of less educated English
1993	never say or write
1996	error (×2); incorrectly using object pronouns in subject positions; mistake
2003	not elegant; not correct

Table D.6: Expressions of attitudes to usage in entries on subject *me*

Expressions of attitudes to usage in entries on the split infinitive

Year	POSITIVE expressions
1898	it sometimes helps the writer over the difficulty; not a violation of any rule of grammar
1911	seems to be growing in favour
1917	admirable; cannot so easily be proved to be a corruption; contributes decidedly to clearness; has a right to a trial in the language; has the distinct advantage of bringing an adverb into an emphatic position; natural; neither an innovation nor a vulgarism; not really an error in grammar; the use of a split infinitive does not necessarily put us among illiterates, ignoramuses, and violators of English undefiled; very clear; very convenient
1927	clearness and emphasis
1938	the meaning is more clearly expressed by inserting the adverb between the preposition and the infinitive
1949	a single adverb may not do violence to the statement of an idea
1957	the rule against splitting an infinitive contradicts the principles of English grammar and the practice of our best writers; natural
1963	good writers, in fact, prefer using split infinitives in sentences where not doing so would result in ambiguity or awkwardness; there is no point in revising a sentence just to avoid splitting an infinitive
1966	expressive; has its place in good composition; [the split infinitive] should be used; sometimes splitting is called for
1975	feel free to split the infinitive; perfectly good English
1978	sentences can be improved by splitting the infinitives; some [infinitives] should be [split]; smoothly and clearly
1980	infinitives may be split when splitting makes the sentence read more smoothly
1981	keeps your meaning clear; sometimes necessary; clear
1984	if it is the clearest and most natural construction, use it boldly; the split infinitive is [not] a grammatical error
1988	a construction that is fully established in the language
1989	nothing grammatically wrong; the objection to the split infinitive has never had a rational basis; you can split [infinitives] when you need to
1991	better [split]
1992	adds emphasis; excusable with good reason
1993	an adverb may split an infinitive if required by natural position
1994	a sentence that would read more smoothly with the infinitive split
1997	in this case, splitting the infinitive is the most accurate way of expressing what happened; it is better to split an infinitive

1998	[the English language gives us] the inestimable advantage of being able to put adverbs where they will be most effective; the rule against split infinitives contradicts the principles of English grammar and the practice of our best writers; correct and acceptable English; in full accord with the spirit of modern English; no harm in separating them; perfectly proper; the universal adoption of this usage is as certain as anything in the future well can be; there is no point in rearranging a sentence just to avoid splitting an infinitive unless it is an awkward one
1999	it is acceptable; it is usually better (and sometimes necessary) to place [the adverb] between *to* and the verb; neither a major error, nor a grammatical blunder
2000	an improvement of English expression; you need not avoid splitting an infinitive if you have good reason to split it; can also be helpful; expresses your meaning more clearly; sounds more natural
2001	often the most natural position to place an adverb
2003	not strictly speaking an error; often more expressive and graceful
2004	more often than not, in my opinion, infinitives are better split; natural
2005	has a strong rhythm that reinforces the meaning; hard to see what exactly is wrong; meaning is clear
2008	there's no reason to go out of your way to avoid it; don't let anyone tell you that it's forbidden; it is OK to split infinitives; it's fine to split infinitives
2010	do not be afraid to split infinitives

Year	**NEGATIVE expressions**
1856	an adverb should not be placed immediately after *to*; the rule is violated
1867	another of the blunders; preposterous
1868	mistake
1884	do not put an adverb between *to* and its infinitive; never separate *to* from the infinitive with which it belongs
1895	errors; fault
1901	intolerable; no author who uses English with propriety and regard for established correct usage, ever separates the particle from the verbal word
1910	condemnation of the split infinitive is now pretty general
1911	awkward; mistake
1916	has long been frowned on
1917	an offence against philology; barbarous practice; aesthetically ugly; intolerably awkward; objected to only when it produces clumsiness; vulgarism ($\times 2$)
1920	blunder; condemned; finds no place in such expressions as...; reprehensible; should not [be used]; [infinitives are] strictly inseparable
1927	contrary to the history of the construction; suspended syntax; wrong
1934	wrong

1937	indefensible
1942	awkward
1949	error
1955	avoid split infinitives that are obviously awkward
1957	a grammatical mistake; deafening; lapse; the heinousness of this offense
1966	fault
1975	pedantic bogey
1978	awkward; should not be split
1980	believe they will not go to heaven if they split the infinitive
1981	infinitive should be kept intact
1984	avoid the split infinitive wherever possible; grammatical error
1988	the notion that only the illiterate and ill-bred split an infinitive
1989	avoid split infinitives; anyone who aspires to be a bad writer should split as many infinitives as possible; condemned (×2); would produce bad writing
1990	awkward constructions; in general, you should avoid splitting infinitives
1991	implication of ignorance; sloppiness
1992	questionable; unnecessary; unnecessary split
1993	eliminates all possibility of ambiguity; sometimes can cause very clumsy sentences
1994	incorrect
1997	interrupt the flow; may be taken as ignorance; separating its parts can weaken it
1998	displays carelessness; generally to be avoided
1999	error; has sufficient weight of opinion against it
2000	can be awkward; somewhat discomforting
2003	better to avoid; people are offended by split infinitives
2006	can make awkward reading; inelegant; ungrammatical

Table D.7: Expressions of attitudes to usage in entries on the split infinitive

Dimensions of usage in entries on *ain't*

Year	References to FREQUENCY
1901	frequently heard
1911	however much it may be employed
1927	it is true nevertheless that many educated persons permit themselves this habit; used as a contracted form of *am not*, *is not*, and *are not*
1948	in spite of its use

1950	fairly common among educated speakers; the commonest and most easily identifiable vulgate words
1957	it is heard; insist on using; used for *isn't*
1963	used in non-standard English
1975	used
1978	regularly use *ain't*; it is never used in formal writing
1980	sometimes boldly used; relatively rare
1987	occasionally used
1988	even more restricted; occasionally, however, those who are certain of their status as cultivated speakers of standard English dare to use it
1989	word in general use; in widespread use; common among the less educated and among children; most common in fiction; common in fiction; this use pops up unsurprisingly in advertising and in political slogans; often heard; *ain't* occurs frequently in inverted expressions, such as questions; use of *ain't* that many handbooks agree is common is facetious or jocular or humorous use
1998	used; *ain't* is used
2002	common; appears
2003	still commonly used; everyone uses it occasionally
2005	*ain't* continues to appear; leads a vibrant life
2014	not that *ain't* is used as a standard contraction

Year	**References to MODE**
1901	speech
1920	never to use them in public speech
1937	not yet been promoted to writing; speech
1948	in written or spoken
1950	conversation
1955	rarely needed in writing
1975	orally; in speech; writing; *ain't* in writing
1978	in speech; writing (×2)
1980	in writing (×2); in speech
1987	speech and writing
1988	speech or writing
1998	orally
2002	speech; in writing; in speech; used orally; may be suitable for writing
2005	in the speech; in speech; writing
2006	distinction between spoken and written usage; spoken as well as written; appearances in print; used orally

Year	References to REGISTER
1847	in the pulpit or at the bar
1901	colloquial speech
1920	never to use them in public speech
1927	on the low colloquial level; cultivated colloquial; literary use; low colloquial
1934	not in good use in either colloquial or formal
1950	colloquial; in actual conversation
1955	in dialogue
1963	informal English; colloquial; in formal English
1975	in fictional dialogue
1977	to deliberately informal usage; a device for providing humor, shock, or other special effect
1978	formal writing; ordinary expository prose; in general writing
1980	quoted speech; jocular
1987	colloquial; informally
1988	informal; mainly limited to dialogue and humorous contexts; formal; in clichés; to lighten the tone of their remarks
1989	in ordinary speaking and writing; usually in particular circumscribed ways; educated persons whose regular vocabulary still includes *ain't* use the term in talking to relatives and to peers with whom they are both friendly and on a first-name basis; the use of *ain't* in a letter marks a close and warm relationship; spoken, as in an interview or even in a talk; written, as in an article; to emphasize their informality; most common in fiction; can also be found in other forms of writing; can also be used for characterizing purposes; common in fiction; the characterizing *ain't* can be used in reportage; in advertising; in political slogans; in otherwise rather straightforward prose for purposes of contrast; *ain't* occurs frequently in inverted expressions, such as questions; in popular music; catch phrases and variations on them make up a goodly portion of the word's use, both orally and in writing; when the tag is necessary, *ain't* will probably occur in it in some people's speech; use of *ain't* that many handbooks agree is common is facetious or jocular or humorous use; many educated people, when they use *ain't*, try to use it in such a way as to show that it is not part of their serious day-to-day vocabulary; accomplished by the use of the familiar fixed phrases; from speech
1993	when used unconsciously or unintentionally; jocular uses; using *ain't* in circumstances that do not suggest deliberate choice
1998	show that you have the common touch; to be tongue-in-cheek
1999	in catchphrases; as an affectation
2002	in the most casual of colloquial speech; spoken slang; in many songs
2003	part of a joking phrase

2005	in song lyrics; *ain't* has no substitute in fixed expressions; informal writing
2006	informal conversation; signal of congruent informality; associated with casual and dialectal speech; embedded in quoted speech; quasi-proverbial sayings; appearing freely in utterances quoted in newspapers; in proverbial sayings; gets into print in reference to songs
2014	but it does have some widely established places; in the lyrics of popular songs; even in relatively formal settings; to emphasize that some fact is so obvious as to be beyond further debate

Year	References to SPEAKERS
1847	some persons of education and character
1927	students of English; critical speakers; many educated persons
1950	educated speakers; educated people
1955	some modern users of English
1957	a few bold spirits; most people
1963	some authorities; most users of standard English
1975	America's schoolteachers; by many cultivated speakers
1977	the writer
1978	millions of Americans
1980	by those who are sure of themselves; most readers; cultivated speakers prefer *am I not*
1987	by educated persons
1988	some authorities; many English-speakers; educated speakers; those who are certain of their status as cultivated speakers of standard English
1989	among the less educated; among children; educated persons whose regular vocabulary still includes *ain't* use the term in talking to relatives and to peers with whom they are both friendly and on a first-name basis; many educated people, when they use *ain't*, try to use it in such a way as to show that it is not part of their serious day-to-day vocabulary
1993	by some authorities; Americans (×2)
1998	by cultivated speakers; for most people
2002	by many cultivated speakers
2003	everyone
2005	by upper-class speakers; the lower classes; of ordinary folks; educated and upperclass speakers; educated speakers
2006	American school teachers; between American speakers; writers; by many cultivated speakers

Year	References to VALUE
1927	they reprehend it as careless

1950	if the social objection could be relaxed; prejudice against it among educated people has been almost unanimous for the last century or so
1957	shamefaced reluctance with which these full forms are often brought out; as used for *isn't* is an uneducated blunder and serves no useful purpose; he (or still more she) fears will convict him of low
1963	strong social and educational pressure against *ain't*
1975	cultivated
1980	most readers are likely to consider it the hallmark of the uneducated
1988	to many people, *ain't* doesn't bear discussion: it is simply a mark of illiteracy; the risk of criticism or ridicule is too high a price to pay for using this handy contraction
1989	in ordinary speaking and writing it tends to mark the speaker and writer as socially or educationally inferior; often meant to mark the speaker as belonging to a lower class or being poorly educated or being black or being countrified; but it may also be a code word, used in a sly way to tip off the reader to the fact that the person being quoted is poor, illiterate, or black; most common public uses of *ain't* makes use of the word's ability to attract attention; it's not really an attempt at jocularity or humor, it's an attempt at distancing; the verbal equivalent of a wink or nudge intended to show that you are not so ill-bred as to really use *ain't*
1993	may brand you as a speaker of vulgar English
1998	(1) to be tongue-in-cheek; and (2) to flaunt their reverse snobbery; if you're tempted to use it to show that you have the common touch, make clear that you know better
2003	if you always use it instead of the more proper contractions you're sure to be branded as uneducated
2005	low-class; upper-class; a term used by the lower classes; has come to be regarded as a mark of ignorance.
2006	stigma attached to it in the U.S.; a signal of congruent informality

Year	References to VARIETY
1957	in the United States; standard; modern English
1963	non-standard English; standard English
1975	in most parts of the United States; substandard
1980	British; American; non-standard; non-standard
1987	dialectal; standard speech and writing
1988	dialectal; peculiar to a certain region, community, social group, or the like; non-standard; the language variety of educated speakers
1989	non-standard
1993	Americans; standard English; substandard; standard use; standard American English; vulgar and some Common use; vulgar English
1998	in most parts of the country

1999	current English; Cockney speech; unlikely that *ain't* will be admitted to standard English in the foreseeable future
2002	in most parts of the U.S.
2005	in English
2006	American English; British; dialectal; standard English; in most parts of the U.S.; stigma attached to it in the U.S.; American speakers; more significantly embedded in American English than in British
2014	regional and lower-class English; as a standard contraction

Table D.8: References to dimensions of usage in entries on *ain't*

Dimensions of usage in entries on *like*

Year	References to FREQUENCY
1975	used constantly in the speech of many persons
1991	sometimes used
1998	ubiquitous
2003	common in speech; this habit has spread throughout American society
2005	used frequently

Year	References to MODE
1975	in the speech of many persons
2003	speech
2004	speaking
2005	speech; spoken
2005	writing

Year	References to REGISTER
1993	casual
1998	colloquialism
2004	informal
2005	informal ($\times 2$); limited chiefly to dialogue

Year	References to SPEAKERS
1975	especially young people
1975	many persons
1988	some speakers
1991	adolescents
1993	adolescent

1998	in teenagers; in adults
2002	teenage
2003	hipsters; people of all ages; young people
2005	younger people

Year	References to VALUE
1988	with the approval of almost no one outside their own group
1991	using it amounts to an admission by the speaker that his or her expression is poor
1998	in adults, it shows arrested development
2000	faddish
2003	to be reacted to as a grown-up, avoid this pattern
2004	because a few of these likes in a sentence send all the wrong signals

Year	References to VARIETY
1993	substandard
1998	California

Table D.9: References to dimensions of usage in entries on discourse particle *like*

Dimensions of usage in entries on *literally*

Year	References to FREQUENCY
1918	often incorrectly used
1927	sometimes used
1947	when used, as it often is
1957	we have come to such a pass with this emphasizer; we do not hesitate to insert the very word
1966	*literally* continues to be seen as a mere intensive that means practically, almost, all but
1975	too often used to intensify a statement which is actually a figure of speech
1977	common misuse
1978	it is so often used to support metaphors that its literal meaning may be reversed
1980	seldom is the word employed in its exact sense
1984	all too often used as a kind of disclaimer
1989	seldom is the word employed in its exact sense, which is to the letter, precisely as stated; often used hyperbolically; often improperly used; furthermore, these uses as monitored by our readers outnumber the hyperbolic use by a substantial margin

1991	careless writing and speech it often has the opposite meaning
1994	but many writers use it
2006	yet sensational examples like this don't outnumber those of a more measured kind in the BNC; it has also been used to underscore figures of speech or turns of phrase which could never be taken at face value
2010	how often writers and speakers confuse the antonyms literal and figurative
2014	common

Year	**References to MODE**
1927	in colloquial speech
1991	writing and speech
2006	both written and spoken
2008	writing

Year	**References to REGISTER**
1927	in colloquial speech
2006	not in the most formal prose; but in interactive discourse; media discourse

Year	**References to SPEAKERS**
1966	writers; rhetoricians; writers
1977	writers of such sentences
1978	literal-minded readers
1980	heedless writers
1984	writers
1988	some people use both words, but particularly the adverb, merely as intensives
1994	many writers use it misguidedly
2006	skilled writers; writers/speakers
2010	writers and speakers

Year	**References to VARIETY**
2006	in standard English

Table D.10: References to dimensions of usage in entries on *literally*

Dimensions of usage in entries on negative concord

Year	References to FREQUENCY
1856	a very common mistake
1872	not uncommon
1917	it sometimes crops out at inopportune times in the speech of self-taught men and women
1920	often used for 'any' by the illiterate
1937	dropped out of use
1942	not used in formal and informal English; probably not so common in vulgate English as comic writers suggest; two negatives are very often used to make an emphatic negative; survives in vulgate usage
1963	often used; not used by educated people
1978	very often used in non-standard English; two negative words in the same construction are not used in standard English; no longer used
1980	the more sophisticated are often unwittingly guilty of it
1989	is indeed common; the range of use of the double negative has shrunk considerably in the past 400 years; but it has not disappeared; it still occurs in the casual speech and writing of more sophisticated and better educated people
1993	many speakers still use these constructions today
1999	can easily be found
2006	it has a long history of use; used in many non-standard dialects; survives in casual conversation

Year	References to MODE
1917	speech
1917	speech
1942	lost to written English
1955	writers
1989	speech of the unlettered
1989	speech
1989	speech
1989	writing
1989	speech
1993	spoken and written standard English
2003	speech
2005	spoken English
2005	speech
2006	conversation

2006	strongly associated with speech

Year	References to REGISTER
1917	popular speech
1927	in your conversation
1942	in many speech situations; formal; informal; formal; informal
1963	except in joking mood; in formal and informal English
1989	seem to have gone out of literary favour; restricted to familiar use; conversation; letters; the speech of similar characters in fiction; discursive prose; when talking to your family and friends; casual speech and writing
1993	except in jocular use
2001	standard English; vernacular English
2003	informal
2005	writing dialogue; folksy note; spoken English
2006	survives in casual conversation

Year	References to SPEAKERS
1895	writers
1917	teachers; self-taught men and women
1920	the illiterate
1942	educated people; comic writers
1947	contrary to the present idiom of the educated
1955	writers
1963	educated people
1978	educated people
1980	the unlettered; the more sophisticated
1984	most people know that you shouldn't say; some writers
1988	schoolteachers; people who write street graffiti, like this one in Kingston, Jamaica
1989	grammarians; among the least educated; unlettered; less educated people; more sophisticated and better educated people; the boss; the teacher; the job interviewer; many other grammarians
1991	the speaker's or writer's
1993	eighteenth-century grammarians; many speakers
1998	traditionalists; descriptive linguists
2000	an investigative correspondent; English teachers
2001	vernacular speakers
2003	people

2005	grammarians; readers (×2)
2006	sociolinguists; contemporary grammarians; writers

Year	References to VALUE
1942	not now in fashion among educated people
1957	no one who values public opinion can afford to say; when used with a negative verb put a man beyond the pale
1963	they are out of fashion now in the standard language
1978	not now in fashion among educated people
1989	associated with the speech of the unlettered; you are not likely to impress the boss, the teacher, or the job interviewer
1991	it is nevertheless an immediate indication that the speaker's or writer's diction is substandard
1993	mark speakers of vulgar English
1999	poorly educated
2000	double negative is vulgar and improper
2006	socially stigmatized; they incur more censure than the others through their social connotations – the fact that they're used in many non-standard dialects

Year	References to VARIETY
1937	Modern English
1942	survives in vulgate usage; vulgate English; vulgate way; contrary to the present idiom of the educated
1955	non-standard; vulgate idioms
1957	in all Teutonic languages
1963	non-standard English; non-standard; the standard language; standard
1978	standard English; non-standard English; in standard English
1991	substandard
1993	standard English; not out of the language, but out of standard use
1998	standard English
1999	in all varieties of English used throughout the world; East London English; Black English spoken in the U.S.
2005	dialect or non-standard speech; standard English; standard usage
2006	American and British English; non-standard dialects

Table D.11: References to dimensions of usage in entries on negative concord

Dimensions of usage in entries on object *I*

Year	References to FREQUENCY
1856	frequently heard; how often do we hear even well-educated people say
1868	common
1884	often carelessly used
1911	we sometimes meet with gross errors of this kind; common
1920	common
1927	occurs occasionally; sometimes used; current
1937	common faults
1942	frequently heard
1947	often used
1957	is often said
1975	is often heard
1977	a very common error
1978	frequently heard and has a long history in written English
1980	here are many (bad) examples of the expression; often erroneously used
1984	very common
1988	instead we often hear the subjective forms; more often in speech than in print; why is it so common
1989	it occurred in the past and it occurs now; examples in print, especially recent ones, are hard to find
1991	is heard
1992	common
1994	most common error
1996	commonly occur; so common
1998	it is perennially surprising how many otherwise educated speakers commit them; ubiquitous
2001	common though this form may be in spoken English
2005	phrase occurs quite often in speech
2006	certainly used; unlikely to occur
2010	common
2014	commonly heard phrase

Year	References to MODE
1927	speech
1988	in speech than in print
1989	chiefly spoken; in print; mostly in speech; in print
1993	speech; writing

1996	speech
2001	in spoken English; in careful writing
2005	in speech
2005	writing
2006	in writing
2006	it's to be avoided in writing

Year	References to REGISTER
1911	we sometimes meet with gross errors of this kind in the writings of authors of repute; in conversation
1927	colloquial speech; colloquial; low colloquial
1957	not sanctioned even by colloquial usage
1975	in casual speech
1980	in literary classics; colloquial
1989	in your casual speech; in essays; works of a discursive nature; modern edited prose
2005	formal writing
2006	a formal document

Year	References to SPEAKERS
1856	well-educated people
1911	authors of repute
1947	often used by those who would never dream of saying *between he and I*
1988	even from people who probably know better; editors; proofreaders; editors; authors; schoolteachers; writers of books on grammar and usage; speakers and writers
1989	the ignorant or timid
1991	the well-educated
1993	standard English users
1994	misguided speakers and writers
1996	sophisticated people such as news broadcasters and educators
1998	educated speakers
2003	educated people; people
2006	for some people
2014	speakers; careful writers and speakers ($\times 2$); many speakers; writers

Year	References to VALUE
1977	sometimes a genteelism resorted to by those who think that *I* is somehow a finer or 'more correct' word than *me*
1977	attempt at refinement is misguided
1978	anyone who uses it now is apt to be thought only half-educated
1988	probably no construction is more irritating to those who don't use it than between you and I
1989	in rather more educated varieties of English; if you use it, someone is sure to notice and disparage your character, background, or education.
1994	who think I sounds more formal than me
1998	most people who make this mistake do so out of habit, without thinking, and not because they don't know the difference between I and me
2000	a form of overrefinement
2003	the misuse of 'I' and 'myself' for 'me' is caused by nervousness about 'me'
2005	widely regarded as a sign of ignorance

Year	References to VARIETY
1927	in dialectal speech
1957	not standard English
1989	occurs in rather more educated varieties of English; early modern English
1993	standard English
2003	standard English
2014	contemporary English

Table D.12: References to dimensions of usage in entries on object *I*

Dimensions of usage in entries on subject *me*

Year	References to FREQUENCY
1868	we sometimes hear
1984	not as uncommon as we might hope them to be
1988	sometimes
1989	in actual practice we also find me and someone and someone and me
1996	quite rare
2005	widespread tendency
2006	*me* is sometimes used for *I* when it's the first coordinate of the subject

Year	References to MODE
1988	in print
1989	speech forms
1989	speech and writing
2003	speech
2005	writing
2005	in speech
2006	speech
2006	in writing

Year	References to REGISTER
1989	when used facetiously
2005	formal writing
2006	in conversation; informal; easy-going conversation; casual speech

Year	References to SPEAKERS
1988	editors
1988	proofreaders
1989	children; adults
1991	child's speech; parents; teachers
2003	educated people
2006	some speakers

Year	References to VALUE
1989	associated with the speech of children

Year	References to VARIETY
1989	less educated English
1989	non-mainstream varieties of English
1991	standard English
2006	world Englishes

Table D.13: References to dimensions of usage in entries on subject *me*

Dimensions of usage in entries on the split infinitive

Year	References to FREQUENCY
1867	it is even more common
1869	the liberty is frequently taken
1895	the most common fault
1917	is used by a great many careful writers; the split infinitive is very rare as compared with the other; more and more common among good writers; widely used in colloquial and literary English; used without hesitation by many writers of repute; it is becoming more and more common among good writers; it crops up frequently in scientific journals, daily papers, reports of mercantile societies, and such places; it is used pretty frequently by well educated men not especially careful of their English; the split infinitive is rare; very rare in standard literature; it is spreading in the daily and weekly papers, and in the colloquial English of the intelligent classes; while a good many reputable authors use the split infinitive, they use it rarely
1920	despite the hundreds of uses of this method of expression
1927	in practice; some good writers permit themselves the liberty of placing an adverbial modifier between to and the infinitive
1957	the split infinitive first came into general use
1966	extremely rare
1989	there has always been a question about how frequently the split infinitive construction occurs; only occasionally; the construction is common; frequent in Mark Twain, Thomas Hardy, and Rudyard Kipling as well as Browning; the frequency of the split infinitive is that it noticeably increased in the 19th century; the increase in split infinitives
1993	during the nineteenth and early twentieth centuries great numbers of split infinitives appeared in print; for the popularity of the split infinitive; split infinitives continue to appear often
1998	frequently; has steadily increased during the last hundred years, and goes on increasing still
1999	occasionally
2004	common
2005	people have been splitting infinitives since the 14th century; split infinitives all the time without giving it a thought

Year	References to MODE
1911	everyday speech
1966	spoken; in written work
1978	formal writing; writing
1989	spoken; on the printed page; in the speech of the less educated
1993	in print; in writing; speech

1998	bad writing
1999	best avoided in normal writing and speech

Year	References to REGISTER
1869	in print
1911	in most cases of everyday speech
1917	colloquial; literary English; in scientific journals; daily papers; reports of mercantile societies; in the daily and weekly papers; the colloquial English of the intelligent classes
1966	in good composition
1978	formal writing; in unquestionably reputable general writing; in general English; in formal
1984	a question of style
1989	unless it is in the slangy construction in which an expletive is infixed between the syllables of a word; to literary contexts
1993	edited English; planned, oratorical, and formal levels

Year	References to SPEAKERS
1898	good writers
1917	careful writers; good writers; writers of repute; good writers; well educated men not especially careful of their English; the intelligent classes; good many reputable authors
1957	readers
1963	good writers
1980	many writers
1988	English-speakers who are at home with the language; anyone who has ever spoken or written the language
1989	native speakers; the less educated; users of standard English; many authors
1991	fewer and fewer writers, and few grammarians; sophisticated users; less sophisticated users; writers
1994	many people
1998	writers of English
2001	a minority of people
2005	people have been splitting infinitives since the 14th century

Year	References to VALUE
1957	deplorable breach of etiquette
1966	the one fault that everybody has heard about and makes a great virtue of avoiding and reproving in others

1989	common in the speech of the less educated
1991	occasionally writers seem to go out of their way to put the modifier in an unnatural place, perhaps as a kind of showing off they want their readers to notice that they know enough not to split infinitives.
1998	bad writing

Year	**References to VARIETY**
1955	standard usage
1966	spoken English
1989	spoken English
1993	in standard speech
1998	both in England and America

Table D.14: References to dimensions of usage in entries on the split infinitive

Primary Sources

Allen, Edward Frank. 1955. *How to Write and Speak Effective English*. New York: Fawcett Publications, Inc. First published in 1938.

Allen, Robert, ed. 1999. *Pocket Fowler's Modern English Usage*. Oxford: Oxford University Press.

Anonymous. 1856a. *Five Hundred Mistakes of Daily Occurrence in Speaking, Pronouncing and Writing the English Language, Corrected*. New York: Daniel Burgess & Co.

Anonymous. 1856b. *Live and Learn: A Guide for All, Who Wish to Speak and Write Correctly*. New York: Garrett & Company.

Ayres, Alfred. 1911. *The Verbalist*. New York and London: D. Appleton and Company, revised edn. First published in 1881.

Bache, Richard. 1868. *Vulgarisms and Other Errors of Speech*. Philadelphia: Claxton, Remsen & Haffelfinger.

Ballard, Harlan H. 1884. *Handbook of Blunders Designed to Prevent 1,000 Common Blunders in Writing and Speaking*. Boston: Lothrop, Lee & Shepard Co.

Bardeen, C.W. 1883. *Verbal Pitfalls: A Manual of 1500 Words Commonly Misused*. Syracuse, N.Y.: C.W. Bardeen, Publisher.

Batko, Ann. 2004. *When Bad Grammar Happens to Good People: How to Avoid Common Errors in English*. Franklin Lakes, NJ: Career Press.

Beason, Larry and Mark Lester. 1996. *Now I get it: A Commonsense Guide to Grammar and Usage*. Boston and New York: Bedford/St. Martin's.

Bechtel, John H. 1901. *Slips of Speech: A Helpful Book for Everyone Who Aspires*

to Correct the Everyday Errors of Speaking and Writing. Philadelphia: The Penn Publishing Company. First published in 1895.

Bierce, Ambrose. 1909. *Write It Right: A Little Blacklist of Literary Faults*. New York: Walker Publishing Company, Inc.

Booher, Dianna. 1992. *Good Grief, Good Grammar*. New York: Ballantine. First published in 1988.

Brians, Paul. 2003. *Common Errors in English Usage*. Wilsonville, Oreg.: William, James & Co.

Bryson, Bill. 1984. *The Penguin Dictionary of Troublesome Words*. London: Guild Publishing.

Carter, Bonnie, Craig Skates, and Ted Buchholz. 1990. *The Rinehart Guide to Grammar and Usage*. Fort Worth: Holt, Rinehart and Winston, Inc., 2nd edn. First published in 1988.

Clark, Roy Peter. 2010. *The Glamour of Grammar: A Guide to the Magic and Mystery of Practical English*. New York: Little, Brown and Company.

Compton, Alfred G. 1898. *Some Common Errors of Speech*. New York and London: G.P. Putnam's Sons.

Copperud, Roy H. 1980. *American Usage and Style: The Consensus*. New York: Van Nostrand Reinhold, revised edn.

Corbin, Richard K. and Porter G. Perrin. 1963. *Guide to Modern English*. Chicago: Scott, Foresman, and Company. First published in 1955.

de Mello Vianna, Fernando, Anne D. Steinhardt, and Carole La Mond, eds. 1977. *The Written Word*. Boston: Houghton Mifflin Company.

De Vries, Mary. 1991. *The Complete Word Book*. Englewood Cliffs, N.J.: Prentice Hall, Simon & Schuster.

Ebbitt, Wilma and David Ebbitt. 1978. *Writer's Guide and Index to English*. Glenview, Cook County, Ill.: Scott, Foresman and Company, 6th edn. First published in 1939.

Evans, Bergen and Cornelia Evans. 1957. *A Dictionary of Contemporary American Usage*. New York: Random House.

Fernald, James C. 1907. *Better Say – A Book of Helpful Suggestions for the Correct Use of English Words and Phrases*. New York: Funk & Wagnalls Company.

Fitzgerald, Joseph. 1901. *Word and Phrase: True and False Use in English*. Chicago: C. McClurg & Co.

Fogarty, Mignon. 2008. *Grammar Girl's Quick and Dirty Tips For Better Writing*. New York: St. Martin's Griffins.

Follett, Wilson. 1966. *Modern American Usage: A Guide*. New York: Hill & Wang.

Garner, Bryan. 1998. *A Dictionary of Modern American Usage*. New York: Oxford.

Gilman, E. Ward, ed. 1989. *Webster's Dictionary of English Usage*. Springfield, Mass.: Merriam-Webster.

Gould, Edward. 1867. *Good English, or Popular Errors in Language*. New York: W.J. Widdleton.

Hadida, Sophie. 1927. *Pitfalls in English and How to Avoid Them*. New York and London: G.P. Putnam's Sons.

Hall, John Leslie. 1917. *English Usage: Studies in the History and Uses of English Words and Phrases*. Chicago: Scott, Foresman, and Company.

Horwill, Herbert W. 1935. *A Dictionary of Modern American Usage*. Oxford: Oxford University Press.

Hurd, Seth T. 1847. *A Grammatical Corrector; or, Vocabulary of the Common Errors of Speech*. Philadelphia: E.H. Butler & Co.

Johnson, Edward D. 1991. *The Handbook of Good English: A Comprehensive, Easy-to-use Guide to Modern Grammar, Punctuation, Usage, and Style*. New York: Pocket Books, revised edn. First published in 1982.

Krapp, George Philip. 1927. *A Comprehensive Guide to Good English*. New York: Rand McNally.

Lounsbury, Thomas R. 1908. *The Standard of Usage in English*. New York and London: Harper & Brothers Publishers.

Lovinger, Paul W. 2000. *The Penguin Dictionary of American English Usage and Style*. New York: Penguin.

Mager, Nathan H., Sylvia K. Mager, and John Domini. 1993. *Prentice Hall Encyclopedic Dictionary of English Usage*. Englewood Cliffs: Prentice Hall, 2nd edn.

Meredith, L.P. 1872. *Every-day Errors of Speech*. Philadelphia: J.B. Lippincott & Co.

Morris, William and Mary Morris. 1975. *Harper Dictionary of Contemporary Usage*. London: Harper & Row.

Nicholson, Margaret. 1957. *A Dictionary of American-English Usage: Based on Fowler's Modern English Usage*. New York: Oxford University Press.

O'Conner, Patricia. 1998. *Woe is I*. New York: Riverhead Books. First published in 1996.

Parrish, Thomas. 2002. *The Grouchy Grammarian: A How-Not-To Guide to the 47 Most Common Mistakes in English Made by Journalists, Broadcasters, and Others Who Should Know Better*. Hoboken, N.J.: Wiley.

Partridge, Eric. 1947. *Usage and Abusage in English*. London: Hamilton. First published in 1942.

Payne, Gertrude. 1911. *Everyday Errors in Pronunciation, Spelling, and Spoken English*. San Francisco: Ricardo J. Orozco.

Perrin, Porter G. 1950. *A Writer's Guide and Index to English*. Chicago: Scott, Foresman, and Company, revised edn.

Perrin, Porter G. and George H. Smith. 1955. *The Perrin-Smith Handbook of Current English*. Chicago: Scott, Foresman, and Company.

Peters, Pam. 2004. *The Cambridge Guide to English Usage*. Cambridge: Cambridge University Press.

Pickett, Joseph P., Steven Kleinedler, and Susan Spitz, eds. 2005. *The American Heritage Guide to Contemporary Usage and Style*. Boston and New York: Houghton Mifflin.

Pinker, Steven. 2014. *The Sense of Style*. London and New York: Allan Lane.

Randall, Bernice. 1988. *Webster's New World Guide to Current American Usage*. New York: Prentice Hall, Simon & Schuster.

Shaw, Harry. 1975. *Dictionary of Problem Words and Expressions*. New York: McGraw-Hill. First published in 1975.

Stilman, Anne. 1997. *Grammatically Correct: The Writer's Essential Guide to Punctuation, Spelling, Style, Usage and Grammar*. Cincinnati, Ohio: Writer's Digest.

Stratton, Clarence. 1949. *Guide to Correct English*. New York and Toronto: McGraw-Hill Book Company, Inc.

Straus, Jane. 2008. *The Blue Book of Grammar*. San Francisco: Jossey-Bass.

Strunk, William. 1918. *The Elements of Style*. Geneva, N.Y.: Privately printed at the Press of W.F. Humphrey.

Sutcliffe, Andrea J. 1994. *The New York Public Library Writer's Guide to Style and Usage*. New York: Harper Collins Publishers.

Trask, Robert Lawrence. 2001. *Mind the Gaffe: The Penguin Guide to Common Errors in English*. London and New York: Penguin Books.

Treble, H.A. and G.H. Vallins. 1937. *An A.B.C. of English Usage*. New York: Oxford University Press.

Turck Baker, Josephine. 1910. *The Correct Word: How to Use It*. Chicago: The Correct English Publications.

Utter, Robert P. 1916. *Every-day Words and Their Uses: A Guide to Correct Diction*. Harper & Brothers.

Vermes, Jean C. 1981. *Secretary's Modern Guide to English Usage*. New York: Parker Publishing Company, Inc.

Vizetelly, Frank H. 1920. *A Desk-Book of Errors in English*. New York: Funk & Wagnalls, revised edn. First published in 1906.

Walsh, Bill. 2004. *The Elephants of Style: A Trunkload of Tips on the Big Issues and Gray Areas of Contemporary American English*. New York: McGraw-Hill.

White, Richard Grant. 1870. *Words and Their Uses, Past and Present: A Study of the English Language*. New York: Sheldon & Company.

Wilson, Kenneth G. 1993. *The Columbia Guide to Standard American English*. New York: Columbia University Press.

Witherspoon, Alexander M. 1948. *Common Errors in English and How to Avoid Them*. New York: Permabooks. First published in 1943.

Wood, George B. and Clarence Stratton. 1934. *The Blue Book of Good English*. Garden City, N.Y.: Nelson Doubleday, Inc. First published in 1926.

Secondary Sources

Aarts, Flor. 1994. Relative *who* and *whom*: prescriptive rules and linguistic reality. *American Speech*, 69(1):71–79.

Agheyisi, Rebecca N. and Joshua A. Fishman. 1970. Language attitude studies: a brief survey of methodological approaches. *Anthropological Linguistics*, 12(5):137–157.

Albakry, Mohammed. 2007. Usage prescriptive rules in newspaper language. *Southern Journal of Linguistics*, 30(2):28–56.

Albanyan, Ahmed and Dennis R. Preston. 1998. What is standard American English? *Studia Anglica Posnaniensia*, 34:29–46.

Algeo, John. 1991a. The best of the genre. *English Today*, 7(2):52–57.

Algeo, John. 1991b. Sweet are the usages of diversity. *Word*, 42(1):1–17.

Algeo, John. 1994. Pan-atlantic usage guidance. In Greta D. Little and Michael Montgomery, eds., *Centennial Usage Studies (Publication of the American Dialect Society, No. 78)*, pp. 101–107. Tuscaloosa: University of Alabama Press.

Allen, Harold B. 1935. The standard of usage in freshman textbooks. *English Journal*, 24:564–571.

Allen, Robert E. 2009. Dictionaries of usage. In A.P. Cowie, ed., *The Oxford History of English Lexicography, Vol. II: Specialized Dictionaries*, pp. 340–360. Oxford: Oxford University Press.

Anderwald, Lieselotte. 2002. *Negation in Non-standard British English: Gaps, Regularizations, and Asymmetries*. London and New York: Routledge.

Anderwald, Lieselotte. 2004. The varieties of English spoken in the southeast

of England. In Bernd Kortmann, Kate Burridge, Rajend Mesthrie, Edgar W. Schneider, and Clive Upton, eds., *A Handbook of Varieties of English, Volume 2*, pp. 175–195. Berlin and New York: Mouton de Gruyter.

Anderwald, Lieselotte. 2012. Clumsy, awkward or having a peculiar propriety? Prescriptive judgements and language change in the 19th century. *Language Sciences*, 34(1):28–53.

Anderwald, Lieselotte. 2014. Measuring the success of prescriptivism: quantitative grammaticography, corpus linguistics and the progressive passive. *English Language and Linguistics*, 18(1):1–21.

Anderwald, Lieselotte. 2016. *Language Between Description and Prescription: Verbs and Verb Categories in Nineteenth-Century Grammars of English*. Oxford: Oxford University Press.

Anderwald, Liselotte. 2005. Negative concord in British English dialects. In Yoko Iyeiri, ed., *Aspects of English Negation*, pp. 113–137. Amsterdam and Philadelphia: John Benjamins.

Arnovick, Leslie K. 1997. Proscribed collocations with *shall* and *will*: the eighteenth-century non-standard reassessed. In Jenny Cheshire and Dieter Stein, eds., *Taming the Vernacular: From Dialect to Written Standard Language*. Longman.

Auer, Anita. 2006. Precept and practice: the influence of prescriptivism on the English subjunctive. In Christiane Dalton-Puffer, Dieter Kastovsky, Nikolaus Ritt, and Herbert Schendl, eds., *Syntax, Style and Grammatical Norms: English from 1500–2000*, pp. 33–53. Bern: Peter Lang.

Auer, Anita. 2009. *The Subjunctive in the Age of Prescriptivism*. London: Palgrave Macmillan.

Auer, Anita and Victorina González-Díaz. 2005. Eighteenth-century prescriptivism in English: a re-evaluation of its effects on actual language usage. *Multilingua*, 24(4):317–341.

Austin, Frances. 1994. The effect of exposure to standard English: the language of William Clift. In Dieter Stein and Ingrid Tieken-Boon van Ostade, eds., *Towards a Standard English, 1600–1800*, pp. 285–313. Berlin and New York: Mouton de Gruyter.

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

Bailey, Richard W. 1992. *Images of English: A Cultural History of the Language*. Cambridge: Cambridge University Press (first published by the University of Michigan Press in 1991).

Baker, Anne and Kees Hengeveld, eds. 2012. *Linguistics*. Malden, Mass.: John Wiley & Sons.

Baker, C.L. 1970. Double negatives. *Linguistic Inquiry*, 1(2):169–186.

Baker, Robert. 1770. *Reflections on the English Language, In the Nature of Vaugelas's Reflections on the French*. London.

Baron, Dennis E. 1982. *Grammar and Good Taste: Reforming the American Language*. New Haven and London: Yale University Press.

Beal, Joan C. 2009. Three hundred years of prescriptivism (and counting). In Ingrid Tieken-Boon van Ostade and Wim van der Wurff, eds., *Current Issues in Late Modern English*, pp. 35–55. Bern: Peter Lang.

Berk, Lynne M. 1994. Prescriptive versus descriptive: the role of reference books. In Greta D. Little and Michael Montgomery, eds., *Centennial Usage Studies (Publication of the American Dialect Society, No. 78)*, pp. 108–114. Tuscaloosa: University of Alabama Press.

Biber, Douglas. 1988. *Variation across Speech and Writing*. Cambridge: Cambridge University Press.

Biber, Douglas, Jesse Egbert, Bethany Gray, Rahel Oppliger, and Benedikt Szmrecsanyi. 2016. Variationist versus text-linguistic approaches to grammatical change in English: nominal modifiers of head nouns. In Merja Kytö and Päivi Pahta, eds., *Handbook of English Historical Linguistics*, pp. 351–375. Cambridge: Cambridge University Press.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Longman.

Bloomfield, Leonard. 1927. Literate and illiterate speech. *American Speech*, 2(10):432–439.

Bloomfield, Leonard. 1944. Secondary and tertiary responses to language. *Language*, 20(2):45–55.

Blyth, Carl, Sigrid Recktenwald, and Jenny Wang. 1990. I'm like, "Say what?!": a new quotative in American oral narrative. *American Speech*, 65(3):215.

Bodine, Ann. 1975. Androcentrism in prescriptive grammar: singular 'they', sex-indefinite 'he', and 'he or she'. *Language in Society*, 4(2):129–146.

Boyland, Joyce Tang. 2001. Hypercorrect pronoun case in English? Cognitive processes that account for pronoun usage. In Joan Bybee and Paul J. Hopper, eds., *Frequency and the Emergence of Linguistic Structure*, pp. 383–404. Amsterdam and Philadelphia: John Benjamins.

Breheny, Patrick and Woodrow Burchett. 2017. Visualization of regression models using visreg. *The R Journal*, 9(2):56–71.

Brugman, Claudia. 1984. The very idea: a case study in polysemy and cross-lexical generalizations. *Papers from the Parasession on Lexical Semantics*, pp. 21–38.

Bryant, M.M. 1946. The split infinitive. *College English*, 8(1):39–40.

Bryer, Jason and Kimberly Speerschneider. 2017. *likert: Analysis and Visualization Likert Items*. http://jason.bryer.org/likert.

Buchstaller, Isabelle and Karen Corrigan. 2011. How to make intuitions succeed: testing methods for analysing syntactic microvariation. In Warren Maguire and April McMahon, eds., *Analysing Variation in English*, pp. 30–48. Cambridge: Cambridge University Press.

Burchfield, Robert. 1991. The Fowler brothers and the tradition of usage handbooks. In Gerhard Leither, ed., *English Traditional Grammars: An International Perspective*, pp. 93–111. Amsterdam and Philadelphia: John Benjamins.

Busse, Ulrich. 2015. Transatlantic perspectives on late nineteenth-century english usage: Alford (1864) compared to White (1871). In Marina Dossena, ed., *Transatlantic Perspectives on Late Modern English*, pp. 73–98. Amsterdam and Philadelphia: John Benjamins Publishing Company.

Busse, Ulrich and Anne Schröder. 2009. Fowler's *Modern English Usage* at the interface of lexis and grammar. In Ute Römer and Rainer Schulze, eds., *Exploring the Lexis-Grammar Interface*, pp. 69–87. Amsterdam and Philadelphia: John Benjamins Publishing Company.

Busse, Ulrich and Anne Schröder. 2010. Problem areas of English grammar between usage, norm, and variation. In Alexandra N. Lenz and Albert Plewnia, eds., *Grammar Between Norm and Variation*, pp. 87–102. Peter Lang.

Calhoun, Kendra. 2015. "It is the worst of our time": youth language, language attitudes, and arguments about literally. *Texas Linguistics Forum*, 58:1–10.

Calle-Martín, Javier and Antonio Miranda-García. 2009. On the use of split infinitives in English. In Andrew Kehoe and Antoinette Renouf, eds., *Corpus*

Linguistics: Refinements and Reassessments, pp. 347–364. Amsterdam: Brill Academic Publishers.

Cameron, Deborah. 1995. *Verbal Hygiene*. London and New York: Routledge.

Cameron, Deborah. 2012. The commodification of language: English as a global commodity. In Terttu Nevalainen and Elizabeth Closs Traugott, eds., *The Oxford Handbook of the History of English*, pp. 352–362. Oxford: Oxford University Press.

Campbell-Kibler, Kathryn. 2006. *Listener Perceptions of Sociolinguistic Variables*. Stanford University: PhD Dissertation.

Campbell-Kibler, Kathryn. 2010. The sociolinguistic variant as a carrier of social meaning. *Language Variation and Change*, 22(3):423–441.

Card, William, Raven I. McDavid, and Virginia McDavid. 1984. Dimensions of usage and dictionary labeling. *Journal of English Linguistics*, 17(1):57–74.

Chafe, Wallace. 1984. Speaking, writing, and prescriptivism. In Deborah Schiffrin, ed., *Meaning, Form, and Use in Context: Linguistic Applications*, pp. 95–103. Washington, D.C.: Georgetown University Press.

Chambers, Jack K. 2004. Dynamic typology and vernacular universals. In Bernd Kortmann, ed., *Dialectology Meets Typology*, pp. 127–145. Berlin and New York: Mouton de Gruyter.

Chambers, Jack K. and Peter Trudgill. 1980. *Dialectology*. Cambridge: Cambridge University Press.

Chapman, Don. 2017. Stalwarts, SNOOTS and some readers: How 'Traditional Rules' are traditional. In Ingrid Tieken-Boon van Ostade and Carol Percy, eds., *Prescription and Tradition in Language: Establishing Standards across Time and Space*, pp. 238–252. Bristol: Multilingual Matters.

Cheshire, Jenny. 1981. Variation in the use of *ain't* in an urban British English dialect. *Language in Society*, 10(3):365–381.

Cheshire, Jenny. 1982. *Variation in an English Dialect: A Sociolinguistic Study*. Cambridge: Cambridge University Press.

Claridge, Claudia. 2010. *Hyperbole in English: A Corpus-based Study of Exaggeration*. Cambridge: Cambridge University Press.

Cohen, Jacob. 1988. *Statistical Power Analysis for the Behavioral Sciences*. New York: Academic Press.

Connors, Robert J. 1983. Handbooks: history of a genre. *Rhetoric Society Quarterly*, 13(2):87–98.

Corder, Gregory W. and Dale I. Foreman. 2009. *Nonparametric Statistics for Non-Statisticians: A Step-by-Step Approach*. Hoboken: Wiley.

Creswell, Tomas J. 1975. *Usage in Dictionaries and Dictionaries of Usage*. Alabama: University of Alabama Press.

Crisp, Raymond. 1971. *Change in Attitudes toward English Usage*. University of Illinois: PhD Dissertation.

Croissant, Yves. 2013. *mlogit: multinomial logit model*. R package version 0.2-4.

Crystal, David. 2006. *The Fight for English: How Language Pundits Ate, Shot, and Left*. Oxford: Oxford University Press.

Curme, George O. 1927. The split infinitive. *American Speech*, 2(8):341.

Curme, George O. 1935. *A Grammar of the English Language, in Three Volumes. Vol. II: Parts of Speech and Accidence*. Boston: D.C. Heath and Company.

Curzan, Anne. 2014. *Fixing English: Prescriptivism and Language History*. Cambridge: Cambridge University Press.

Dailey-O'Cain, Jennifer. 2000. The sociolinguistic distribution of and attitudes toward focuser *like* and quotative *like*. *Journal of Sociolinguistics*, 4(1):60–80.

D'Arcy, Alexandra. 2006. Lexical replacement and the like(s). *American Speech*, 81(4):339–357.

D'Arcy, Alexandra. 2007. *Like* and language ideology: disentangling fact from fiction. *American Speech*, 82(4):386–419.

Davies, Mark. 2008–present. *The Corpus of Contemporary American English (COCA): 560 million words, 1990–present*. Available online at https://corpus.byu.edu/coca/.

Davies, Mark. 2009. The 385+ million word *Corpus of Contemporary American English* (1990–2008+): design, architecture, and linguistic insights. *International Journal of Corpus Linguistics*, 14(2):159–190.

Davies, Mark. 2010a. The Corpus of Contemporary American English as the first reliable monitor corpus of English. *Literary and Linguistic Computing*, 25(4):447–464.

Davies, Mark. 2010b. *The Corpus of Historical American English (COHA): 400 million words, 1810–2009*. Available online at https://corpus.byu.edu/coha/.

Davies, Mark. 2012. Expanding horizons in historical linguistics with the 400 million word Corpus of Historical American English. *Corpora*, 7:121–157.

Dekeyser, Xavier. 1975. *Number and Case Relations in 19th Century British English: A Comparative Study of Grammar and Usage*. Antwerp: De Nederlandsche Boekhandel.

Denison, David. 1998. Syntax. In Suzanne Romaine, ed., *The Cambridge History of the English Language, Volume 4*, pp. 92–329. Cambridge: Cambridge University Press.

Dobbins, Austin C. 1956. The language of the cultivated. *College English*, 18(1):46–47.

Dragojevic, Marko, Howard Giles, and Bernadette M. Watson. 2013. Language ideologies and language attitudes: a foundational framework. In Howard Giles and Bernadette Watson, eds., *The Social Meanings of Language, Dialect and Accent: International Perspectives on Speech Styles*, pp. 1–25. New York: Peter Lang.

Drake, Glendon F. 1977a. American linguistic prescriptivism: its decline and revival in the 19th century. *Language in Society*, 6(3):323–340.

Drake, Glendon F. 1977b. *The Role of Prescriptivism in American Linguistics, 1820–1970*. Amsterdam: John Benjamins.

Ebner, Carmen. 2016. Language guardian BBC? investigating the BBC's language advice in its 2003 news styleguide. *Journal of Multilingual and Multicultural Development*, 37(3):308–320.

Ebner, Carmen. 2017. *Proper English Usage: A Sociolinguistic Investigation of Attitudes towards Usage Problems in British English*. University of Leiden: PhD Dissertation.

Facchinetti, Roberta. 2000. The modal verb *shall* between grammar and usage in the nineteenth century. In Arthur Mettinger and Dieter Kastovsky, eds., *The History of English in a Social Context: A Contribution to Historical Sociolinguistics*, pp. 115–133. Berlin and New York: De Gruyter Mouton.

Feagin, Crawford. 1979. *Variation and Change in Alabama English: A Sociolinguistic Study of the White Community*. Washington, D.C.: Georgetown University Press.

Finegan, Edward. 1980. *Attitudes Toward English Usage: The History of a War of Words*. New York and London: Columbia University Teachers College Press.

Finegan, Edward. 1985. Unconscious attitudes toward linguistic variation. In Sidney

Greenbaum, ed., *The English Language Today*, pp. 92–98. Oxford: Pergamon Institute of English.

Fischer, Olga. 2000. Grammaticalisation: unidirectional, non-reversable? The case of *to* before the infinitive in English. In Olga Fischer, Anette Rosenbach, and Dieter Stein, eds., *Pathways of Change: Grammaticalization in English*, pp. 149–169. Amsterdam and Philadelphia: John Benjamins.

Fischer, Roswitha. 2007. To boldly split the infinitive – or not? Prescriptive traditions and current English usage. In Stephan Elspaß, Nils Langer, Joachim Scharloth, and Wim Vandenbussche, eds., *Germanic Language Histories 'From Below' (1700–2000)*, pp. 259–274. Berlin: De Gruyter.

Fitzmaurice, Susan M. 1998. The commerce of language in the pursuit of politeness in eighteenth-century England. *English Studies*, 79(4):309–328.

Fitzmaurice, Susan M. 2000a. The great leveler: the role of the spoken media in stylistic shift from the colloquial to the conventional. *American Speech*, 75(1):54–68.

Fitzmaurice, Susan M. 2000b. Remarks on the de-grammaticalisation of infinitival *to* in present-day American English. In Olga Fischer, Anette Rosenbach, and Dieter Stein, eds., *Pathways of Change: Grammaticalization in English*, pp. 171–186. Amsterdam and Philadelphia: John Benjamins.

Fries, Charles C. 1925. The periphrastic future with *shall* and *will* in Modern English. *Publications of the Modern Language Association of America*, 40(4):963–1024.

Fries, Charles C. 1940. *American English Grammar*. New York and London: D. Appleton-Century.

Fromkin, Victoria A., ed. 2000. *Linguistics: An Introduction to Linguistic Theory*. Malden, Mass.: Wiley.

Fuller, Janet M. 2003. Use of the discourse marker *like* in interviews. *Journal of Sociolinguistics*, 7(3):365–377.

Garrett, Peter. 2001. Language attitudes and sociolinguistics. *Journal of Sociolinguistics*, 5(4):626–631.

Garrett, Peter. 2010. *Attitudes to Language*. Cambridge: Cambridge University Press.

Giles, Howard and Tamara Rakic. 2014. Language attitudes: social determinants and consequences of language variation. In Thomas M. Holtgraves, ed., *The Oxford Handbook of Language and Social Psychology*, pp. 11–26. Oxford: Oxford University Press.

Goatly, Andrew. 1997. *The Language of Metaphors*. London and New York: Routledge.

Gramley, Stephan and Kurt-Michael Pätzold. 1992. *A Survey of Modern English*. London: Routledge.

Greenbaum, Sidney and Janet Whitcut. 1988. *Longman Guide to English Usage*. Harlow: Longman.

Gustafsson, Larisa Oldireva. 2002. *Preterite and Past Participle Forms in English 1680–1790: Standardisation Processes in Public and Private Writing*. Uppsala: Uppsala University.

Harrell, Frank E. 2018. *rms: Regression Modeling Strategies*. R package version 5.1-2.

Hinrichs, Lars, Benedikt Szmrecsanyi, and Axel Bohmann. 2015. *Which*-hunting and the Standard English relative clause. *Language*, 91(4):806–836.

Hock, Hans Henrich and Brian D. Joseph. 2009. *Language History, Language Change, and Language Relationship: An Introduction to Historical and Comparative Linguistics*. Berlin: De Gruyter.

Hodson, Jane. 2006. The problem of Joseph Priestley's (1733–1804) descriptivism. *Historiographia Linguistica*, 33(1):57–84.

Howe, Darin M. 1997. Negation and the history of African American English. *Language Variation and Change*, 9(2):267–294.

Howe, Darin M. and James A. Walker. 2000. Negation and the creole-origins hypothesis: evidence from early African American English. In Shana Poplack, ed., *The English History of African American English*, pp. 109–140. Malden, Mass.: Blackwell.

Huddleston, Rodney and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge.

Ilson, Robert. 1985. Usage problems in British and American English. In Sidney Greenbaum, ed., *The English Language Today*, pp. 166–182. Oxford: Pergamon Institute of English.

Irvine, Judith. 1989. When talk isn't cheap: language and political economy. *American Ethnologist*, 16(2):248–267.

Israel, Michael. 2002. *Literally* speaking. *Journal of Pragmatics*, 34(4):423–432.

Iyeiri, Yoko. 2001. *Negative Constructions in Middle English*. Fukuoka: Kyushu University Press.

Jespersen, Otto. 1940. *A Modern English Grammar on Historical Principles, Part V: Syntax (Fourth Volume)*. London and Copenhagen: George Allen & Unwin Ltd./Ejnar Munskgaard.

Kachru, Braj B. 1992. Models for non-native Englishes. In Braj B. Kachru, ed., *The Other Tongue. English Across Cultures*, pp. 48–74. Urbana, Ill.: University of Illinois Press.

Kato, Kazuo. 2001. *Not to be* or *to not be*: more on split negative infinitives. *American Speech*, 76(3):312–315.

Kjellmer, Goran. 1997. On contraction in modern English. *Studia Neophilologica*, 69(2):155–186.

Klima, Edward S. 1964. Relatedness between grammatical systems. *Language*, 40:1–20.

Konopka, Marek. 1996. *Strittige Erscheinungen der Deutschen Syntax im 18. Jahrhundert*. Tübingen: Max Niemeyer Verlag.

Kostadinova, Viktorija. 2015. Literally, too big a fuss about nothing. *English Today*, 31(2):3–4.

Kostadinova, Viktorija. 2018a. Correcting English: Josephine Turck Baker (1873–1942) and the early American usage guide tradition. In Linda Pillière, Wilfrid Andrieu, Valérie Kerfelec, and Diana Lewis, eds., *Standardising English: Norms and Margins in the History of the English Language*, pp. 171–190. Cambridge: Cambridge University Press.

Kostadinova, Viktorija. 2018b. Usage problems in American English. In Ingrid Tieken-Boon van Ostade, ed., *English Usage Guides: History, Advice, Attitudes*, pp. 155–175. Oxford: Oxford University Press.

Kostadinova, Viktorija. forthcoming. Prescriptivism as a constraint in language variation and change: moving beyond binary oppositions. In Don Chapman and Jacob D. Rawlins, eds., *Values and Multiplicity: Identity and Fluidity in Prescriptivism and Descriptivism*. Bristol: Multilingual Matters.

Kostadinova, Viktorija. in press. Attitudes to usage vs. actual language use: the case of *literally* in American English. *English Today*.

Kroch, Anthony and Cathy Small. 1978. Grammatical ideology and its effect on speech. In David Sankoff, ed., *Linguistics Variation: Models and Methods*, pp. 45–55. New York: Academic Press.

Kroskrity, Paul. 2004. Language ideologies. In Alessandro Duranti, ed., *A Companion to Linguistic Anthropology*, pp. 496–517. Malden, Mass.: Blackwell.

Labov, William. 1972a. Negative attraction and negative concord in English grammar. *Language*, 48(4):773–818.

Labov, William. 1972b. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Labov, William, Paul Cohen, Clarence Robins, and John Lewis. 1968. *A Study of the Non-standard English of Negro and Puerto Rican Speakers in New York City*. New York: Columbia University.

Laitinen, Mikko. 2009. Singular YOU WAS/WERE variation and English normative grammars in the eighteenth century. In Minna Palander-Collin, Minna Nevala, and Arja Nurmi, eds., *The Language of Daily Life in England (1400–1800)*, pp. 199–217.

Landau, Sidney I. 2001. *Dictionaries: The Art and Craft of Lexicography*. Cambridge: Cambridge University Press.

Lavandera, Beatriz R. 1978. Where does the sociolinguistic variable stop? *Language in Society*, 7(2):171–182.

Leech, Geoffrey, Marianne Hundt, Christian Mair, and Nicholas Smith. 2009. *Change in Contemporary English: A Grammatical Study*. Cambridge: Cambridge University Press.

Leech, Geoffrey and Nicholas Smith. 2009. Change and constancy in linguistic change: how grammatical usage in written English evolved in the period 1931–1991. In Andrew Kehoe and Antoinette Renouf, eds., *Corpus Linguistics: Refinements and Reassessments*, pp. 173–200. Amsterdam: Brill Academic Publishers.

Leonard, Sterling A. 1929. *The Doctrine of Correctness in English Usage, 1700–1800*. New York: Columbia University.

Leonard, Sterling A. 1932. *Current English Usage*. Chicago: The Inland Press.

Levshina, Natalia. 2015. *How to Do Linguistics with R: Data Exploration and Statistical Analysis*. Amsterdam and Philadelphia: John Benjamins.

Lippi-Green, Rosina. 1997. *English with an Accent: Language, Ideology, and Discrimination in the United States*. London and New York: Routledge.

Long, Daniel and Dennis R. Preston. 2002. *Handbook of Perceptual Dialectology, Volume 2*. Amsterdam and Philadelphia: John Benjamins.

Lukač, Morana. 2018. *Grassroots Prescriptivism*. University of Leiden: PhD Dissertation.

Macaulay, Ronald. 2001. You're like 'why not?' The quotative expressions of Glasgow adolescents. *Journal of Sociolinguistics*, 5(1):3–21.

MacKay, Donald G. 1980. On the goals, principles, and procedures for prescriptive grammar: singular *they*. *Language in Society*, 9(3):349–367.

Mair, Christian. 2006. *Twentieth-Century English: History, Variation, and Standardization*. Cambridge: Cambridge University Press.

Mair, Christian and Geoffrey Leech. 2006. Current changes in English syntax. In Bas Aarts and April M.S. McMahon, eds., *The Handbook of English Linguistics*, pp. 318–342. Malden, Mass. and Oxford: Blackwell.

Malmstrom, Jean. 1964[1958]. Linguistic atlas findings vs texbooks pronouncements on current American usage. In Harold B. Allen, ed., *Readings in Applied English Linguistics*, pp. 316–323. New York: Appleton-Century-Crofts.

Malone, Kemp. 1941. The split infinitive and a system of clauses. *American Speech*, 16(1):52–54.

Marckwardt, Albert H. 1968. Language standards and attitudes. In Albert H. Marckwardt, ed., *Language and Language Learning*, pp. 1–22. Champaign, Ill.: National Council of Teachers of English.

Marckwardt, Albert H. and Fred G. Walcott. 1938. *Facts About Current English Usage*. New York: Appleton-Century-Crofts.

McArthur, Tom. 1986. The usage industry. *English Today*, 2(3):8.

McCarthy, Michael and Ronald Carter. 2004. "There's millions of them": hyperbole in everyday conversation. *Journal of Pragmatics*, 36(2):149–184.

McDavid, Raven I. 1941. *Ain't I* and *aren't I*. *Language*, 17(1):57–59.

McDavid, Virginia G. 1973. Variations in dictionary labeling practices. In Raven I. McDavid and Audrey R. Duckert, eds., *Lexicography in English*, pp. 187–207. New York: New York Academy of Sciences.

McKenzie, Robert M. 2010. *The Social Psychology of English as a Global Language*. Dordrecht: Springer.

Meyers, Walter E. 1991. Usage glossaries in current college handbooks. *American Speech*, 66(4):341–358.

Meyers, Walter E. 1995. Linguistics in textbooks: a forty-year comparison. *American Speech*, 70(1):30–68.

Milroy, James. 2001. Language ideologies and the consequences of standardization. *Journal of Sociolinguistics*, 5(4):530–555.

Milroy, James and Lesley Milroy. 1985. *Authority in Language: Investigating Language Prescription and Standardisation*. London and New York: Routledge.

Milroy, James and Lesley Milroy. 2012. *Authority in Language: Investigating Standard English*. London and New York: Taylor & Francis, 4th edn.

Mittins, W.H., Mary Salu, Mary Edminson, and Sheila Coyne. 1970. *Attitudes to English Usage*. London: Oxford University Press.

Molencki, Rafał. 2003. Proscriptive prescriptivists: on the loss of the 'pleonastic' perfect infinitive in counterfactual constructions in Late Modern English. In Marina Dossena and Charles Jones, eds., *Insights into Late Modern English*, pp. 175–196. Bern: Peter Lang.

Nerlich, Brigitte and Pedro J. Chamizo Domínguez. 2003. The use of *literally*: Vice or virtue? *Annual Review of Cognitive Linguistics*, 1(1):193–206.

Nevalainen, Terttu. 1997. The processes of adverb derivation in Late Middle and Early Modern English. In Matti Rissanen, Merja Kytö, and Kirsi Heikkonen, eds., *Grammaticalization at Work*, pp. 145–190. Berlin and New York: Mouton de Gruyter.

Nevalainen, Terttu. 2000. Gender differences in the evolution of standard English: evidence from the corpus of early English correspondence. *Journal of English Linguistics*, 28(1):38–59.

Nevalainen, Terttu. 2006a. Historical linguistics and language change. In Ans van Kemenade and Bettelou Los, eds., *The Handbook of the History of English*, pp. 558–588. Oxford: Blackwell.

Nevalainen, Terttu. 2006b. Negative concord as an English "vernacular universal": social history and linguistic typology. *Journal of English Linguistics*, 34(3):257–278.

Nevalainen, Terttu and Helena Raumolin-Bunberg. 2003. *Sociohistorical linguistics: Language Change in Tudor and Stuart England*. London and New York: Routledge.

Pateman, Trevor. 1982. MacKay on singular *they*. (discussion). *Language in Society*, 11(3):437–438.

Perales-Escudero, Moisés D. 2011. To split or to not split: the split infinitive past and present. *Journal of English Linguistics*, 39(4):313–334.

Percy, Carol. 2009. Periodical reviews and the rise of prescriptivism: the *Monthly* (1749–1844) and *Critical Review* (1756–1817) in the eighteenth century. In Ingrid Tieken-Boon van Ostade and Wim van der Wurff, eds., *Current Issues in Late Modern English*, pp. 117–150. Bern: Peter Lang.

Peters, Pam. 2006. English usage: prescription and description. In Bas Aarts and April M.S. McMahon, eds., *The Handbook of English Linguistics*, pp. 759–780. Malden, Mass. and Oxford: Blackwell.

Peters, Pam. 2012. Usage guides and Australian English: prescription and description. In Anne Schröder, Ulrich Busse, and Ralf Schneider, eds., *Codification, Canons and Curricula: Description and Prescription in Language and Literature*, pp. 247–264. Bielefeld: Aisthesis Verlag.

Peters, Pam. 2014. Usage guides and usage trends in Australian and British English. *Australian Journal of Linguistics*, 34(4):581–598.

Peters, Pam and Wendy Young. 1997. English grammar and the lexicography of usage. *Journal of English Linguistics*, 25(4):315–331.

Poplack, Shana. 2006. How English became African American English. In Ans van Kemenade and Bettelou Los, eds., *The Handbook of the History of English*, pp. 452–479. Oxford: Blackwell.

Poplack, Shana and Nathalie Dion. 2009. Prescription vs. praxis: the evolution of future temporal reference in French. *Language*, 85(3):357–387.

Powell, Mava Jo. 1992. Folk theories of meaning and principles of conventionality: encoding literal attitude via stance adverb. In Adrienne Lehrer and Eva Feder Kitay, eds., *Frames, Fields, and Contrasts: New Essays in Semantic and Lexical Organization*, pp. 333–353. New York and London: Routledge.

Preston, Dennis R. 1991. Sorting out the variables in sociolinguistic theory. *American Speech*, 66(1):33–56.

Preston, Dennis R. 1999a. *Handbook of Perceptual Dialectology, Volume 1*. Amsterdam and Philadelphia: John Benjamins.

Preston, Dennis R. 1999b. A language attitude analysis of regional US speech: is northern US English not friendly enough? *Cuadernos de Filologia Inglesa*, 8:129–146.

Pullum, Geoffrey K. 2010a. The land of the free and *The Elements of Style*. *English Today*, 26(2):34–44.

Pullum, Geoffrey K. 2010b. These 'rules' are already broken. *Times Higher Education*, p. 56.

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

R Core Team. 2013. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Roberts, Sam. 2009. 'The Elements of Style' turns 50. *The New York Times*, p. C3.

Romaine, Suzanne and Deborah Lange. 1991. The use of *like* as a marker of reported speech and thought: a case of grammaticalization in progress. *American Speech*, 66(3):227–279.

Sankoff, Gillian. 1972. Above and beyond phonology in variable rules. In Charles-James N. Bailey and Roger W. Shuy, eds., *New Ways of Analyzing Variation in English*, pp. 44–61. Washington, D.C.: Georgetown University Press.

Schourup, Lawrence. 1983. Common discourse particles in English conversation. *Ohio State Working Papers in Linguistics*, 28.

Schweiger, Beth Barton. 2010. A social history of English grammar in the early United States. *Journal of the Early Republic*, 30(4):533–555.

Seright, Orin Dale. 1966. Double negatives in Standard Modern English. *American Speech*, 41(2):123.

Silverstein, Michael. 1976. Shifters, linguistic categories, and cultural description. In Keith H. Basso and Henry A. Selby, eds., *Meaning in Anthropology*, pp. 11–56. Albuquerque: University of New Mexico Press.

Silverstein, Michael. 1979. Language structure and linguistic ideology. In Paul R. Clyne, William F. Hanks, and Carol L. Hofbauer, eds., *The Elements: a Parasession on Linguistic Units and Levels Including Papers from the Conference on Non Slavic Languages of the USSR*, pp. 193–248. Chicago: Chicago Linguistic Society.

Silverstein, Michael. 1996. Monoglot "standard" in America: standardization and metaphors of linguistic hegemony. In Donald Brenneis and Ronald H.S. Macaulay, eds., *The Matrix of Language: Contemporary Linguistic Anthropology*, pp. 284–306. Boulder: Westview Press.

Silverstein, Michael. 2003. Indexical order and the dialectics of sociolinguistic life. *Language and Communication*, 23(3):193–229.

Singh, Rajendra. 1996. *Lectures Against Sociolinguistics*. New York: Peter Lang.

Smith, Jennifer. 2001. Negative concord in the old and new world: evidence from Scotland. *Language Variation and Change*, 13(2):109–134.

Speelman, Dirk, Adriaan Spruyt, Leen Impe, and Dirk Geeraerts. 2013. Language attitudes revisited: auditory affective priming. *Journal of Pragmatics*, 52:83–92.

Stenetorp, Pontus, Sampo Pyysalo, Goran Topić, Tomoko Ohta, Sophia Ananiadou, and Jun'ichi Tsujii. 2012. brat: a web-based tool for nlp-assisted text annotation. In *Proceedings of the Demonstrations Session at EACL 2012*.

Stevens, Martin. 1954. The derivation of 'ain't'. *American Speech*, 29(3):196.

Straaijer, Robin. 2009. Deontic and epistemic modals as indicators of prescriptive and descriptive language in the grammars of Joseph Priestley and Robert Lowth. In Ingrid Tieken-Boon van Ostade and Wim van der Wurff, eds., *Current Issues in Late Modern English*, pp. 57–87. Bern: Peter Lang.

Straaijer, Robin. 2014. *Hyper Usage Guide of English*. University of Leiden: http://huge.ullet.net.

Straaijer, Robin. 2015. *The Hyper Usage Guide of English Database: User Manual*. University of Leiden. Online publication.

Straaijer, Robin. 2018. The usage guide: evolution of a genre. In Ingrid Tieken-Boon van Ostade, ed., *English Usage Guides: History, Advice, Attitudes*, pp. 11–29. Oxford: Oxford University Press.

Sundby, Bertil, Anne Kari Bjørge, and Kari E. Haugland. 1991. *A Dictionary of English Normative Grammar, 1700–1800*. Amsterdam and Philadelphia: John Benjamins.

Szmrecsanyi, Benedikt. 2016. About text frequencies in historical linguistics: disentangling environmental and grammatical change. *Corpus Linguistics and Linguistic Theory*, 12(1):153–171.

Tagliamonte, Sali A. 2012. *Variationist Sociolinguistics: Change, Observation, Interpretation*. Malden, Mass.: Wiley-Blackwell.

Thomason, Sarah G. 2007. Language contact and deliberate language change. *Journal of Language Contact – THEMA*, 1:41–62.

Thomason, Sarah G. 2011. *Language Contact*. Edinburgh: Edinburgh University Press.

Tibbetts, A. M. 1966. Were nineteenth-century textbooks really prescriptive? *College English*, 27(4):309–315.

Tieken-Boon van Ostade, Ingrid. 1982. Double negation and eighteenth-century English grammars. *Neophilologus*, 66:278–285.

Tieken-Boon van Ostade, Ingrid. 1994. Standard and non-standard pronominal usage in English, with special reference to the eighteenth century. In Dieter Stein and Ingrid Tieken-Boon van Ostade, eds., *Towards a Standard English, 1600–1800*, pp. 217–242. Berlin and New York: Mouton de Gruyter.

Tieken-Boon van Ostade, Ingrid. 2000. Normative studies in England. In Sylvian Auroux, E.F.K. Koerner, Hans-J. Niederehe, and Kees Versteegh, eds., *History of the Language Sciences/Geschichte der Sprachwissenschaften/Histoire des sciences du langage*, pp. 876–887. Berlin and New York: Walter de Gruyter.

Tieken-Boon van Ostade, Ingrid. 2002. *You was* and eighteenth-century normative grammar. In Katja Lenz and Ruth Möhlig, eds., *On Dyuersite & Chaunge of Langage: Essays Presented to Manfred Görlach on the Occasion of his 65th Birthday*, pp. 88–102. Heidelberg: C. Winter Universitätsverlag.

Tieken-Boon van Ostade, Ingrid. 2006. Eighteenth-century prescriptivism and the norm of correctness. In Ans van Kemenade and Bettelou Los, eds., *The Handbook of the History of English*, pp. 539–557. Oxford: Blackwell.

Tieken-Boon van Ostade, Ingrid. 2008a. The codifiers and the history of multiple negation, or, why were 18th-century grammarians so obsessed with double negation? In Joan C. Beal, Carmela Nocera, and Massimo Sturiale, eds., *Perspectives on Prescriptivism*, pp. 197–214. Bern: Peter Lang.

Tieken-Boon van Ostade, Ingrid, ed. 2008b. *Grammars, Grammarians, and Grammar-writing in Eighteenth-century England*. Berlin and New York: Mouton de Gruyter.

Tieken-Boon van Ostade, Ingrid. 2010. The usage guide: its birth and popularity. *English Today*, 26(2):14–23.

Tieken-Boon van Ostade, Ingrid. 2011. *The Bishop's Grammar: Robert Lowth and the Rise of Prescriptivism*. Oxford: Oxford University Press.

Tieken-Boon van Ostade, Ingrid. 2012a. The codification of English in England. In Raymond Hickey, ed., *Standards of English: Codified Varieties Around the World*, pp. 34–54. Cambridge: Cambridge University Press.

Tieken-Boon van Ostade, Ingrid. 2012b. Codifying the English language. In Anne Schröder, Ulrich Busse, and Ralf Schneider, eds., *Codification, Canons and*

Curricula: Description and Prescription in Language and Literature, pp. 61–77. Bielefeld: Aisthesis Verlag.

Tieken-Boon van Ostade, Ingrid. 2012c. English at the onset of the normative tradition. In Lynda Mugglestone, ed., *The Oxford History of English*, pp. 240–273. Oxford: Oxford University Press.

Tieken-Boon van Ostade, Ingrid. 2015. *Five Hundred Mistakes Corrected*: an early American usage guide. In Marina Dossena, ed., *Transatlantic Perspectives on Late Modern English*, pp. 55–72. Amsterdam and Philadelphia: John Benjamins.

Tieken-Boon van Ostade, Ingrid. 2016. Usage guides and the age of prescriptivism. Paper presented at the *19th International Conference on English Historical Linguistics*, University of Duisburg-Essen, Germany.

Tieken-Boon van Ostade, Ingrid. forthcoming. *Describing Prescriptivism*.

Tieken-Boon van Ostade, Ingrid and Carmen Ebner. 2017. Prescriptive attitudes to usage. In *Oxford Research Encyclopedia of Linguistics*. Oxford: Oxford University Press. Available online at: http://linguistics.oxfordre.com.

Tottie, Gunnel. 1997. Literacy and prescriptivism as determinants of linguistic change. In Uwe Böker and Hans Sauer, eds., *Anglistentag, Dresden 1996*, pp. 83–93. Trier: Wissenschaftlicher Verlag Trier.

Trudgill, Peter. 2009. Vernacular universals and the sociolinguistic typology of English dialects. In Markku Filppula, Juhani Klemola, and Heli Paulasto, eds., *Vernacular Universals and Language Contacts: Evidence from Varieties of English and Beyond*, pp. 304–322. New York and London: Routledge.

Tucker, G. Richard and Wallace E. Lambert. 1972. White and Negro listeners' reactions to various American-English dialects. In Joshua A. Fishman, ed., *Advances in the Sociology of Language*, pp. 175–184. The Hague and Paris: Mouton.

Ueda, Reed. 1987. *Avenues to Adulthood: The Origins of the High School and Social Mobility in an American Suburb*. Cambridge: Cambridge University Press.

Ukaji, Masatomo. 1999. On the scope of negative concord. In Ingrid Tieken-Boon van Ostade, Gunnel Tottie, and Wim van der Wurff, eds., *Negation in the History of English*, pp. 269–294. Berlin and New York: Mouton de Gruyter.

Underhill, Robert. 1988. *Like* is, like, focus. *American Speech*, 63(3):234.

van der Wouden, Ton. 1997. *Negative Contexts: Collocation, Polarity and Multiple Negation*. London and New York: Routledge.

Vorlat, Emma. 1979. Criteria of grammaticalness in 16th- and 17th-century English. *Leuvense Bijdragen: Tijdschrift voor Germaanse Filologie*, 68(2):129–140.

Walker, James A. 2005. The *ain't* constraint: *not*-contraction in early African American English. *Language Variation and Change*, 17(1):1–17.

Weiner, Edmund. 1988. On editing a usage guide. In E.G. Stanley and T.F. Hoad, eds., *Words: For Robert Burchfield's Sixty-fifth Birthday*, pp. 171–183. Cambridge: D.S. Brewer.

Weldon, Tracey. 1994. Variability in negation in African American Vernacular English. *Language Variation and Change*, 6(3):359–397.

Wickham, Hadley. 2009. *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Wild, Catherine. 2010. *Attitudes towards English Usage in the Late Modern Period: the Case of Phrasal Verbs*. University of Glasgow: PhD Dissertation.

Wolfram, Walt. 1969. *A Sociolinguistic Description of Detroit Negro Speech*. Washington, D.C.: Center for Applied Linguistics.

Wolfram, Walt. 1974. *Sociolinguistic Aspects of Assimilation: Puerto Rican English in New York City*. Arlington, Va.: Center for Applied Linguistics.

Wolfram, Walt. 2004. Social varieties of American English. In Edward Finegan and John R. Rickford, eds., *Language in the U.S.A.: Themes for the 21st Century*, pp. 58–75. Cambridge: Cambridge University Press.

Wolfram, Walt and Donna Christian. 1976. *Appalachian Speech*. Arlington, Va.: Center for Applied Linguistics.

Wolfram, Walt and Ralph W. Fasold. 1974. *The Study of Social Dialects in American English*. New York: Prentice-Hall.

Wolfram, Walt and Natalie Schilling. 2016. *American English: Dialects and Variation*. Malden, Mass.: Wiley Blackwell, 3rd edn.

Woolard, Kathryn A. and Bambi B. Schieffelin. 1994. Language ideology. *Annual Review of Anthropology*, 23(1):55–82.

Yáñez-Bouza, Nuria. 2015. *Grammar, Rhetoric and Usage in English: Preposition Placement 1500–1900*. Cambridge: Cambridge University Press.

Samenvatting in het Nederlands

De laatste jaren neemt de belangstelling voor het prescriptivisme gestaag toe en dit proefschrift hoopt een bijdrage te leveren aan het wetenschappelijke werk op dit vlak. Het onderzoek richt zich op de invloed die het prescriptivisme heeft op het Amerikaans-Engelse taalgebruik en op de opvattingen van taalgebruikers in de Verenigde Staten hierover.

Deze invloed wordt onderzocht met behulp van zowel corpus- als sociolinguïstische methodologieën. Het onderzoek volgt een drieledige aanpak, en biedt analyses van (a) het metalinguïstische discours in een corpus van zeventig tussen 1874 en 2014 verschenen Amerikaanse taalgidsen, (b) de gevolgen van het prescriptivisme op gesproken en geschreven Amerikaans-Engels taalgebruik door middel van corpusstudies en (c) de opvattingen van 79 sprekers van het Amerikaans-Engels over zes bekende, maar niet aan elkaar gerelateerde, gestigmatiseerde vormen en constructies, namelijk *ain't*, het partikel *like*, het niet letterlijke *literally*, dubbele ontkenning, *I* als lijdend voorwerp en *me* als onderwerp, en de opgebroken infinitief constructie (*the split infinitive*); voorbeelden hieronder.

(1) He thinks he **ain't** a man any more.

(2) Is there some way I can, **like**, throw a bouquet to him at the hearing today?

(3) I don't even know who these people are, and suddenly they have **literally** exploded into the American consciousness.

(4) I've got a snug estate, and **don't owe nobody** anything.

(5) You know, **Bernie and me** used to talk, and he'd say "Hey Jerry, I know you feel the same way".

(6) It's important **to significantly improve** working conditions.

Er is voor deze zes gestigmatiseerde constructies gekozen om de invloed te kunnen onderzoeken van het prescriptivisme op zowel van oudsher bekende ongewenste constructies, zoals de dubbele ontkenning en de opgebroken infinitief constructie, als op nieuwere vormen van taalgebruik, zoals *literally* en *like*.

De eerste drie hoofdstukken schetsen de achtergrond van het onderwerp. In hoofdstuk 1 worden de uitgangspunten en de relevantie van het onderwerp besproken. Het tweede hoofdstuk bevat een discussie van relevante literatuur over taalgidsen, het prescriptivisme, taalvariatie en -verandering, ideologieën, en opvattingen over taal. Het proefschrift benadert gestigmatiseerd taalgebruik als taalkundige variatie, en pleit ervoor om in de context van variatiepatronen de rol van het prescriptivisme te onderzoeken op basis van empirisch bewijs. In hoodstuk 3 wordt een overzicht gegeven van het tot nu toe verrichte wetenschappelijk onderzoek naar de zes constructies.

In hoofdstuk 4 worden de methodologieën voor elk van de drie onderzochte aspecten uiteengezet. Ten eerste wordt zowel de verzameling taalgidsen beschreven die voor dit onderzoek is gebruikt als de manier waarop het material is geanalyseerd. Hiermee wordt de basis gelegd voor de discussie die in hoofdstuk 5 gevoerd wordt. Ten tweede bevat het hoofdstuk een beschrijving van het corpus dat de empirische basis vormt voor het vaststellen van de feitelijke gebruikspatronen die in hoofdstuk 6 nader worden besproken. Tenslotte beschrijft hoofdstuk 4 de wijze waarop het onderzoek naar de opvattingen van taalgebruikers is uitgevoerd en hoe de verkregen data zijn geanalyseerd.

Hoofdstukken 5, 6 en 7 behandelen de bevindingen met betrekking tot elk van de gestigmatiseerde constructies en worden afgesloten met een korte synthese. Hoofdstuk 5 bevat een metalinguïstische analyse van de bespreking van de zes constructies in de taalgidsen. Drie belangrijke punten komen daarbij naar voren. Ten eerste blijkt dat de behandeling van de ongewenste constructies in de loop der tijd verandert. *Ain't* en de opgebroken infinitief constructie raken geleidelijk aan meer geaccepteerd, terwijl andere, zoals *literally* and *like*, negatiever benaderd lijken te worden. Ten tweede blijkt dat ondanks de gesignaleerde verschillen in behandeling, het genre taalgidsen als zodanig bezig is op te schuiven naar een meer beschrijvende, op feiten gebaseerde benadering. Dat valt onder meer af te leiden uit de stijgende aantallen verwijzingen in taalgidsen naar daadwerkelijk taalgebruik.

Hoofdstuk 6 bespreekt de resultaten van de studie naar de variatie in het gebruik

van de zes constructies. Twee verschillende benaderingen zijn gebruikt: een tekstuele (*text-linguistic*) en een context-variabele (*variationist*). De analyse toont aan dat de constructies verschillende variatie- en veranderingspatronen hebben ondergaan. De dubbele ontkenning, *I* als lijdend voorwerp en *me* als onderwerp worden consequent relatief weinig genoemd, zowel per genre als door de tijd heen, terwijl de frequentie van *ain't* laag is door de tijd heen, maar wel gevoeligheid voor teksttype laat zien. Daarentegen laten *like*, *literally* en de opgebroken infinitief constructie alle drie een stijgende frequentie zien. Dat zou erop kunnen wijzen dat het prescriptivisme meer invloed op sommige gestigmatiseerde constructies heeft dan op andere. Met andere woorden, de frequenties op zich leveren niet veel bewijs voor de invloed van het prescriptivisme.

Om dit nader te kunnen onderzoeken is de opgebroken infinitief constructie in een *case study* gebruikt, waarbij bekeken is hoe het gebruik daarvan in de afzonderlijke teksten van het corpus correleert met het gebruik van andere constructies die door prescriptivisten worden afgekeurd. De richting van de correlatie doet inderdaad prescriptieve invloed vermoeden: teksten waarin *shall* and *whom* (vaker) voorkomen bevatten ook meer opgebroken infinitief constructies dan teksten die geen *shall* and *whom* bevatten. Een mogelijke verklaring hiervoor kan zijn dat taalgebruikers die ertoe neigen voorgeschreven items zoals *shall* en *whom* te gebruiken, ook "verboden" constructies zoals de opgebroken infinitief constructie vermijden. Teksten waarin verboden vormen zoals *ain't* en *literally* gebruikt worden, hebben een grotere kans ook opgebroken infinitief constructies te bevatten. Met andere woorden, taalgebruikers die opgebroken infinitief constructies vermijden, hebben de neiging ook andere ongewenste constructies te vermijden. Tenslotte bleek dat terwijl het gebruik van de opgebroken infinitief constructie in de loop der tijd stijgt, dit niet geldt voor alle delen van het corpus: kranten laten een teruglopend gebruik zien. Dit lijkt te suggereren dat de invloed van het prescriptivisme per genre variëert.

Hoofdstuk 7 bespreekt aan de hand van de enquête onder en interviews met 79 sprekers van het Amerikaans Engels hun houding van ten opzichte van de ongewenste constructies. De resultaten hiervan komen grotendeels overeen met de conclusies van andere studies naar de opvattingen van taalgebruikers, zij het met nieuwe kwalitatieve en kwantitatieve gegevens. De analyse laat ook zien dat de houding van sprekers vaak overeenkomt met algemene prescriptieve opvattingen, in het bijzonder zoals die in de taaladviesliteratuur worden aangetroffen.

De algemene conclusie is dat de tweedeling prescriptief-descriptief opnieuw beoordeeld en geijkt moet worden. Het is tijd om te beseffen dat prescriptieve

opvattingen en feitelijk taalgebruik elkaar wederzijds beïnvloeden. Wanneer we de invloed van het prescriptivisme willen meten, moeten we onderscheid maken tusssen taalvariatie en taalverandering. Hoewel de conclusie van het hier gepresenteerde onderzoek is dat het prescriptivisme geen blijvende effecten heeft gehad op taalverandering, dat wil zeggen op de ontwikkeling van individuele veranderingen, datzelfde precriptivisme heeft wel invloed op taalvariatie met betrekking tot andere sociale categorieën of categorieën van taalgebruik, in het bijzonder teksttype.

Curriculum Vitae

Viktorija Kostadinova graduated in English Language and Literature from the Ss. Cyril and Methodius University in Skopje in 2010. During her undergraduate studies, she spent one semester as an exchange student at Ghent University, thanks to a European Union scholarship. She obtained an MA degree in English Language and Literature in 2011, and an MA degree in Cultural Studies in 2013, both from the Katholieke Universiteit Leuven, Belgium. In September 2012, she started work on her PhD project at Leiden University. In 2013, she spent one month at the LSA Linguistic Institute at the University of Michigan.

She has been teaching English in various forms and at various levels since 2004. In Macedonia, she taught English to primary and secondary students, as well as adult learners of English. During her undergraduate studies in Skopje, she taught courses in English phonetics and phonology, and English grammar. At Leiden University, she taught courses in sociolinguistics, language ideology and prescriptivism, and corpus linguistics, and in the last two years she has been teaching academic writing in the BA International Studies programme. While at Leiden University, she was also member and later chair of the PhD Council, and worked as Information Officer for the MA Linguistics programme. Since 2017 she has been teaching English linguistics to students in the English Language and Culture programme at the University of Amsterdam, as well as academic writing courses to media studies students at this university.