MASTERARBEIT | MASTER'S THESIS

Titel | Title

What are the odds (ratios)? A corpus-based diachronic exploration of the collocational strengths of the pronouns *he* and *she* with select constructions to quantify gender bias

verfasst von | submitted by

Jayou Oh BA

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Arts (MA)

Wien | Vienna, 2024

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 812
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium English Language and Linguistics
Betreut von | Supervisor:	Univ.-Prof. Mag. Dr. Nikolaus Ritt

Abstract

This thesis contributes to the literature on the topic of gender bias in language. Gender bias and asymmetries in language have received scholarly attention since the 1970s, with renewed interest in detecting such linguistic inequalities arising from the 2010s onwards following developments in machine learning. Much of the research has focused on gender bias associated with nouns, verbs and adjectives on the basis of corpus data, through which it has been demonstrated that nouns and pronouns denoting males and females differ in their collocational tendencies with certain words. Beyond the level of individual words, there is conversely little written about how overtly gendered terms interact with different constructions; research on the historical progression of gender bias is also sparse. This research thus addresses these overlooked aspects through diachronic investigations of bias trends specifically observable for various verb constructions, using odds ratio as a statistical measure to quantify gender bias. The exploratory study involved a dataset consisting of 26 constructions as well as 36 verbs. Data were collected from the Google Books (American English) Corpus. By examining to what degree the gendered subject pronouns *he* and *she* collocate with select constructions containing grammatical features such as tense, aspect, modality and negation, it was possible to determine whether grammatical constructions exhibit gender bias, which constructions appear more biased than others, and how bias trends changed in the span of two centuries. The findings suggest that verb constructions do display small yet pervasive and statistically measurable bias, gender bias decreased overall between the 19th and 21st century, and simple tense constructions register greater bias than progressive, perfect or modal constructions. Moreover, the past tense is more associated with female bias while present tense constructions tend to show male bias, negative constructions may be more biased than positive ones, and the gradual reduction in biased language use can seemingly more likely be attributed to weakening female bias than to jointly diminishing bias for both genders. The results so indicate that constructions are not as innocuous as they seem and should be scrutinised further in matters relating to gender bias.

Zusammenfassung

Diese Arbeit trägt zur Literatur des Themas geschlechtsspezifische Voreingenommenheit in der Sprache bei. Seit den 1970ern haben Gender-Bias und geschlechtsspezifische Asymmetrien in der Sprache wissenschaftliche Aufmerksamkeit erhalten, wobei das Interesse an der Erkennung solcher sprachlichen Ungleichheiten seit den 2010ern im Zuge der Entwicklungen des Bereiches maschinelles Lernen erneut zugenommen hat. Ein Großteil der Forschung hat sich auf Gender-Bias verbunden mit Substantiven, Verben und Adjektiven in Korpusdaten konzentriert. Dabei wurde gezeigt, dass Substantive und Pronomen, die Männer und Frauen bezeichnen, sich in ihren Kollokationstendenzen mit bestimmten Wörtern unterscheiden. Über die Ebene der einzelnen Wörter hinaus gibt es hingegen wenig Recherche darüber, wie geschlechtsspezifische Begriffe mit verschiedenen Konstruktionen interagieren; so ist auch die Forschung zum historischen Verlauf des Gender-Bias spärlich. Diese Forschung befasst sich daher mit diesen übersehenen Aspekten durch diachrone Untersuchungen von Bias-Tendenzen, die speziell für verschiedene Verbkonstruktionen beobachtbar sind, wobei das Chancenverhältnis als statistisches Maß zur Quantifizierung des Bias verwendet wird. Die Sondierungsstudie beinhaltet einen Datensatz mit 26 Konstruktionen sowie 36 Verben. Die Daten wurden aus dem Google Books (American English) Corpus gesammelt. Durch das Untersuchen inwieweit die Subjektpronomen *he* und *she* mit ausgewählten Konstruktionen zusammenfallen, die grammatische Merkmale wie Zeitform, Aspekt, Modalität und Verneinung enthalten, war es möglich festzustellen, ob Konstruktionen Vorurteile auf Geschlecht beruhend aufweisen, welche Konstruktionen stärker befangen erscheinen als andere und wie sich die Bias-Trends im Verlauf von zwei Jahrhunderten veränderten. Die Ergebnisse deuten darauf hin, dass Verbkonstruktionen zwar kleine, aber dennoch verbreitete und statistisch messbare Befangenheit zeigen. Das geschlechtsspezifische Vorurteil nahm insgesamt zwischen dem 19. und 21. Jahrhundert ab, und einfache Zeitkonstruktionen weisen einen größeren Bias auf als progressive, perfekte oder modale Konstruktionen. Darüber hinaus ist die Vergangenheitsform stärker mit weiblichem Bias verbunden, während Konstruktionen im Präsens tendenziell männlichen Bias zeigen. Negative Konstruktionen können verzerrter sein als positive, und die allmähliche Reduktion der befangenen Sprachverwendung kann scheinbar eher der Abschwächung des weiblichen Bias als einer abnehmenden Bias beider Geschlechter zugeschrieben werden. Die Ergebnisse zeigen somit, dass Konstruktionen nicht so unverfänglich sind, wie sie scheinen, und in Bezug auf geschlechtsspezifische Befangenheit weiter untersucht werden sollten.

Table of contents

List of tables

List of figures

1. Introduction

The complex relationship between language and gender as well as the various underpinnings and implications thereof have consistently garnered substantial scholarly interest since the 1970s, coinciding with the growth of feminist movements in the United States (Eckert & McConnell-Ginet 2013: 37). As a result, perceived and perceivable gendered imprints and effects on language have been extensively explored and studied over the years with considerable focus on linguistic differences, stereotypes and biases, with most research chiefly concerned with the male and female genders. Indeed, there is a wide array of literature on topics such as purported gender variation in language use and investigations to that end (e.g. Lakoff 1973; Mulac & Lundell 1994; Biber & Burges 2000; Mondorf 2004; Newman et al. 2008; Leone 2022; Schmauss & Kilian 2023) and how men and women – and other such similarly gendered words themselves – are described, represented or written about (e.g. Sigley & Holmes 2002; Pearce 2008; Baker 2010; Jockers & Kirilloff 2016; Norberg 2016; Ye et al. 2018; Ritt et al. 2024), in addition to works identifying and addressing pervasive linguistic bias or sexism and gender stereotyping indicating male dominance in language (e.g. Bodine 1975; Pauwels 2003; Talbot 2003; Cameron 2023). These types of gender disparity in language itself, in particular, have been recognised as a problem, with Lakoff (1973: 73) remarking that these "linguistic imbalances" simultaneously hint at "real-world imbalances and inequities", being symptoms of a greater issue. Consequently, gender inequality in language is thus also considered to be a point of concern affecting society at large, necessitating addressing (Pauwels 2003: 553-554; Doughman et al. 2021: 34; Toro Isaza et al. 2023: 6510).

While some studies report that linguistic sexism and gender bias have generally become less pronounced since the 1970s based on the frequency of gendered terms or gender stereotypes in specific language data (Baker 2010: 133; Garg et al. 2018: 3639; Ritt et al. 2024: 17), gender bias in language continues to be an issue of relevance today, especially within the context of natural language processing and large language models as they also have implications for artificial intelligence (Prates, Avelar & Lamb 2020: 6363; Doughman et al. 2021: 34). For instance, the reproduction of gender stereotypes involving occupations and adjectives in machine translations from languages with gender-neutral pronouns to English – which has gendered pronouns – has been examined before (Cho et al. 2019; Prates, Avelar & Lamb 2020; Endriga & Rosario Jr. 2023). As a whole, it has been widely observed that language models and systems inherit then reflect (Bolukbasi et al. 2016: 4357; Zhao et al.

2019: 632; Babaeianjelodar et al. 2020: 752; Touileb 2022: 423) and even amplify societal biases from data sources (Basta, Costa-jussà & Casas 2019: 33; Bender et al. 2021: 617; Doughman et al. 2021: 36), leading researchers to devise and introduce methods to mitigate biases (e.g. Bolukbasi et al. 2016; Zhao et al. 2019) with the understanding that language models – memorising form but not comprehending meaning – trained on "large, uncurated, static datasets […] encode hegemonic views that are harmful" to society and detrimental to already-marginalised groups in particular (Bender et al. 2021: 615).

It is against this background of the importance of identifying and quantifying gender bias in the domain of language, along with addressing the very issue, that the research for this thesis was undertaken. The thesis presents a primarily quantitative exploratory study determining the collocational strengths of the gendered subject pronouns *he* and *she* with select verbs and various verb constructions accounting for tense, aspect, modality and negation, aided by diachronic corpus data and employing odds ratio as the key statistical measure for analysis. The idea is that the odds ratio values represent the collocational affinities of the pronouns to the constructions investigated and thus can be interpreted as indicating the level of gender bias by comparing male and female constructions, i.e. constructions preceded by either *he* or *she*, respectively. The corpus used for the study was the 155-billion-word Google Books (American English) Corpus by Davies (2011-) which provides written language data of published content in American English spanning two centuries, specifically from the 1810s until the 2000s. Since utilising diachronic corpora to study language also provides opportunities to track some possible changes in language use over time (cf. Hilpert 2020), the prospect of visualising and following trends and trajectories of gender bias associated with verb constructions was especially enticing.

This research conducted to ascertain gender bias on the level of verb constructions is warranted, considering, for example, that template-based gender bias probes for language models have been noted to be sensitive to both changes in grammatical tense and the introduction of negation to the test clauses (Touileb 2022: 426), and that female subjects appear to be more visible than their male counterparts before past tense constructions than they are before present tense constructions (Ritt et al. 2024: 25). These findings suggest that grammatical categories – such as tense – may perhaps be a factor in whether female or male terms are more inclined or likely to occur before certain verb constructions. This is not to say that this is a conscious or intentional decision on our part as we use language. Rather, it is

that precisely this aspect of language use in which gender bias is not or should not necessarily be expected is worth illuminating and exploring further.

Examining these tendencies and patterns displayed by pronouns collocating with select verb constructions is all the more intriguing since there is a remarkable paucity of research on collocational affinities of gendered terms with verb constructions despite the abundance of corpus-based research highlighting differences and stereotypes in the representations of gendered terms in texts which seemingly reflect social norms and biases (e.g. Sigley & Holmes 2002; Pearce 2008; Jockers & Kirilloff 2016; Norberg 2016; Piits 2016; Čermáková & Fárová 2017; Ye et al. 2018; Ritt et al. 2024). Most of these works focus on the collocational behaviours and patterns of gendered nouns and pronouns with adjectives and verbs, generally corroborating the postulation that gendered terms – *he*, *she*, *man*, *woman*, *boy*, *girl*, etc. – collocate more strongly with certain adjectives and verbs depending on whether the terms denote female or male entities. As such, it has been observed that adjectives and verbs relating to power, deviancy, physical activity and motion tend to take male subjects (Pearce 2008: 7-9; Jockers & Kirilloff 2016: 21; Norberg 2016: 298) while female entities are depicted as being more emotional or more likely to be victims of violence (Sigley & Holmes 2002: 148; Pearce 2008: 10; Jockers & Kirilloff 2016: 21; Norberg 2016: 302). Those findings indicate that the semantics of select words can roughly predict whether they show stronger association towards male terms or female ones. However, the question remains whether latent gender bias can likewise be discerned from certain grammatical constructions at the syntactic level.

Thus, the purpose of this research is mainly to explore what effects different verb constructions formed through variations of grammatical categories have on the collocational affinities of the gendered pronouns *he* and *she* with said constructions, which of the constructions are associated with higher levels of bias, which select verbs collocate more strongly with which pronoun, and how the degrees of identifiable biases have shifted over a period of two hundred years, all on the basis of written language data of published American English material. The methodological foundation and inspiration for the thesis are provided by research from Ritt et al. (2024), who, in investigating the levels of gender bias associated with nouns, verbs and adjectives in the Google Books corpora of American and British English, also utilise odds ratio as the prime statistical measure of bias. In effect, the analytic approach taken is an adapted variant of that employed for collostructional analyses as expounded by Stefanowitsch and Gries (2003), the difference being the measure of

association used. While the exploratory study constituting the thesis is principally quantitative in nature, aspects of qualitative analysis are also present to make sense of the results.

The thesis is structured as follows. After the introduction, background information to further understanding of concepts relevant to the subject matter is provided for context. A review of previous corpus-based studies on gender differences, biases and stereotypes in language data is presented in the third chapter. Subsequently, the research design comprising the data, queries and methodology utilised is provided in the fourth chapter which is then followed by the presentation of results and analyses. The conclusion forming the sixth and final chapter summarises the main points and findings of the thesis.

2. Background information

To facilitate understanding of the research and its context and to provide the background and rationale behind the thesis, major fundamental concepts deemed relevant to the study conducted are presented in this chapter.

2.1. Language and gender

Much has been said and written about the relationship and association between language and gender – in effect, how language is used by and about both men and women – as evidenced by the amount of literature on the topic (e.g. Eckert & McConnell-Ginet 2013; Coates 2016; Cameron 2023) in the decades since Robin Lakoff's seminal paper "Language and woman's place" was published in 1973. In it, Lakoff (1973: 49) argues that "women's language" differs from the language of men in various aspects such as in the lexicon and syntax. Lakoff (1973: 46-47), acknowledging that she bases her claims on her own observations and introspective examination of her speech and that of her acquaintances, cites examples such as colour words, "weaker" profanities and certain evaluative adjectives as being largely limited to women's vocabulary (Lakoff 1973: 49-51), noting that these factors impart a sense of frivolity, triviality and powerlessness to women's speech (Lakoff 1973: 52-53). In particular, the frequent use of tag questions (Lakoff 1973: 54-55) and polite language marked by hedging (Lakoff 1973: 56-57) are also given as being characteristic of female language, suggesting that women's speech effectively prevents women from being assertive in their statements (Lakoff 1973: 57). In addition to language use by women, Lakoff (1973: 46) also writes that general language use about women, e.g. select words referring to women which

have male counterparts, tends to be rife with inequity and negative connotations which "relegate women to subservient functions", rendering women as beings "not to be taken seriously" (Lakoff 1973: 60) and whose social identities are constructed by their relationship with men and male roles (Lakoff 1973: 68). These examples serve as evidence, as Lakoff (1973: 46) writes, "that women experience linguistic discrimination in two ways: in the way they are taught to use language, and in the way general language use treats them".

It is understood that Lakoff's (1973) paper, published during a time of rising social awareness, served as a catalyst for the development of the study of language and gender within the context of the United States (Eckert & McConnell-Ginet 2013: 37). With the pioneering article as basis, subsequent literature sought to address the claims made by Lakoff (1973), either to confirm or disprove them through empirical studies (Eckert & McConnell-Ginet 2013: 38). Perhaps unsurprisingly, many research works on language and gender thus have had – and continue to have – the tendency to focus on investigating the suspected discrepancies between and differences in language as used by men and by women or the lack thereof, a propensity also remarked upon by Eckert and McConnell-Ginet (2013: 50). Indeed, there are numerous studies examining or questioning potential gender differences (or similarities) along the lines of Lakoff's (1973) observations, with regard, for example, to syntax at large (Mondorf 2004), vulgar language (Ginsburg, Ogletree & Silakowski 2003), tag questions (Dubois & Crouch 1975; Hepburn & Potter 2012), hedging (Holmes 1986; Schmauss & Kilian 2023) and tentative speech and language (Preisler 1986; Leaper & Robnett 2011), including the use of modals in spoken language data (Leone 2022). There also are additional related studies on gender differences in earlier forms of English backgrounded by topics mentioned in Lakoff's (1973) article as surveyed by Nevalainen (2002), findings which show that evidentiality markers in British English – a varied set of select adverbs, cognitive verbs, speaker-subject and speaker-recipient verbs, and modals – are used by women "statistically significantly more often than men" (Berglind Söderqvist 2017: 30), and a study on Peninsular Spanish requests which found that there is "a significant difference in the distribution of verb forms" between women's and men's requests with women favouring the use of the "softened" subjunctive and modal conditionals and men more frequently using the direct imperative for requests (Czerwionka, Staszkiewicz & Shamloo 2023: 19). It can thus be seen that although scholarship has since moved on from the initial approach of studying language and gender taken by Lakoff (1973) which Coates (2016: 6) identifies as an emblematic example of the "deficit approach", i.e. implying that women's language is "weak

and unassertive [and] deficient" relative to male language,[1] the interest in neutrally, descriptively exploring gender difference in language clearly remains.

This interest extends to discussions and analyses of writing as well as literature and fiction at large; there are contentious assumptions about gender differences in the very style of writing, that the feminine literary style is supposedly marked by greater use of "epistemic modality" and "grammatically complex, meandering" sentences (Livia 2003: 144). On the basis of informal written essays which involved describing landscape photographs, Mulac and Lundell (1994: 304) found linguistic differences in male and female writing, with men more frequently employing quantitative terms, references to location and judgmental adjectives whereas women made greater use of verbs marking uncertainty, progressive verbs and long sentences among other features. Mulac and Lundell (1994: 307) also note, however, that a "substantial degree of cross-gender linguistic overlap" exists, as evidenced by students' inability to correctly guess the sex of the writers. In a related study, Newman et al. (2008: 223), utilising mixed written and spoken language data, identify differences wherein women were more likely to use pronouns, social terms, references to psychological processes, verbs and negations than men while male language was more concerned with immediate events and marked by more frequent profanities, differences considered to be "small but systematic" (Newman et al. 2008: 233). There is also a piece of research on dramatic dialogue in plays by Biber and Burges (2000: 26) in which they found that male writers have tended to favour male-to-male dialogue whereas female authors "portray women as talking much more equitably to both men and women". In addition to aforementioned examples where the aspect of language use is concerned, literary works also occasionally show that norms and expectations of linguistic gender – represented primarily through gendered pronouns – can be masked or subverted through novel means so as to make the language used about characters intentionally ambiguous (Livia 2003: 150-151).

The intriguing nature of gender-based differences notwithstanding, it is of course unwise to posit that female and male uses of language must always be different or that there are profound inherent differences in the ways that women and men speak or write. Cameron (2007: 163-164) asserts that "[t]here is a great deal of similarity between men and women" and that "many differences are context-dependent", stressing the fact that other variables, such as group-belonging and communities within the genders, produce as many differences

[1] Cameron (2005: 485) rather views Lakoff's (1973) approach as having been an example of the "dominance approach", "in which the emphasis was placed on the effects of gender inequality".

as those between men and women. That gender is but one of several determinant factors in linguistic variation has been observed early on, such as in the sociolinguistic variationist work on the English spoken in Norwich by Trudgill (1972) which also took into account social class and age. Romaine (2003: 103), noting that women have the tendency to use prestige language variants whilst being perceived as possessing lower status than men in many cultures, also writes that "both gender and language comprise rather complex social practices and performances", such that the effect of gender on language choice "is not straightforward but mediated through other identities and ideologies" (Romaine 2003: 116). Other variables existing in tandem with gender in the study of language and gender gained special prominence from the 1990s onwards in what Cameron (2005: 483) refers to as the "'postmodern' turn".

The turn away from focusing solely on gender differences in the study of language and gender symbolised a paradigm shift in which the "diversity of gendered and sexual identities" became central to the study (Cameron 2005: 482). This shift in the approach to language and gender was motivated by the understanding and view that gender and sexuality are characterised by fluidity and diversity (Cameron 2005: 494; Eckert & McConnell-Ginet 2013: 41) as well as the fact that language practices are varied in an increasingly globalised world (Cameron 2005: 499), giving rise to the modern prevailing "dynamic approach" to inquiries on the relationship and interaction between language and gender which entails adopting a "social constructionist perspective" as described by Coates (2016: 6). The current approach to language and gender thus rejects the view of gender as a static binary (Holmes 2007: 55), with the knowledge that while gender hierarchy still does exist, "the extremely overt and rigid gender hierarchies of the past" are no longer present in many contexts today (Cameron 2005: 490). Obviously, this is not to say there are no differences between genders or that hierarchies have been completely eliminated. As Eckert and McConnell-Ginet (2013: 20-21) point out, "gender is asymmetrical" and the gender order maintains the "inequality […] built into gender at a very basic level". Holmes (2007: 53) invokes the persistence of the gender order – explaining it as the "repressive ideology which ensures that deviation from gender norms […] entails penalties" – in making the argument that while the expanded paradigm to examine language and gender has proven instrumental in helping deepen recognition and understanding of the complex and fluid concept of gender as well as emphasising the "heterogeneity within women's and men's linguistic practices", "a focus on women's ways of

talking is [still] warranted" as workplaces, for instance, are one setting in which "gendered expectations about normative ways of speaking persist" (Holmes 2007: 59-60).

Nevertheless, the acknowledgement and acceptance of the concept of gender as a fluid social construct (Eckert & McConnell-Ginet 2013: 41-42, 50; Coates 2016: 6) is a particularly important point, considering the tendency and prevalent inclination to treat and portray gender as a male-female binary in discourse for oversimplified categorisation purposes (Eckert & McConnell-Ginet 2013: 214). To limit gender categorisation into a dichotomy of male and female is understandably problematic. Yet, such a division is to be expected when conducting research on gender-based linguistic differences for the main reason that the binary represents the societal norm, such that many sets of available language data reflect this. Most research on gender differences or gender bias therefore restricts gender to a binary system for operational purposes, simultaneously noting that gender is or should not be construed as such. For example, Berglind Söderqvist (2017: 19) explicitly states that although "gender is generally seen as a fluid concept", the study itself utilises a binary gender variable owing to a lack of "large spoken corpora […] with information about speakers' age and gender that include more than two gender categories" and established norms. Garg et al. (2018: 3636) similarly indicate that their analysis is confined to "the two major binary categories due to technical limitations"; Czerwionka, Staszkiewicz and Shamloo (2023: 23) likewise point out that while "gender is not a binary category", their research "addresses the language of women and men because these categories represent the [study] participants". Binary categorisation is a noted issue also in the field of computational linguistics, particularly in natural language processing – Stańczak and Augenstein (2021: 2) cite the constant adherence to "modelling gender as a binary variable" as being especially problematic since doing so has the potential to cause harm such as "erasure via invalidation or obscuring of non-binary gender identities", criticising existing research on gender bias for only considering two genders and not accounting for "the fluidity and continuity of gender as a variable".

It must be noted here that due to limitations posed by the data necessary for analysis, this thesis adopts the same approach as other papers have done in categorising gender as a binary for the purposes of this research on determining and quantifying gender differences in language used about men and women while remaining cognisant that gender is fluid.

2.2. Gender bias and stereotyping in language

One aspect of language and gender that has often been underscored, discussed and scrutinised is the prevalence of gender bias and stereotyping in language. As Romaine (2003: 112) observes, language does not merely reflect society but also constructs it. This interplay is a salient point as it suggests that both societal roles or norms and the language used by society at large affect each other interdependently; social observations and perceptions as well as imbalances and inequities become projected onto language, which then can itself cause these values to become internalised, ingrained, accepted and propagated through usage (Lewis & Lupyan 2020: 1021). This way, societal biases and stereotypes are then perpetuated and made pervasive, becoming difficult to decouple and eliminate.

It is accepted that bias is a steadfast fixture in everyday lives, something cognitively "hardwired" into humans to allow us to easily compartmentalise and categorise experience at the cost of reproducing prejudices (Doughman et al. 2021: 34; Einersen, Villesèche & Huopalainen 2021: 44). Gender biases and stereotypes in language can be variably seen as manifesting themselves as symbolic representations of gender norms (Einersen, Villesèche & Huopalainen 2021: 46), appearing as a mirror to existing societal gender hierarchy (Doughman et al. 2021: 34) or arising due to patriarchal prescriptivism (Bodine 1975: 133). Gender bias is understood, for example, to be "exclusionary, implicitly prejudicial" and offering a generalised "representation of a specific gender as a function of various societal stereotypes" (Doughman et al. 2021: 35) and "the preference […] toward one gender over the other" (Sun et al. 2019: 1630), whereby even the term "gender binarism" has been used synonymously to gender bias in at least one paper (Ritt et al. 2024: 3). Closely related to bias is the concept of stereotyping, a practice that Talbot (2003: 470) describes as involving "simplification, reduction, and naturalization"; as such, while they "are not lies" *per se*, they are considered to be "exaggerations with a purpose" (Eckert & McConnell-Ginet 2013: 58) with the harmful potential to discriminate and reproduce unflattering depictions of a group (Stańczak & Augenstein 2021: 5). Stereotypes pertinent to gender, then, are tied to societal expectations of gender and perpetuate the notion of gender differences (Talbot 2003: 472). Thus, in Talbot's (2003: 472) view, gender stereotypes are utilised to maintain "hegemonic male dominance and female subordination", such that men are frequently popularly deemed "authoritative language users" compared to women who are "garrulous, frivolous, and illiterate language users" (Pauwels 2003: 550), or that men are active doers and women are passive experiencers (Eckert & McConnell-Ginet 2013: 174).

The idea of male dominance in language – that language is biased towards and favours men – has long been discussed in conjunction with gender differences, with Lakoff (1973: 74) having drawn attention to the use of the masculine pronoun as the norm when referring to mixed-sex groups as well as words like *man* and *mankind* which encompass both men and women. The use and history of the masculine generic *he* was further expounded upon by Bodine (1975), pointing out "androcentrism in prescriptive grammar" – that is to say, enforced male-centredness in English. In her article, Bodine (1975: 134-135) highlights the role prescriptive grammarians from the 18th century onwards played in seeking the eradication of both the inclusive phrase *he or she* and the singular and sex-indefinite *they* which had previously been in common currency and instead attempting to impose the "linguistic embodiment of an androcentric world-view" through the generic *he*. As Bodine (1975: 133) puts it herself, "linguistically, human beings were to be considered male unless proven otherwise" by those grammarians. On an interesting note, Bodine (1975: 141) further alludes to possibilities of "pronominal changes" in the future, mentioning the pushback against the sex-indefinite *he*; it would appear that she was somewhat correct in her predictions. The more frequent use of *they* or even newly-created pronouns were entertained already in the late 1970s and found considerable acceptance in the 1980s (Harrigan & Lucic 1988: 138), the frequency of *he or she* in printed works increased substantially since the 1970s (Ritt et al. 2024: 18) as well as in spoken academia at the turn of the 21st century (Pauwels 2003: 564), and the use of *they* alongside other newly-derived pronouns for non-binary individuals is common practice today (Lindqvist, Renström & Gustafsson Sendén 2019: 111). With the women's movement of the 1970s spearheading the drive to challenge blatant male bias in language as well as its regulation by demonstrating the discriminatory nature of such biases and stereotypes not just in English but in other languages as well (Pauwels 2003: 551, 553; Cameron 2023: 90-91), in several ways, the countermovement against "default male language" did achieve some success with the ongoing language reform efforts toward inclusive language in general (Cameron 2023: 84).

Despite great strides made in reducing overtly sexist or male-marked words over the years, e.g. in the field of professions (Pauwels 2003: 563), it is nonetheless important to remember that the pervasiveness of male bias in language is not confined to its appearance in pronouns or other nouns. The persistence of male dominance is still identifiable, for instance, in the word order of mixed-gender binomial phrases where the male term tends to precede the female term (Doughman et al. 2021: 40) when examining language at a smaller level, and

also at the macro-level of language in the presumption that being an effective leader is predicated on adopting "stereotypically masculine discourse behaviours" (Holmes 2007: 58). In everyday colloquial speech, androcentric biases continue to be found in the form of genericised male labels such as *man*, *dude* and *guy*, used by both women and men to refer to both genders (MacArthur, Cundiff & Mehl 2020: 89). Male bias in language comprehension has been found to exist even in neutral noun constructions, i.e. words without overt gender markings such as *applicant* and singular *they*, the bias only getting eliminated when newly-created gender-neutral pronouns are used (Lindqvist, Renström & Gustafsson Sendén 2019: 111). Male prevalence in language is perhaps most visible in quantifiable terms in collections of text; in a corpus subsample of drama texts, Biber and Burges (2000: 26) mention the ubiquity of male over female language in dramatic dialogue, with male authors especially favouring male discourse over female discourse in their works, and the fact that male writers are much better represented in the subcorpus (Biber & Burges 2000: 24). Similarly, Baker (2010: 132, 143) found that – in terms of absolute frequency – male terms such as male pronouns, the words *man* and *men*, and the title *Mr*, although collectively decreasing in usage, are nevertheless more commonly represented in corpora than their female counterparts.

Beyond obvious displays of bias in language, there is additionally the issue of implicit or unconscious bias. Albeit based on a study of a large dataset of doctoral theses written in Spanish – a language with grammatical gender – it has nevertheless been recognised that "unconscious bias is characteristic of academic texts" when taking into account non-inclusive language (Orgeira-Crespo et al. 2021: 15). Hauser and Schwarz (2022: 2) suggest that the set of words which collocates with a particular word both reflect and reinforce implicit biases in human minds; they write that the subtle, non-explicit company kept by words – "biased company in language" – suffice in producing such implicit biases in cognition such that linguistic biases involving gender inequities, e.g. collocations associating men in some occupations and women in others, fosters implicit bias with implications for children's literature as well (Hauser & Schwarz 2022: 9). In sentence completion tasks for measuring implicit causality bias, Ferstl, Garnham and Manouilidou (2011: 133) found that participants completing sentence fragments ending in the conjunction *because* attributed negative events to being more likely to be caused by men and positive events by women. A follow-up study on implicit consequentiality bias which had participants complete sentence fragments ending in *and so* report similar results such that verbs likely to be followed by male referents suffering the consequences have the tendency to be negative and verbs likely to followed by

female referents tend to be more positive (Garnham, Vorthmann & Kaplanova 2021: 1537). Other unobvious ways in which gender bias is effectuated are of social relevance, such as in employment – it has been suggested that women in academia face hurdles in their jobs through these implicit inequalities which see maintenance in language (Einersen, Villesèche & Huopalainen 2021: 44); for example, Rubini and Menegatti (2014: 638), examining language abstraction in academic selection judgements for associate professorial positions, found that in rejection statements, there was greater use of concrete negative descriptions for male applicants than for women, suggesting "discrimination against female applicants was primarily perpetrated through the use of more abstract negative language". They add that the findings demonstrate how "women's academic careers can be hindered through the implicit yet strategic use of language abstraction" (Rubini & Menegatti 2014: 642).

Any realisation of linguistic bias is understandably considered to be detrimental and disadvantageous to any select group of people (Stańczak & Augenstein 2021: 5-6; Hauser & Schwarz 2022: 2). It has been shown to be the case in Polish that the use of verbs in the generic masculine to address both women and men is possibly cognitively unfavourable to women on the basis of men and women reading sentences with verbs in concordance with their own genders faster than they do when there exists a gender mismatch in the verbs (Szuba, Redl & de Hoop 2022: 832). This is an issue considering that the "feminine marking is never used when addressing men" (Szuba, Redl & de Hoop 2022: 822) and indeed, as feminine terms are sex-specific unlike their masculine equivalents, the use of the feminine as a generic in languages with grammatical gender – such as German – receives constant resistance from some swathes of the population (Cameron 2023: 85-86). As can be seen, gender bias in language and the problems associated with it certainly are not limited to English. It has been noted that many of the most-used languages of the world do generally exhibit gender stereotyping and prejudice (DeFranza, Mishra & Mishra 2020; Lewis & Lupyan 2020), lending credence to the idea that gender bias in language is cross-linguistically widespread and perhaps inherently human. For example, in a study of 25 languages, Lewis and Lupyan (2020: 1025) found that languages which mark occupational or professional terms by gender contribute to stronger implicit career-gender associations by the speakers of the language. In another study of 45 languages using textual corpus data, DeFranza, Mishra and Mishra (2020: 17-18) additionally found that "gender prejudice appears more in gendered rather than genderless languages", as languages with grammatical gender unsurprisingly

highlight gender and its associated elements, with male terms displaying higher positive associations. So, there is reason to view gender bias as a salient and universal social issue.

It is thus rather reassuring to know that much attention has been paid to the issue of gender bias in language in recent years, specifically in the field of computational linguistics. There is indeed a plethora of works on detecting and addressing the problem within the context of natural language processing, particularly focusing on word embeddings encoding societal biases in data that they are trained on (e.g. Bolukbasi et al. 2016; Basta, Costa-jussà & Casas 2019; Sun et al. 2019; Zhao et al. 2019; Zhou et al. 2019; Babaeianjelodar et al. 2020; Lu et al. 2020; Stańczak & Augenstein 2021; Limisiewicz & Mareček 2022).[2] A little aside here: simply put, a word embedding is trained on collocations – words co-occurring with another word – in large corpus data, whereby the words are converted into numbers and represented as multidimensional vectors placed in semantic vector space (Bolukbasi et al. 2016: 4356; Lewis & Lupyan 2020: 1022). The vector differences between two words then represent word relationships and words with similar semantic meaning – often words appearing in similar positions and which could replace each other in a linguistic structure – have vectors which are situated closer together than with other words (Bolukbasi et al. 2016: 4356). Word embeddings have numerous applications in natural language processing: language modelling and machine translation (Lu et al. 2020: 189), identification of advertisement audiences on the web and filtering resumés for job applications (Babaeianjelodar et al. 2020: 752), question answering and sentiment analysis (Bender et al. 2021: 611), as well as natural language generation (Sheng et al. 2019: 3407), among many other tasks. Hence, considering the current widespread use of word embeddings in natural language processing tasks, the intense focus on biases – gender biases, in particular – in these systems is certainly warranted.

The problem identified by researchers is that, since models for the aforementioned processes are trained on human texts, word embeddings would then learn and encode the biases present in those very texts (Zhao et al. 2019: 629; Touileb 2022: 423) with the tendency to amplify the biases (Basta, Costa-jussà & Casas 2019: 34; Bender et al. 2021: 613). In a canonical case of such a bias being reflected by a system, Bolukbasi et al. (2016: 4356) cite the analogy "man is to computer programmer as woman is to homemaker". Another example provided by Limisiewicz and Mareček (2022: 17) is that models generating text will tend to associate

[2] To go over the mechanisms and processes behind these topics in computational linguistics in a detailed manner would be far beyond the scope of this thesis. Thus, the following paragraphs only provide some information relevant to this subchapter.

being a nurse with being female and doctor with being male. In yet another egregious case, testing for biases specifically in language generation systems, Sheng et al. (2019: 3410) found that there often are negative associations of women when accounting for the context of occupation, with prostitution appearing frequently. These biases also have repercussions for machine translations, notably when translating from languages without grammatical gender to gendered languages or those with gendered pronouns. For instance, Prates, Avelar and Lamb (2020: 6378) identified gender asymmetries relative to demographics in Google Translate as it was in 2018, finding that in translating professions from languages without grammatical gender to English, "male defaults are not only prominent, but exaggerated in fields suggested to be troubled with gender stereotypes" such as in the sciences, and certain adjectives such as *shy* tend to be translated with female pronouns while words like *guilty* and *cruel* "almost exclusively translated with male ones" on the basis of 12 languages. Similar findings have been noted translating from Korean to English by Cho et al. (2019: 177), words with negative sentiment more often being associated with men. Observations have also been made for translating from Filipino to English by Endriga and Rosario Jr. (2023: 95-96), with most verbs, occupational terms and adjectives being translated by default with male pronouns and noting that the pronouns default to the male variant even when context is provided in previous sentences.

That these biases and stereotypes have harmful and detrimental consequences – what with all the applications of language models today – are well-known (Bender et al. 2021: 617; Stańczak & Augenstein 2021: 5-6). To address the issue, researchers have constructed different formulas to firstly measure and evaluate bias (Zhou et al. 2019: 5278; Babaeianjelodar et al. 2020: 755; Lu et al. 2020: 194; Stańczak & Augenstein 2021: 13-14; Limisiewicz & Mareček 2022: 19) and developed methods to mitigate bias in the systems (see Sun et al. 2019; Stańczak & Augenstein 2021), beginning with Bolukbasi et al. (2016: 4363) devising debiasing algorithms which were found to be effective in reducing gender bias in word embeddings. Other methods followed, with bias mitigation methods suggested for contextualised word embeddings (Zhao et al. 2019: 632), for languages with grammatical gender (Zhou et al. 2019: 5279), for language models without losing actual gender information (Limisiewicz & Mareček 2022: 21-22), and further approaches involving debiasing the datasets the models are trained on or adjusting the algorithms for natural language processing (see Sun et al. 2019: 1634, 1636). However, despite all the extensive effort exerted into bias mitigation approaches, methods and strategies, Stańczak and

Augenstein (2021: 2) still note that existing research is fraught with limitations, with varying methodologies. For example, Touileb (2022: 423) points out that template-based bias probes for language models seem to have difficulty accounting for negation and can be especially sensitive to relatively minor word changes, including changes to grammatical tense, recommending that grammatical tense be taken into account for future probes (Touileb 2022: 427). Nevertheless, increasing awareness of the dangers posed by machines and systems perpetuating biases and constant endeavours to reduce biases in said systems – for natural language processing and language models are ubiquitous in their utilisation today – demonstrate the importance of detecting gender bias in society and in language.

Now, as fascinating and useful as they are, the exploratory research for this thesis does not employ word embeddings, although studies such as by Garg et al. (2018) do show that diachronic studies of bias and stereotyping are possible using the methodology. Instead, to assess gender bias, this research takes what Ritt et al. (2024: 3) refer to as the "traditional" corpus linguistics approach, deriving observations from the collocational strengths of the gendered third-person pronouns in English with select verb constructions.

2.3. Corpus linguistics, collocations and collostructions

Lindquist and Levin (2018: 1) state that, unlike other branches of linguistics where the word preceding the word *linguistics* indicates what the subject of linguistic study is, corpus linguistics is primarily a methodology – a means with which to study language rather than the study itself. Specifically, as the name suggests, corpus linguistics "refers to a body of language" comprising of texts serving as "authentic" samples of language reflecting language use as grounded in reality, i.e. how the language actually is used by the population (Baker 2014: 7). Thus, Baker (2014: 7) highlights similarities between corpus linguistics and the branches of social sciences, describing corpus linguistics as "an approach that is grounded in empiricism". In spite of criticisms from generative linguists – Noam Chomsky being one such notable example – which meant that "many early corpus linguists almost felt as if they had to work in secret cells" (Lindquist & Levin 2018: 9), corpus linguistics endures as a valid methodology for research. Indeed, since the 1960s – when the first electronic corpus was compiled using written texts – the number of corpora has increased dramatically both in their quantity as well as in the collections of texts comprising the corpora (Egbert, Larsson & Biber 2020: 4). At the same time, although Egbert, Larsson and Biber (2020: 4) do state that "all other things being equal, a bigger corpus is preferable", considering the longstanding aim

among linguists to strive for substantially large corpora, they also draw attention to the issue of representativeness.

To begin, this is not to say that corpora sizes do not matter. Rather, Egbert, Larsson and Biber (2020: 14-15) are explicit in stating that the focus on the purported representativeness of corpora should not detract from the fact that the size of the corpora ultimately does matter once the issue of representativeness has been dealt with as large samples are still favoured. The fact remains, however, that representativeness has been a recurrent point of contention and discussion in the field of corpus linguistics. As infinitely many sentence combinations are possible in a language, it is accepted that no corpus can ever attain the size that could contain the entirety of realisations of a particular language (Lindquist & Levin 2018: 10). Yet, as Baker (2014: 9) points out, it is nevertheless possible for some corpora to "claim total representativeness of a genre while others cannot", citing an example of a hypothetical corpus of an authors' cumulative written output over a given timespan. Egbert, Larsson and Biber (2020: 14) also emphasise that there are necessary trade-offs to be made when taking the representativeness of corpora into consideration against practicality and feasibility, noting that "[t]he ideal of representativeness may be more or less easy to achieve depending on the subfield of corpus linguistics", in line with what Baker (2014: 9) mentions. As Lindquist and Levin (2018: 43) perspicaciously put it, "representativity is relative: it varies between corpora and is never absolute"; obviously, for example, specialised corpora lay claim to different levels of representativeness in contrast to general corpora.

As already mentioned, there are different types of corpora. In addition to specialised and general corpora, there also exist other groupings such as written, spoken and historical or diachronic corpora (Lindquist & Levin 2018: 11). As the name suggests, specialised corpora have a focus on providing language samples from a restricted set of data, such as only looking at academic or learners' English (Lindquist & Levin 2018: 18). In contrast, general corpora seek to encompass both written and spoken language data and aim to be as balanced and representative as possible of the language variety or type from which the data are collected (Lindquist & Levin 2018: 14). As for historical or diachronic corpora, they can either be of the type which is limited to looking at an older form of the language or alternatively, one which enables diachronic studies from the present-day until a certain point in the past, facilitating comparisons between language stages and allowing for historical trends to be tracked (Lindquist & Levin 2018: 19). It is thus possible to take both synchronic

– describing language as it appears or appeared at a certain point in time – and diachronic approaches – exploring how language has changed over time – in studying language with corpora (Lindquist & Levin 2018: 174). This thesis primarily takes the diachronic approach in investigating how the trends and trajectories of gender bias appear based on corpus data, utilising the Google Books (American English) Corpus (Davies 2011-).

That being said, it is also important to take a cautious approach when undertaking diachronic studies that connect society and language utilising large corpora. Hilpert (2020: 6) outlines primary issues which must be taken into consideration when analysing diachronic data from corpora, of which several are noteworthy: that corpus frequencies do not necessarily reflect real-world frequencies, that corpus frequencies should be examined according to specific sense and construction where polysemous words are concerned, and that social changes are fairly difficult to decouple from linguistic changes when scrutinising linguistic records. To the first point, Hilpert (2020: 8) makes a distinction between semasiological and onomasiological frequencies; the former refers to frequencies in texts reflecting how often a concept is verbalised whereas the latter concerns "how often speakers experience a given concept" – in other words, actions committed on a frequent basis in real life may not necessarily be reflected in textual form as used in corpora. The third issue is somewhat related to the first, as it could be the case that linguistic changes occur independently of changes in the social sphere rather than language functioning as a reflector of societal trends with several possible factors in play (Hilpert 2020: 12-13). The second concern is of relevance as there is an abundance of words in English which happen to be polysemes, i.e. words with the same written and pronounced form which carry related but ultimately different meanings; Hilpert (2020: 9) supplies an example with the verb *get*, which has a variety of meanings and constructed forms depending on context and use. For these reasons and other similar known problems, Hilpert (2020: 25) remarks that – as tempting as it is to make use of large corpora to track social change – there are numerous factors which may complicate matters owing to the complex relationship between language change and social change that need to be addressed to conduct proper analyses. Indeed, Hilpert's (2020) own research shows that, taking precautions and all things considered, diachronic studies are both reasonable and valid.

In conjunction with either synchronic or diachronic analyses of linguistic structures, corpora are useful tools which crucially allow investigations involving collocations. Collocations

essentially describe the relationship of the degree of co-occurrence between a word and other words (Lindquist & Levin 2018: 58) and are important as there is a "tendency in speakers to use groups of words together as multi-word units or phrases" (Lindquist & Levin 2018: 72). In examining collocations and collocates, attention is paid to either the words in the vicinity of a given word or words which immediately precede or follow a given word (Lindquist & Levin 2018: 74). As such, collocations are related to frequencies and are "useful in revealing how meaning is acquired through repeated uses of language" (Baker 2014: 13). Consequently, collocations are especially important for exploring bias in language; considering that the company of words surrounding a word accords it with certain associations positive, negative and ostensibly neutral, it has been suggested that even if the word itself may often be interpreted neutrally, biased collocations may very well reproduce implicit bias in cognition (Hauser & Schwarz 2022: 10). As Egbert, Larsson and Biber (2020: 26) explain, however, it does not suffice to simply look at the absolute frequencies of any given collocate of a word – rather, statistical methods are used to identify collocations as actual valid examples of collocations. Therefore, in collocational analyses, the primary focus is on calculating and measuring the strength of association – the collocational strength between words (Lindquist & Levin 2018: 88). Indeed, Schmid and Küchenhoff (2013: 532) state that one of the goals of corpus linguistics has long been "to measure the degree of mutual attraction between lexical elements in text" and this, for the most part, appears to remain the case.

An extension of collocational analysis is provided by Stefanowitsch and Gries (2003) in the form of collostructional analysis. The underlying principle is similar to that of analysing collocations with the notable difference of investigating not merely the relationship between words, but rather the "interaction of lexemes and the grammatical structures associated with them" (Stefanowitsch & Gries 2003: 209), treating both lexemes and structures – words and associated grammatical constructions – as independent analysable units for which co-occurrence patterns can be identified (Stefanowitsch & Gries 2003: 213). Rather than looking at mere raw frequencies of immediate collocates which tend to overlook grammatical structures, the key feature of collostructional analysis involves examining which lexemes are attracted to or repelled by being in a particular position in a given construction (Stefanowitsch & Gries 2003: 214). Among several statistical means to measure the degree of attraction or repulsion, Stefanowitsch and Gries (2003: 218) utilise Fisher's exact test despite the computational intensity and expense of the measure as "[i]t neither makes any distributional assumptions, nor does it require any particular sample size", making it

especially amenable to use in cases of infrequent attestations of a collostruction in a corpus,[3] although other alternative measures – such as odds ratio – do exist for similar purposes (Schmid & Küchenhoff 2013: 548). In addition to demonstrating case studies of their methodology applied for investigations into idioms and complementation patterns, Stefanowitsch and Gries (2003: 230-235) also show that collostructional analyses are possible for more abstract structures such as verb constructions accounting for tense, aspect and modality. Through their analysis, for instance, the researchers were able to confirm that stative verbs – mainly those to do with cognition – are particularly not attracted to being used in the progressive (Stefanowitsch & Gries 2003: 230) and additionally found that certain verbs seem to appear in corpus data as showing more affinity or less towards the past tense (Stefanowitsch & Gries 2003: 235). These examples demonstrate that there is tremendous wealth in analysing word affinities particularly when grammatical constructions are also taken into special consideration.

Technically, collostructional analyses form the overarching methodological foundation of the research for this thesis as the degrees of attraction of the gendered pronouns *he* and *she* towards various verb constructions are examined to ascertain the level of gender bias in language. However, there are some deviations. As the research here is of the exploratory type, and while utilising statistical measures indeed applicable for collostructional analysis, it is perhaps safer to say that collocational, and not strictly collostructional, strengths are being measured, bearing in mind that the main focus is on the degree of attraction of the pronouns immediately preceding the constructions treated as chunks without the surrounding context of the attestations taken into account. Furthermore, referring to collocational strengths is also in line with the research by Ritt et al. (2024), who also write of collocational strengths and affinities for noun-construction relationships. For these reasons, the thesis uses the terms collocational strengths or affinities to refer to the relationship between the pronouns and the verb constructions under investigation.

2.4. Tense, aspect, modality and negation

As verb constructions are of special interest to this thesis, it is worth reviewing the systems of tense, aspect and modality as well as verbal negation. Only some selected facets of the grammatical categories deemed salient are included here since an exhaustive description of the verbal systems is beyond the purview of this paper.

[3] See Stefanowitsch and Gries (2003: 218-219) for a description and example of the calculation methodology.

Where the categories of tense, aspect and modality are concerned, Huddleston (2002: 116) makes a distinction between four verbal systems: primary tense, secondary tense, aspect and mood. These systems can themselves be respectively subdivided into preterite and present, perfect and non-perfect, progressive and non-progressive, and modal and non-modal types (Huddleston 2002: 116). Additionally, Huddleston (2002: 118) distinguishes between the forms – tense, aspect and mood – and the meanings conveyed by these forms – time, aspectuality and modality. These forms provide information about situations, which Radden and Dirven (2007: 171) describe as "events that happen and states things are in" which also possess temporal structure (Radden & Dirven 2007: 176) and are thus crucial to communication. More specifically, the information specified by the different forms also determines when the situation took place, what view the speaker takes of a given situation, and the factuality or reality of a situation (Radden & Dirven 2007: 171). As with the verbal systems, both Huddleston (2002: 118) and Radden and Dirven (2007: 176) classify different types of situations; a first distinction can be made between states and events – the term used by Radden and Dirven (2007) – or occurrences – used by Huddleston (2002) – and then there are achievements, accomplishments and activities as subtypes of events. The verbal grammatical categories in combination are then used accordingly for each situation type, although overlaps may occur.

As has been said before, tense indicates time. Specifically, the primary tense marks time that is referred to usually from the viewpoint of the time of orientation, the default case of the deictic time, which itself can be understood as the "time of utterance" in normal cases (Huddleston 2002: 125). Similarly, Radden and Dirven (2007: 204) write that deictic time is related to speech time, remarking that "speech time always establishes the present time" while pointing out that event time – the time being referred to – may also "be earlier or later than speech time". Thus, the three possible times that can be referred to from the time of orientation correspond to the canonical past, present and future simple tenses (Radden & Dirven 2007: 204). Huddleston (2002) deviates slightly in interpreting the tenses; the tenses indeed serve the purpose of establishing relations between the time referred to and the time of orientation (Huddleston 2002: 127), but Huddleston (2002: 208) does not recognise the future tense in English. Rather, he states that although the future time can be indicated through various means, "there is no grammatical category that can properly be analysed as a future tense", instead subsuming the auxiliaries *will* and *shall* under expressions of mood

(Huddleston 2002: 209) – hence the aforementioned categorisation of the primary tense into only two types: preterite and present.

The simple present tense has a multitude of uses, for both static and dynamic situations – i.e. states and events – whereby the meaning conveyed must be distinguished from each other despite the same form employed in both settings (Huddleston 2002: 127). For events, the simple present tense usually faces constraints – Huddleston (2002: 128) cites some main cases in which the tense is nonetheless used: as performatives and to provide running, demonstrative or informal commentaries. The present tense is also noted for its "timeless" use in written works or when describing such works, such as synopses, captions and chronicles which accord events with a certain level of permanence (Huddleston 2002: 129-130; Radden & Dirven 2007: 222), allowing its utilisation for everlasting and indefinitely lasting states as well (Radden & Dirven 2007: 210). Most often, situations construed in the present tense include accomplishing activities – additionally in the progressive – and also, quite prominently, regular habitual states, which are in the simple present (Radden & Dirven 2007: 210). Huddleston (2002: 130-131) also notes that the present tense can be extended to also cover past and future situations, an example of the former being in narratives and the latter for e.g. cyclic natural events. The use of the present in narratives has likewise been pointed out by Fludernik (2012: 83-84), highlighting the increase in use of the potentially jarring present tense – distinguished from the "historical present tense" – as the main tense as opposed to the preterite in fictional narratives since the 19th century.

The preterite or the past tense is used to indicate or refer to events occurring before the moment of speech or the time of orientation (Huddleston 2002: 137). Unlike the case for the present tense, Huddleston (2002: 137) mentions that the past tense does not have constraints on its use for dynamic situations – it is employed for both states and events regardless whether the situation is finished or ongoing, as long as the situation being referred to has happened in the past. Indeed, Radden and Dirven (2007: 218-219) draw attention to the properties of the past tense in that it is firstly focused on the past time, i.e. not immediate reality, and that it is additionally characterised by "its detachment from the present". Unsurprisingly, the preterite often sees use in narration and especially fictional narratives as it allows the text to be uprooted from reality, bestowing upon narratives their characteristic quality of not being "anchored in the here and now of the author and the reader" (Fludernik 2012: 84). In addition to its use in narratives, the distality expressed by the past tense makes

it appropriate to signify politeness as a marker of diffidence; Huddleston (2002: 138) writes that the distancing effect produced by the past tense helps "avoid the risk of appearing too direct, possibly brusque". A further role of the preterite has been recognised in the form of backshifting, e.g. in cases of indirect reported speech, which can be considered a special use of the past tense that is not simply akin to converting the original tense from present to past (Huddleston 2002: 151-152).

Although there is no inflected future tense in English, such that some modern grammars such as by Huddleston (2002: 208) do not recognise it, this thesis nevertheless treats the *will*-future as a tense for convenience. Huddleston (2002: 209) does make the argument that *will* falls into the category of modal auxiliary verbs both grammatically and semantically, warranting its analysis together with mood rather than tense. Nonetheless, *will* still may be used to express futurity, itself falling within the realm of modality (Huddleston 2002: 188). Radden and Dirven (2007: 225), while acknowledging that *will* and *shall* additionally have the function of modals and thus may convey different meanings as opposed to the common view of their use as future tense markers, state that the original meaning of the two future tense forms – that of "volition and insistence", respectively – have an added layer of predictive implication on the part of the hearer. They argue that this connection is conventionalised to the extent that the verbs are now used neutrally as indicating the predictive future with the semantic element of volition having disappeared (Radden & Dirven 2007: 225). As is known, there also exists another future form: *be going*. Huddleston (2002: 211-212) notes some distinctions between the two forms, such as that *be going* frequently invokes a sense of immediacy which is not readily available with *will*, or that the *be going* future also indicates the aspect of intentionality or future arrangement. Radden and Dirven (2007: 225) likewise remark on the contingent nature of *be going* as opposed to *will*, writing that the *be going* future is more evidence-based and indicates a higher level of confidence.

As for the secondary tense, there is the perfect which is formed periphrastically in English through the use of the auxiliary *have*. Huddleston (2002: 116) points out in a footnote that the perfect is occasionally viewed as an aspectual feature rather than a tense; indeed, such a perspective is taken by some scholars who assert that there, at the very least, exists a state of overlap in the case of the perfect (Binnick 2006: 251; Michaelis 2006: 223). This may likely be due to the conflation of the perfect and the perfective, the latter being an unambiguous marker of aspectuality. Huddleston (2002: 124), elucidating the difference between the

perfective and imperfective aspectualities which is predicated on whether a situation is presented in its totality or not, notes that the perfective and the perfect do not necessarily share the same sense. Effectively, the perfect is a realisation of the past tense along with the preterite (Huddleston 2002: 139). The perfect, in contrast to the preterite, is normally a non-deictic past tense, i.e. the time referred to is not placed before the immediate present or now (Huddleston 2002: 140). Rather, as Radden and Dirven (2007: 207) put it, the time of the event precedes both the time of reference and time of speech in the complex tenses of the present and past perfect. For the present perfect, Huddleston (2002: 143) lists several established types for various uses: continuative, experiential, resultative and the recent past. The gravamen is that the present perfect is often applied in cases of ways a past situation possess "current relevance" (Huddleston 2002: 143), a point included by Radden and Dirven (2007: 212) in characterising the present perfect alongside other properties such as the focus on the present and indefiniteness. The past or preterite perfect is then doubly in the past and so allows "doubly remote" interpretations of a situation (Huddleston 2002: 150), enabled by retrospectively viewing situations taking place further in the past from an already past perspective (Radden & Dirven 2007: 222). Through these means, the chronology of a narrative is made clear; events placed in the past perfect "typically relate to the narrative sequence of events" (Radden & Dirven 2007: 222). Fludernik (2012: 88) alludes to this deployment of the past perfect which "is used to signal […] a shift into the previous history", adding that by using the past perfect, the narrative shifts into the past "and that shift correlates with a linking between two plotlines"; thus, the past perfect in narrative can also signal flashbacks (Fludernik 2012: 95). The perfect tense as a whole can so be seen as a tool for organising and sequencing events having begun in the past from a current point of view.

In addition to tense, aspect is another salient grammatical feature with regard to verbs. Radden and Dirven (2007: 175) define aspect as the "grammatical form used by a speaker in taking a particular view of a situation". Huddleston (2002: 117) also foregrounds the situation in defining aspect as a "system where the basic meanings have to do with the internal temporal constituency of the situation". Both definitions point to aspect as being a feature with respect to the viewpoint of the speaker, whether an internal or external view of the situation is being taken. Radden and Dirven (2007: 175) specifically use the term "viewing frame" to refer to how a speaker views the events. With maximal viewing frames, events are viewed "externally and in their entirety" and are temporally bounded with an emphasis on the endpoint of an event whereas events are seen "internally and provide a close-up view on the

progression of the event" with restricted viewing frames, being "unbounded events" with implicit boundaries (Radden & Dirven 2007: 178). States, too, can be viewed with viewing frames; temporary states take a restricted view whereas lasting states are seen with a maximal viewing frame (Radden & Dirven 2007: 179). Therefore, as Radden and Dirven (2007: 175) explain, bounded events and lasting states are expressed in the non-progressive aspect while unbounded events and temporary states are in the progressive. Both Huddleston (2002: 124) and Radden and Dirven (2007: 178) mention the perfective and imperfective aspectualities in their writing. Radden and Dirven (2007: 178) specify that bounded events may also be referred to as perfective and unbounded events as imperfective, and Huddleston (2002: 124) similarly notes that the perfective aspectuality externally presents the situation in its totality while the imperfective is not seen as a whole but "viewed from within". Crucially, Huddleston (2002: 124) points out that unlike some other languages, English does not make the perfective-imperfective distinction in verbs, both the present and preterite being employed either way with regard to perfectiveness. The one instance of aspectuality which does exist in English is the progressive, which Huddleston (2002: 124) describes as "a special case of imperfectivity […] where a dynamic situation is presented as ongoing, in progress".

English is somewhat of an oddity among European languages in that a specific progressive aspect arose over its history (Radden & Dirven 2007: 5). As a type of imperfective, the progressive aspect – formally indicated by the auxiliary *be* preceding a "gerund-participle" ending in *-ing* (Huddleston 2002: 162) – variously marks situations in progress, dynamic and in limited duration (Huddleston 2002: 163). Among its diverse uses, the progressive can also allow the interpretation and common reading of a situation as a single occurrence (Huddleston 2002: 164) and "change a punctual situation into a durative one" (Huddleston 2002: 166). Duration is an especially salient aspect of the progressive, such that the progressive aspectuality may be applied for serial states as well, in which case the situation is "accompanied by an emotive overtone" (Huddleston 2002: 166). The focus on duration means that the progressive aspect can figuratively stretch bounded events into unbounded activities (see Radden & Dirven 2007: 184); for example, accomplishments – bounded events which normally have an endpoint – may be turned into accomplishing activities (Radden & Dirven 2007: 184). Another type of bounded event, achievements, can be extended from being punctual events to being unbounded culminating activities by focusing on not the end result but rather the phase leading to the culmination of an action through the use of the progressive (Radden & Dirven 2007: 188). Huddleston (2002: 167) also highlights that the

progressive cannot normally be used for "purely static situations" as it expresses dynamicity, which contrasts the progressive aspectuality with the imperfective since imperfective states do exist. At the same time, it has been noted that static situations can be made to convey dynamic expression through the progressive where agentive activity or the temporary state of the situation are particularly emphasised (Huddleston 2002: 167). Furthermore, it is known that there are certain verbs which are ambiguous as to their usage in their progressive forms, whereby the progressive suggests an activity and the non-progressive usually expresses a state; these, as Huddleston (2002: 169) lists, are primarily verbs of feeling, of hurting and of cognition, emotion or attitude, as well as verbs denoting stance or position. All in all, it is understood that situations – whether they are various types of states or events – are central to the understanding of aspect and aspectuality.

Mood and modality are a different matter from tense and aspect as marked moods denote kinds of "non-factuality or non-actuality" (Huddleston 2002: 117) and mainly deal with potential reality, detaching them from situations (Radden & Dirven 2007: 173). There are two major types of modality which are especially important: epistemic modality, which signals possibility and potentiality, and deontic modality, which marks direction and obligation (Huddleston 2002: 178; Radden & Dirven 2007: 233). To convey modality, there exist a set of modal verbs in the form of auxiliaries (Huddleston 2002: 92) which are set apart from lexical verbs by, among other distinguishing features, lacking non-finite forms (Radden & Dirven 2007: 242) and possessing separate negative forms (Huddleston 2002: 75): *can*, *may*, *will*, *shall*, *must*, *ought*, *need*, and *dare*. Although modality expresses the speaker's attitude vis-à-vis the "factuality or actualisation of the situation" for the most part (Huddleston 2002: 173), it may also encompass other interpretations, in addition to the aforementioned concepts of possibility and necessity, such as objectivity in some cases (Huddleston 2002: 173). Now, it must be stated that considering that Huddleston (2002: 172) himself mentions that mood as an inflected grammatical category has been largely extinguished from English and markers for mood are provided for by the modal auxiliaries, this thesis – for convenience – uses the term *modality* to cover both the category of form and meaning in lieu of *mood* for the former. Furthermore, only the modal verbs are of relevance here; modality may indeed be conveyed through other means such as certain adjectives, adverbs and lexical verbs (Huddleston 2002: 173), but their use as indicating modality is not covered in this thesis.

Put succinctly, different modal verbs obviously express different types and dimensions of modality. There is a distinction to be made between strong and weak modality which correspond to the expressions of necessity and possibility (Huddleston 2002: 175). Among the modal verbs which convey necessity and possibility are *must* and *need* for the former concept and *may* and *can* for the latter (Huddleston 2002: 180). Of those verbs, *must* and *may* express epistemic necessity and possibility, respectively (Huddleston 2002: 180), and Huddleston (2002: 181) writes that they are "most often used subjectively, to express the speaker's judgement". As expressions of deontic modality, *must* conveys the meaning of strong obligation while *may* indicates permission, which can likewise be expressed by *can* in informal settings (Huddleston 2002: 182). The modal *can* has a variety of uses where possibility is concerned, such as its use to indicate "what is reasonable or acceptable", "what is circumstantially possible", "what is sometimes the case" (Huddleston 2002: 184) as well as the ability on the part of subject to perform a particular action (Huddleston 2002: 185).

The preterite forms of *can* and *may* – *could* and *might*, respectively – and indeed all the preterite forms of the modal auxiliaries, including *would* and *should*, generally function the same way as other verbs in the past tense would, such as marking past time (Huddleston 2002: 196) and to backshift (Huddleston 2002: 198), but also are used to signal remoteness (Huddleston 2002: 198) with which "tentativeness, diffidence [and] extra politeness" can be expressed (Huddleston 2002: 200), owing to the distancing effect mentioned earlier for the preterite in general. The form *would* in particular is interesting as it stems from *will*, expressing futurity; as such, *would* may also be used to convey the same, which is indeed the case in narratives and similar genres (Huddleston 2002: 198). Since *will*, in addition to marking the future, can also signal volition, i.e. willingness or the lack thereof, as well as intention (Huddleston 2002: 192), and propensity, i.e. the "characteristic or habitual behaviour of animates or general properties of inanimates" (Huddleston 2002: 194), *would* also carries those associated meanings (Huddleston 2002: 197). The case of *should* is also of interest; Radden and Dirven (2007: 252) note that *should* has inherited the "original sense of indebtedness" present in *shall* and that its use today denotes weak obligations that foreground the "speaker's idea of desirability about an advisable course of action". Huddleston (2002: 186) shares the view that *should* is used principally in the deontic sense, although epistemic uses of *should* are certainly possible in certain instances (Huddleston 2002: 187; Radden & Dirven 2007: 253). As can be seen, modals thus generally convey a variety of meanings – those of necessity, possibility, potentiality or aspects of non-factuality and non-actuality –

and provide information about the speaker's attitude, the possible state of objects, or situations in actual or hypothetical realities.

The paragraphs above have hitherto been concerned with the grammatical categories of tense, aspect and modality. Negation, specifically verbal negation, is a separate topic which is still relevant to verb constructions. In simplest terms, negation, as canonically marked by the negator *not*, exists to help "imagine the positive counterpart of a situation" (Radden & Dirven 2007: 235). Negative clauses, i.e. clauses with negative polarity, are considered marked in a language with positive constructions considered the default as they are deemed "structurally and semantically simpler than negative ones" (Pullum & Huddleston 2002: 786). With regard to negation, the focus of this thesis is on verbal negation, where the negation is "grammatically associated with the verb, the head of the clause" (Pullum & Huddleston 2002: 788), and even more specifically on primary verbal negation.

Simply, in primary verbal negation, positive clauses are turned negative through one of two ways: analytically or synthetically (Pullum & Huddleston 2002: 799). The difference is that analytic primary negation is realised through the addition of *not* after the auxiliary – either after the supportive *do* in clauses containing lexical verbs or immediately after the auxiliary verb in clauses already with auxiliary verbs – while synthetic primary negation is effectuated by inflecting the auxiliary to its negative form which exist separately from the positive primary forms (Pullum & Huddleston 2002: 799). It must be underscored here that it is not a mere contraction of the negator *not* which occurs in the case of synthetic primary negation but rather actual inflection. Indeed, it has already been demonstrated fairly early on by Zwicky and Pullum (1983: 510) that the *-n't* contraction is not a mere clitic – as in the case of *-'ve*, for example – but that it is an inflectional suffix. This is generally the case for auxiliaries,[4] of which *do* is one – when a lexical verb is negated, the supportive *do* takes on the inflection of the lexical verb, including the negation which can be realised either analytically or synthetically like the other auxiliaries (Pullum & Huddleston 2002: 800). As Pullum and Huddleston (2002: 800) also point out, the main difference between the analytic and synthetic forms is to do with style – inflected negation is viewed as a marker of informality. At the same time, Pullum and Huddleston (2002: 801) also do mention that there are cases in which the analytic and synthetic forms are not interchangeable, such as in cases

[4] Pullum and Huddleston (2002: 800) suggest that *mayn't* is not deemed "fully acceptable". Moreover, negative forms of *be* and *have* are also possible.

of inversion, and also state that analytic negations can sound both unusual and unnatural in conversations (Pullum & Huddleston 2002: 800).

It additionally appears that modals and negation have a fascinating relationship, something alluded to by Radden and Dirven (2007: 258) as well. For instance, they note that generally, deontic modality negated has an obligatory – or rather, prohibiting – lean whereas negation coupled with epistemic modality primarily concerns possibility (Radden & Dirven 2007: 260); in other words, the reading in the former case is that there is an obligation not to do something (Radden & Dirven 2007: 260), and in the latter instance, the impossibility of something being the case is what is implied (Radden & Dirven 2007: 261). This distinction for modals and negation is technically that of internal and external negation, i.e. whether "the modal has scope over the negative" or whether "the negative has scope over the modal", respectively (Huddleston 2002: 175). In non-modal instances, this is fairly straightforward: the negation has scope over the immediate verb and the clause at large, provided there is no coordination (Pullum & Huddleston 2002: 791). For modals, the distinction means that the negation may apply to the verb following the modal, in which case it is internal or propositional negation, and if the negation applies to the modal, then it would be external or – co-terminously – modal negation (Huddleston 2002: 176). Huddleston (2002: 204) also notes that "[t]here is a general tendency for strong or medium modals to take internal negation, and for weak modals to take external negation". Hence, the negation of modal verbs like *should*, *must* and *may* commonly leads to an internal negation reading (Huddleston 2002: 180, 187) while *need* and *can* negated, i.e. *it is not necessary or possible that*, are generally interpreted as external negations, with negations of *must* and *need* additionally understood as examples of deontic modality (Huddleston 2002: 180-181).

It is clear that the inner mechanisms and underlying principles of verb constructions are diverse and complex. There are many factors to take into account when investigating the use of tense, aspect and modality, and exceptions and anomalies abound. Precisely for this reason, the main research for this thesis is primarily concerned with the form and not necessarily the content or context of the verb constructions. Nevertheless, it is still possible to hypothesise, speculate and interpret with the salient information presented above.

2.5. Research motivation and interest

The preceding subchapters have provided information about the necessary underpinnings for the thesis. To reiterate, the primary purpose of this exploratory study is to determine and quantify the collocational strength of *he* and *she* with select verbs and their constructions diachronically based on statistical computations using corpus data with a focus on the construction form. The idea is that collocational affinity in this context can be a measure of gender bias in language, specifically language about men and women in published material in American English. In addition to changing trends, the underlying assumption is that while lexical verbs may themselves indicate levels of gender bias, i.e. some verbs collocate more regularly with *he* and others with *she*, there should be nothing inherent in verb constructions for gender bias to be displayed through them. There are reasons to suspect that this is worth exploring. In a study on probing language models in Norwegian and English for bias solely through gender correlations with occupational terms, Touileb (2022: 426) found that even minimally altering bias probes with regard to grammatical tense and negation, e.g. changing *he*/*she works as* [job] to *worked as*, *will work as* or *did not work as*, etc., introduces shifts in correlations towards male or female pronouns when the tense of the probe is changed from the present to the past and future, respectively. A wide-ranging corpus research by Ritt et al. (2024: 25) on gender bias discovered, among other findings, that there is greater relative visibility for *she* preceding the past and past perfect tense than before present and present perfect tense forms.

The motivation for this study thus stems from the fact that although gender differences and gender bias in language have been studied quite extensively, there exists a significant research gap where gender and verb constructions are concerned despite the availability of numerous research methods and means. In fact, this thesis can be seen as partially responding to calls for further research by Touileb (2022: 427), who recommends investigating "diachronic gender-based bias correlations with occupations" as biases and stereotypes change over time, and Ritt et al. (2024: 29), who note that "gender bias in different constructions" has hitherto not been adequately pursued. Rather than focusing solely on gender bias correlations with occupations as Touileb (2022) suggests, this exploratory study expands on investigating the role of the other factor – grammatical tense, and by extension, aspect, modality and negation as well – on gender bias correlations in different verbs and verb constructions. Therefore, this study investigates whether gender bias in constructions in written language is quantifiable and to what extent, in which constructions gender bias is

more prominent, which verbs among the ones selected for analysis exhibit high levels of bias, and how trends have changed over two centuries, among other such questions. This thesis so aims to be a welcome addition to the literature on the topic of gender bias in language.

3. Review of previous research

With the understanding of *corpus* in linguistics to very roughly mean any collection of language data, corpora have been used to study gender in many different ways, with a wide variety of research works to that end. The study topics range from the fairly obscure, such as investigating the phenomenon of the inanimate use of female pronouns for weather terms which suggests the existence of the "emotional value of the feminine pronoun" (Gardelle 2015: 353), to the more general, such as examining the role of gender in language change and development of Standard English on the basis of early English correspondence (Nevalainen 2000). The wide-reaching potential of corpus linguistics in studying gender is evident.

More specifically, gender differences, bias and stereotyping have been popular subjects of study in the field of corpus linguistics, utilising both spoken and written language data; the language datasets are as varied as the field, with there being research employing drama texts (Biber & Burges 2000), correspondence (Nevalainen 2000), searches on the web (Gardelle 2015), literary works (e.g. Jockers & Kirilloff 2016; Čermáková & Fárová 2017), transcripts of public government bioethics commission meetings (Schmauss & Kilian 2023), and even fairy tales (Toro Isaza et al. 2023). There is therefore an extensive body of research which examine gender phenomena through various means using many distinct corpora, such that, for example, there also is a monograph by Baker (2014) which precisely focuses on investigating gender and sexuality through various methods in corpus linguistics and presents several case studies to that effect. There are other general works which complementarily dedicate separate sections or chapters to analysing gender with corpora. For instance, Curzan (2014: 114-136), writing about gender-inclusive language in the wider context of prescriptivism in English, additionally conducts simple corpus research in a style reminiscent of Baker's (2010) investigation on the frequency of gendered terms over time in British English to examine historical frequency trends of inclusive pronoun forms – *he or she* and its variants – and gender-neutral occupation terms, e.g. *police officer*, *firefighter* and *flight attendant*, to see what effect style guides have had on their use in American English. In another case, Lindquist & Levin (2018: 156-173) incorporate a chapter in their work on corpus linguistics in general to study gender relations using corpus data, testing Lakoff's

(1973) observations on language use – corroborating that women use more colour terms but having inconclusive findings on the use of hedges – as well as examining, among others, how men and women are described and referred to in corpora with a focus on adjectives and verbs in the company of *man* and *woman*. Here, the distinction can be drawn between investigating gender differences in language used by men and women – i.e. male and female language use – and language used about men and women, i.e. how men and women are portrayed.

Most corpus linguistics research on gender differences or variation thus broadly fall into two categories: those exploring gender-based active language use differences (e.g. Biber & Burges 2000; Nevalainen 2000; Newman et al. 2008; Berglind Söderqvist 2017; Leone 2022; Schmauss & Kilian 2023), or those investigating how gendered nouns and pronouns are described and represented in the data (e.g. Sigley & Holmes 2002; Pearce 2008; Baker 2010; Norberg 2016; Piits 2016; Ye et al. 2018; Raphael 2023). Occasionally, research incorporates both elements: taking into account both the gender of the language user and seeing how they then represent men and women (e.g. Jockers & Kirilloff 2016; Čermáková & Fárová 2017). Although there is absolutely no limitation, it is usually the case for the former that spoken language data are especially well-suited considering that spoken language is deemed superior and more representative of actual language use than written language is (see Linell 2005: 27). Research of the latter type concerned with how men and women are described in language also makes use of both spoken and written language data, whereby the interest is on collocations – looking at which collocates appear in the vicinity of gendered words.

Research on collocates is likewise varied. In particular, adjectival and verbal collocates of gendered nouns have been of interest in many cases. Sigley and Holmes (2002), for example, explore how *girl(s)* and *boy(s)* are used and represented in corpora of American, British and New Zealand English, finding that *girl* occurs variably as an object of desire or being emotionally weak or vulnerable, in addition to appearing in contexts establishing traditional roles, regardless of the age of the referent (Sigley & Holmes 2002: 147-148). The researchers also found that adjectival collocates for *girls* tend to highlight their appearance more than those for *boys* and that *girls* get married or teased (Sigley & Holmes 2002: 150, 152), concluding that there is "continued sexist usage of *girl* for adult women", with *girl* "more than three times more likely than *boy* to refer to an adult" (Sigley & Holmes 2002: 153).

In another study examining the collocational behaviours of MAN and WOMAN as lemmata in British English, Pearce (2008) offers a sobering view on the description and representation of

men and women. In the paper, pointing out that in terms of appearance, men are depicted as being strong and physical while for women, breast sizes and weight are emphasised (Pearce 2008: 17) – the attention to female physical appearances later also remarked upon by Norberg (2016: 305) – Pearce (2008: 11-12) also notes that women are far more likely than men to be referred to by adjectives highlighting their marital or reproductive status, being described as *married*, *childless*, *celibate* or *single* among other terms. Later findings on gendered marital terms by Raphael (2023: 1194) show that modifiers for *wife* include *pregnant*, *dependent*, *mad* and *nagging* while *husband* is preceded by *ideal*, *brutal* and *violent*; *spinsters* are *sex-starved*, *cold-hearted* or *waspish* while *bachelors* are *eligible*, *agreeable* or *handsome*; and *bride* follows adjectives such as *beautiful*, *bartered* and *jilted* in contrast to *bridegrooms* who are *forceful*, *nervous* or *perfect*. Raphael (2023: 1194-1195) thus writes that men are described as being dominant and aggressive whereas women are perceived to be emotionally ill-tempered, weak or subordinated. Those research aspects corroborate Pearce's (2008: 7) own findings on the representations of power in relation to gender based on corpus data – in terms of verbs, men are represented as being more active and physical, doing manual labour and working with tools as *chase*, *climb*, *jump*, *dig*, *hammer*, *plough* and *haul* attest to.

Not only do verbs associated with domination and leadership collocate more strongly with men (Pearce 2008: 8), but also – rather worryingly – so do verbs which indicate criminal or deviant activity, which is perhaps an accurate representation of demographic facts (Pearce 2008: 9). Men are thus more often depicted as being *condemned*, *dangerous*, *violent*, *arrested* and *convicted*, and are more likely to *abduct*, *attack*, *fight* and *kill* than women, while *bludgeoning*, *mutilating*, *raping* and *strangling* are exclusively done by men on the basis of corpus data (Pearce 2008: 9). On the other hand, both men and women are represented as victims of violence in object position (Pearce 2008: 10), with the former being *killed*, *knifed* and *wounded*, and the latter being *gagged*, *raped* and *violated* (Pearce 2008: 20). The male inclination to kill is noted even in fairy tales, where male characters are stereotypically shown to battle and kill while female characters cook and clean (Toro Isaza et al. 2023: 6515). These findings of gender stereotypes are not too dissimilar from the ones presented in Norberg's (2016) research on the representation of BOY and GIRL. Investigating collocates using a corpus of web-based English, Norberg (2016: 298) found that activity verbs such as *kill*, *jump*, *run*, *leap*, *kick* and *fight* collocate either more strongly or exclusively with BOY whereas *dance* and *ride* appear alongside GIRL. When viewing BOY and GIRL as syntactic objects, Norberg (2016: 302) also highlights that girls are often presented as victims of violence,

being *raped*, *trafficked*, *assaulted*, *forced* and *kidnapped*, among others, revealing similar results to those from Pearce (2008: 20).

Even in descriptions of personality, gender differences and biases can be identified. For instance, Pearce (2008: 13) draws attention to personality traits grouped according to the dimensions of agreeableness, extraversion, conscientiousness, neuroticism, and openness to experience most strongly associated with MAN and WOMAN. Ye et al. (2018: 6-7), specifically exploring how men and women have been represented diachronically through adjectives associated with the aforementioned traits by utilising select Google Books corpora, define agreeableness as the ability to be accommodating, extraversion as the tendency to engage socially and "to lead and have power", conscientiousness as being associated with discipline and trustworthiness, neuroticism as "susceptibility to negative emotion" and stress, and openness as encompassing intellect and adaptability. Pearce (2008: 14), writing of agreeableness as comprising good-naturedness and kindness, mentions that while adjectives pertaining to humour and happiness are strongly associated as male traits, men also collocate strongly with disagreeable traits. In terms of extraversion, men are described as being powerful, that men are verbally more aggressive, and that women are talkative in the sense of *nagging* and *gossiping* (Pearce 2008: 10). Additional findings include general male associations with conscientiousness traits and those referring to openness and intellect, and that – where neuroticism is concerned – women are more often depicted as constantly and perhaps inherently suffering mentally (Pearce 2008: 16). Interestingly, however, Ye et al. (2018: 12-13) found through their research with diachronic corpus data that with regard to gender differences, descriptors of agreeableness, extraversion, conscientiousness and neuroticism appear in greater frequencies for men than for women, with no perceivable gender difference in terms of openness. The other crucial discovery on the part of Ye et al. (2018: 13) is that those gender differences reflected in usage frequencies of adjectives associated with various traits – with the exception in the agreeableness factor – have decreased over time.

It is worth pointing out here that employing language data to investigate and interpret gender differences and biases is not an endeavour restricted to English; indeed, there are research works which examine similar tendencies, for example, in Russian (Kuznetsova 2015), Estonian (Piits 2016) and Czech alongside English (Čermáková & Fárová 2017). Kuznetsova (2015) explores gender stereotypes linked to verbs as indicated by the past tense gender

marker in Russian by deriving a simple feminine-to-masculine ratio of their use. Of the top one hundred verbs with either a male or female human subject lean, Kuznetsova (2015: 415) observes that verbs with a male bias are mainly those which signify or highlight professional skills or roles, leadership, physical strength, criminal and negative behaviour, and argumentative communication among others. In contrast, Russian verbs with a female skew are mostly indicative of domesticity, housework, witchcraft, emotionality (e.g. crying or sobbing), feminine appearances, and being birdlike in movement and speech (Kuznetsova 2015: 429). Once again, these findings are quite similar in content to those in English, with men commonly represented as being powerful and physically active and more likely to engage in criminal behaviour, and women having their appearances emphasised whilst being represented as being emotional.

These typical collocational behaviours are also observable in non-Indo-European languages without grammatical gender such as Estonian. For instance, in her article exploring how men are described in Estonian media via adjectival and verbal collocates of *mees*, the Estonian word for *man*, Piits (2016: 148-149) likewise identifies "adjectives associated with criminal or deviant behaviour" appearing as frequent collocates alongside verbs denoting aggressive activities or behaviours with men as subjects, with no aggressive verb collocating exclusively with women subjects. In the domain of emotions, too, the results are not too different; in the Estonian dataset, men are portrayed *commanding* and *threatening* while women *sob*, *despair*, *scream* and *fear* (Piits 2016: 149-150). Furthermore, in a similar vein to findings by Pearce (2008: 11-12), Piits (2016: 150) also notes that women's – but not men's – marital status is often mentioned in the texts with adjectives such as *single* or *new*, the latter in women's case a reference to being a *new wife*, whereas *new* collocating with *man* refers to the man filling a job vacancy. The findings thus lend credence to the notion that representations of gender stereotypes are possibly cross-linguistically ubiquitous.

Conducting corpus linguistics research for literature likewise reveals gender differences. In a study of character representations in Czech and English contemporary fiction by male and female writers, Čermáková and Fárová (2017) take the approach of examining noun collocates of the possessives *his* and *her*. Their findings demonstrate that in British English fiction, the tendency is for *breasts* and *hair* to follow *her*, and that nouns such as *arms*, *car* and *office* are more likely to collocate with *his* (Čermáková & Fárová 2017: 21). In Czech literature, the noun *beauty* collocates with female characters and *brain* with male ones,

leading the researchers to suggest that attributes of femininity and masculinity in contemporary Czech literature are built around such words (Čermáková & Fárová 2017: 30). Beyond nouns, verbal collocates in fiction texts have also been examined. For instance, by looking at direct speech verbs, Eberhardt (2017) investigates gendered representations in the *Harry Potter* series of books. While recognising that similarities characterise the female and male sidekicks as they share common goals, Eberhardt (2017: 242) draws attention to the more fine-grained representations of the female character – Hermione – as speaking more *shrilly*, *gasping* and *crying* more, as well as *screaming* and *squealing* (Eberhardt 2017: 236, 239-240), whereas the male sidekick – Ron – is depicted as *roaring* and *bellowing*, both *shouting* and *muttering* more, and as using more expressive verbs (Eberhardt 2017: 233, 236, 239). Eberhardt (2017: 243) concludes by stating that as minute the differences are in terms of verbal collocations in general, the use of select verbs and descriptors in greater frequencies for one gender over another reinforces gender stereotypes.

In another study that chiefly deals with the connection between the gender of literary characters in 19th century novels and the characters' agency, Jockers and Kirilloff (2016) explore which verbs collocate with the gendered pronouns in nominative and oblique form as well as the nouns MAN and WOMAN, predicated on the idea that "what characters do influences the way we perceive who characters are" (Jockers & Kirilloff 2016: 1). Once all sentences in the novels were grammatically parsed, the researchers then pinpointed and examined the pronoun-verb pairings, subsequently finding that *wept*, *sat*, *felt* and *cried* collocate more frequently with female pronouns while *took*, *rode* and *walked* show higher degrees of association with male pronouns (Jockers & Kirilloff 2016: 13). It was also found that, although not an absolute given, female and male authors tend to write characters corresponding to their own gender as doing things commensurate with the norms of the time (Jockers & Kirilloff 2016: 16). Jockers and Kirilloff (2016: 20-21) thus conclude that "there is a strong correlation between character gender and verbs" in English novels of the 1800s, such that verbs of emotion collocate more commonly with female pronouns whereas verbs of action or motion are more likely to appear in the vicinity of male pronouns. While there admittedly are exceptions, these observations of similar tendencies in various studies seem to confirm that gender biases on the basis of nominal, verbal and adjectival collocations do exist and are indeed identifiable and quantifiable from texts.

Of special interest is therefore the research by Ritt et al. (2024) who apply the above logic in their wide-ranging study comprising investigations of various nominal, adjectival and verbal constructions. Ritt et al. (2024) diachronically track tendencies and trajectories of gender bias as distributed through collocational patterns identifiable from both the American English and British English Google Books corpora, utilising odds ratios as a statistical measure for determining the level of relative collocational strength, i.e. gender bias. Ritt et al. (2024: 12) firstly report that "[b]ias trajectories of individual items reflect gender stereotypes", discovering, for example, that nouns such as *profession* more often follow *his* while many family terms such as *children* follow *her* (Ritt et al. 2024: 13-14), and that many nouns for parts of the body – particularly *eyes* and *hair* – tend to collocate more strongly with *her* as well – with the exception of *head* which has stronger associations with *his* (Ritt et al. 2024: 15). They also find that gender bias has aggregately decreased over two centuries (Ritt et al. 2024: 16), a trend that has been noticed by Baker (2010: 145) as well, although Ritt et al. (2024: 18) nonetheless point out that "[b]iased word use is typical rather than exceptional" regardless of the degree of bias.

Arguably the most interesting findings of the research by Ritt et al. (2024) concern gender bias as expressed from the different constructions examined. The findings demonstrate that attributive adjectives show higher levels of gender-based disparities than adjectives used predicatively (Ritt et al. 2024: 21) and that adjectives denoting permanent or inherent qualities as opposed to those characterising transient properties also seem to exhibit greater rates of gender bias (Ritt et al. 2024: 22); these are unsurprising considering that attributive adjectives generally indicate inherent properties and predicative adjectives temporary ones (Taylor 2002: 455). Furthermore, males feature as subjects of active or passive constructions more than females do, and gender differences are less pronounced when males or females are in object position syntactically (Ritt et al. 2024: 24). Ritt et al. (2024: 25) also mention one other interesting finding: gender bias is not as prominent when a pronoun collocates with past tense constructions than those in the present tense. This observation hints at a possible correlation between verb constructions and gender bias, making it a prime object of study.

The work by Ritt et al. (2024) thus evidently encompasses many interesting aspects and offers preliminary insights and steps into further studies on similar matters. As there is a general paucity of research on the collocational behaviour and affinity of gendered pronouns and verb constructions themselves, this thesis addresses the research gap in that area of study.

4. Research design

To determine the collocational strengths of the gendered third-person singular subject pronouns *he* and *she* with certain verb constructions, a number of verbs and constructions were selected for study. Necessary data were gathered from the Google Books (American English) Corpus by Mark Davies (2011-), henceforth referred to as GBA. Most of the computations and quantitative analyses were conducted with the aid of Microsoft Excel. The relevant data, including the chosen verbs and constructions, and the research procedure are elucidated below.

4.1. Data, corpus and queries

As this exploratory study principally focuses on diachronically tracking and assessing how strongly the pronouns *he* and *she* collocate with certain verb constructions accounting for tense, aspect and modality as well as negation, the constructions to be examined were selected first. Thirteen constructions each with their negated counterparts were selected for study for a total of 26 constructions: present simple, past simple, *will*-future, present progressive, past progressive, present perfect, past perfect, *would*, *can*, *could*, *must*, *might*, and *should*, all along with their negated forms. Owing to technical limitations regarding periphrastic constructions, apostrophes and contractions – as it is the case, for example, with negated constructions that the auxiliary may be cliticised to the pronoun (*she's not*) or the auxiliary itself may be inflected to reflect negation (*she isn't*)[5] – only the uninflected separated forms were used for the negated constructions, i.e. the full negator *not* followed the auxiliaries DO, BE, HAVE and the other modal verbs including *will*. One exception was made for the negated counterpart of *can* for which *cannot* as one word was used.

Moreover, some possible constructions were not included due to a dearth of attestations, especially in earlier decades; consequently, several future constructions – such as the future progressive and the future perfect – were excluded. Combined constructions, e.g. progressive perfects, were likewise also omitted to simplify the research procedure. It must additionally be mentioned that the modals selected for investigation were arbitrary choices for the most part; they were included in the study despite the crucial caveat that they are not disambiguated as to whether they express epistemic or deontic modality. Nevertheless, by confining investigations of collocational affinity primarily to the formal, structural level, it

[5] See Zwicky and Pullum (1983) for a thorough exposition of the topic.

was considered worthy to examine modal constructions as well, although only a small sample of the modals were selected for more detailed analyses for the reason stated.

To enable finer analysis for the study, 36 verbs were also chosen for closer inspection alongside the constructions. The verbs listed in Table 1 were admittedly selected fairly arbitrarily and unsystematically, albeit not totally haphazardly; several factors were taken into consideration.

Table 1. List of specific verbs selected for study.

drink	*throw*	*read*	*help*
eat	*scream*	*run*	*try*
dance	*shout*	*watch*	*wait*
sing	*say*	*write*	*kill*
cook	*speak*	*play*	*listen*
talk	*work*	*take*	*ask*
think	*consider*	*look*	*carry*
complain	*walk*	*start*	*share*
cry	*give*	*bring*	*choose*

For instance, it was a conscious decision to mostly include lexical or main verbs and thus exclude those which have the same form as auxiliaries such as *do* or *have*. At the same time, verbs such as *throw*, *give* and *take* were included despite being what Butt (2010: 72) refers to as light verbs, i.e. verbs which have diminished lexical semantic content but are not exactly auxiliaries, hence the appellation to signify that they are light, semantically speaking, but not completely devoid of semantic meaning (Butt 2010: 48). Those verbs were included since – as metaphorically as they may be used in many situations – they may either still be used as lexical verbs or their more abstract uses do somewhat relate back to the intended semantic content expressed. Furthermore, stative verbs, with the likes of *love*, *know* and *want*, were generally avoided primarily so that there would be sufficient attestations in the progressive form available for analysis. It is worth pointing out that some of the verbs chosen have also appeared in previous studies cited in the preceding chapter; a number of them – such as *dance*, *cry*, *scream*, *shout*, *walk*, *run* and *kill*, among various others – have been shown to exhibit gender bias based on their collocational behaviour with gendered terms. In addition, some verbs were chosen as obvious pairs, e.g. *drink* and *eat*, *dance* and *sing*, *scream* and *shout*, *read* and *write*, and as verbs of similar meaning or action as in the case of *talk*, *say* and *speak*. As to the number of verbs, the idea was to compile a set for investigation consisting of at

least 30 verbs with the understanding that 30 samples are used as a general rule-of-thumb to demarcate a minimum sample size (see Crawley 2015: 9).

The data were collected from GBA which contains texts from the 1810s up until the 2000s, thus providing 20 decades' worth of data. The corpus was chosen to be used for this study for several reasons: it happens to be one of the largest corpora of English at over 155 billion words; there exists a British English counterpart which, although smaller in size, would enable comparison between American and British English; and it lends itself nicely to diachronic studies and tracking trends with data spanning two centuries. It should be highlighted, however, that the Google Books corpora have their issues as reported by Pechenick, Danforth and Dodds (2015). The main issue is that the Google Books set of corpora are effectively libraries, meaning that the data presented do not accurately reflect or indicate actual cultural popularity as there is, as it were, "only one of each book […] available" (Pechenick, Danforth & Dodds 2015: 2). Related to this is the possible issue of fictional characters being overrepresented in the corpus by virtue of the prolificacy of an author. Some famous authors may publish multiple books featuring the same character over a span of a decade, potentially posing the problem of skewing the corpus data; indeed, Pechenick, Danforth and Dodds (2015: 19) do observe such a case.

Another issue pointed out by Pechenick, Danforth and Dodds (2015: 23) concerns the prevalence of scientific literature skewing data which also affects word choice and collocations, such that "the contents of the Google Books corpus do not represent an unbiased sampling of publications". This latter aspect must be borne in mind as GBA was not filtered for this research; the results must therefore be viewed and interpreted with caution. The issues with the Google Books corpora notwithstanding, accepting that the corpora are by no means perfect, and despite the perception of examining written language as opposed to spoken language (cf. Linell 2005: 29), it was thought that GBA – as a large depository of general published material in American English – would still be useful for the purposes of this study even if it may not be truly representative of American English, let alone written American English. The interest mainly lies, after all, in the tendencies and trends of structural collocational behaviours and patterns rather than the content of the works.

There is one other issue with the corpus worth noting: the Google n-grams, with which the corpus is comprised, are not tagged by part-of-speech. This is especially problematic when trying to limit the search of a word with the same form but different functions to a specific

use, an issue also remarked upon by Ritt et al. (2024: 10) with the case of *her* which can either be a possessive determiner or an object pronoun, for which it is not possible to disambiguate using GBA alone.[6] It is also sometimes the case with verbs that GBA interprets words used as adjectives to be verbs due to the form (e.g. present and past participles). Nevertheless, while combined full word and part-of-speech queries are not possible,[7] searching for just part-of-speech tags in slots is thankfully unaffected; one must still be aware, however, that as the tagging system for the Google Books corpora is not perfectly accurate, one should be cautious drawing any conclusions. The structure of the 26 queries used to attain the total number of male and female constructions necessary for quantitative analysis is presented in Table 2.

Table 2. List of queries run to retrieve the total number of male and female constructions for individual verbs.

Positive constructions	Negative constructions
[he]/[she] [vvz*]	[he]/[she] does not [v*]
[he]/[she] [vvd*]	[he]/[she] did not [v*]
[he]/[she] will [v*]	[he]/[she] will not [v*]
[he]/[she] is [v?g*]	[he]/[she] is not [v?g*]
[he]/[she] was [v?g*]	[he]/[she] was not [v?g*]
[he]/[she] has [v?n*]	[he]/[she] has not [v?n*]
[he]/[she] had [v?n*]	[he]/[she] had not [v?n*]
[he]/[she] would [v*]	[he]/[she] would not [v*]
[he]/[she] can [v*]	[he]/[she] cannot [v*]
[he]/[she] could [v*]	[he]/[she] could not [v*]
[he]/[she] must [v*]	[he]/[she] must not [v*]
[he]/[she] might [v*]	[he]/[she] might not [v*]
[he]/[she] should [v*]	[he]/[she] should not [v*]

Each query was run separately for *he* followed by the strings and *she* with the same, whereby the bracketed queries search for lemmata. This was a necessary step for *he* and *she* as searching them without brackets would only return results with the pronouns in lowercase and would thus discount the pronouns in sentence-initial position. Accordingly, [he] and [she] return the minuscule, capitalised, and majuscule forms of the pronouns: *he*, *He* and *HE*, and *she*, *She* and *SHE*.[8] Occasionally, the queries would result in an error ostensibly due to the

[6] In fact, the query [v*] *er.[p*] does not return a single instance of *her* following a verb, whereas [v*] *im.[p*] does result in [VERB] *him*. This indicates that no instance of *her* is 'tagged' as a pronoun.
[7] GBA does however allow queries involving wildcards for word forms, i.e. substrings, tagged by part-of-speech as in the examples above.
[8] In isolation, the bracketed pronouns also return *hE*, *sHe*, *SHe*, *sHE* and *shE*. These forms, while attested, are not included as they are practically negligible on a wider scale.

number of bracketed terms; in those instances, [he] and [she] were respectively replaced by he|He|HE and she|She|SHE to account for all attested canonical forms of the gendered pronouns. For queries involving modals – including *will* – the simple [v*] tag, which looks for all verbs, was used following the modal verb.

That the third-person singular pronouns mark the present tense is exceptionally felicitous as there exists a separate [vvz*] tag for that very purpose, particularly useful for words like *read* which would otherwise have the same orthographic form in both the present and past tense with other pronouns and therefore pose disambiguation problems. For past tense queries, the [vvd*] tag was employed. The [vvz*] and [vvd*] tags were chosen for the simple present and simple past constructions in lieu of the general [v?z*] and [v?d*] tags as the former set of tags limit the search to lexical verbs and exclude instances of *be*, *have* and *do*. This was done in case these verbs, even if they are used as main verbs in many instances as is the case with *have*, would interfere with calculations for the simple present and past tenses, considering that the general tags naturally contain subsets of progressive, perfect and passive constructions as well. Therefore, it ought to be emphasised that when looking at individual verbs, the simple present and past constructions of the verbs are measured against present and past tense constructions of solely lexical verbs, which, although fairly restrictive, is a suitable compromise for the purposes of this research. Furthermore, so that passive constructions would explicitly be ignored, searching for progressive and perfect constructions required the respective use of the [v?g*] tag returning the *-ing* forms of verbs, and the [v?n*] tag which marks the past participle forms of verbs.

In addition to the queries listed in Table 2, further queries were made to get an overall view of the level of gender bias with verb constructions. For such global perspective, the query [he]/[she] [v*] was set as a benchmark against which other constructions were compared. The present, past and future tense searches were retained as they were, the simple progressive constructions were queried with [he]/[she] [be] [v?g*], perfect constructions with [he]/[she] [have] [v?n*], modals with [he]/[she] [vm*] [v*] with instances of *will*-constructions subtracted, and negative constructions with [he]/[she] [v*] not [v*]. Here, too, the progressive and perfect constructions did not include the future forms to maintain balance.

It is also worth mentioning that all these queries follow a strict structure and thus do not take into consideration cases where, for example, the pronoun and verb are separated by an adverb. Male and female verb constructions for the purposes of this study are therefore understood to

be limited to constructions in which the verb immediately follows the gendered subject pronouns and, in the case of periphrastic constructions, the lexical verb likewise immediately follows the modals or auxiliaries. The only exception was made for negative constructions: the negator *not* must appear between the auxiliary and the lexical verb and cannot be put in another position, such that negative constructions where one or more words separate the negator from either the auxiliary or the lexical verb are not taken into account. Each time the aforementioned construction queries were run for both *he* and *she*, the frequencies per decade were collected and transferred to a Microsoft Excel spreadsheet for quantitative analysis.

4.2. Methodology

As this study is taken to be an extension of Ritt et al. (2024), similar methodological procedures and processes for quantitative analyses were adopted and applied. In particular, the same statistical measure of collocational strength was utilised to quantify gender bias: odds ratio.

Schmid and Küchenhoff (2013: 547), assessing previous methodology utilised by Stefanowitsch and Gries (2003) for collostructional analysis, detail several issues and shortcomings concerning the Fisher's exact test method taken by the latter authors vis-à-vis interpretability, directionality and sample size. For this reason, Schmid and Küchenhoff (2013: 552) present odds ratio values as one of several alternative statistical measures to the *p*-value resulting from Fisher's exact tests for determining and quantifying collostructional strength. The underlying principle behind odds is that they "relate probabilities to converse probabilities" (Schmid & Küchenhoff 2013: 553), i.e. indicate relative probabilities. In other words, odds express the probability of occurrence against the probability of non-occurrence as a ratio; the odds ratio, as the name suggests, is therefore a ratio of ratios. Where collostructional analysis is concerned, the odds of a lexeme occurring with a given construction and the odds of the construction occurring without the lexeme are calculated respectively and then divided to obtain the odds ratio and hence, in this context, the strength of association between the lexeme and the construction (Schmid & Küchenhoff 2013: 553). Only if the resultant odds ratio value is equal to 1 can it be said that there is no attraction between the lexeme and the construction (Schmid & Küchenhoff 2013: 555). As the odds ratio approach to measuring collostructional strength is bidirectional – both lexeme-to-construction and construction-to-lexeme strengths are taken into account in mathematical terms (Schmid & Küchenhoff 2013: 554) – and also frequency-adjusted, with the frequencies

of lexemes of interest or constructions not affecting the measures of strength – Schmid and Küchenhoff (2013: 555) hold the view that odds ratios are generally superior to other statistical approaches – such as the Fisher's exact test – for collostructional analysis.

Odds ratios have been used to measure and analyse gender bias; the study by Ritt et al. (2024) is one such example where it is utilised. In it, Ritt et al. (2024: 5) present their adapted odds ratio methodology, where the odds of male constructions are divided by those of female constructions to derive odds ratio values which effectively quantify gender bias based on observable corpus data. Ritt et al. (2024: 8) also point out that the odds ratio is a mathematically simple statistical measure, and this is indeed true. Investigating the collocational affinities of only the subject pronouns *he* and *she* with verb constructions as in the case of this study, only four values are needed to calculate the odds ratio: the frequency of *he* and of *she* followed by the construction of interest and the total number of constructions of the same kind collocating with *he* and with *she*.

So, in the most general terms, the odds ratio (OR) for this study is calculated thus:

$$OR = \frac{x_m}{(X_M - x_m)} \Big/ \frac{x_f}{(X_F - x_f)},$$

where x_m denotes instances of the male construction of interest, x_f the female construction of interest, X_M the total number of male constructions of the same kind, and X_F the total number of female constructions of the same kind. As it is necessarily so that $x_m \subset X_M$ and $x_f \subset X_F$, the frequencies of male and female constructions of interest must respectively be subtracted from the total number of male and female constructions of the same kind. What the total number of constructions actually is relates to how a construction is defined and which cases count towards being a construction, and as Stefanowitsch and Gries (2003: 218) write, this all depends on the researcher.[9] For this thesis, queries from the previous subchapter serve the purpose of retrieving these overall instances of constructions.

Once all the values are in place, the calculation process is straightforward. For example, if one were interested in the overall degree of gender bias exhibited by past tense constructions in general on the basis of GBA data from the 1810s, the occurrences would need to be retrieved as presented in Table 3. A binary system of categorisation where there are only two

[9] How the cell intended for the total number of constructions should be filled has long been a point of concern in collostructional analysis and more generally in corpus linguistics research utilising contingency tables, having been noted and covered by Schmid and Küchenhoff (2013: 541-544).

target lexemes greatly simplifies the statistic procedure since all constructions that do not begin with *he* in this context is simply understood as all those beginning with *she*.

Table 3. Absolute GBA construction frequencies from the 1810s to calculate the overall level of gender bias with past tense constructions.

1810s		
	[he]	[she]
[he]/[she] [v?d*]	786,323	102,177
[he]/[she] [v*]	2,118,832	244,037

From here on, computing the odds ratio is simple. One need only input the values into the formula previously mentioned and run the calculation

$$OR_{past.overall} = \frac{786{,}323}{(2{,}118{,}832 - 786{,}323)} \Big/ \frac{102{,}177}{(244{,}037 - 102{,}177)}$$
$$\approx 0.59011 : 0.72027$$
$$\approx 0.81929$$

such that the *he*-to-*she* odds ratio for past tense constructions overall can be obtained, which, in this scenario, is approximately 0.81929. The male construction odds are pitted against – i.e. divided by – female odds, leading to the value; in this instance, as the odds ratio is below 1, this means that taking written language samples of published American English material from the 1810s as included in GBA, it is more likely for *she* to precede a construction in the past tense than *he*. The same interpretation is made if the female-to-male construction odds ratio were to be calculated; in that case, the odds ratio would be approximately 1.22057, indicating that *she*, as opposed to *he*, has about 1.22 times the odds of appearing before a past tense construction.[10] It must be stated that this research conventionally takes the male-to-female odds ratios, i.e. male odds divided by female odds and not vice-versa, and that such a choice has nothing to do with value judgement.

With individual verbs, calculating the odds ratio is the same procedure. It simply involves getting, for every decade, the number of occurrences of the verb in a particular construction following *he* or *she* and the total frequency of the construction preceded by *he* or *she*. Taking

[10] The same value is obtainable when dividing 1 by the male-to-female odds ratio. This determines by how much the odds of *she* appearing before a particular construction are higher compared to *he* should odds ratios of below 1 be returned after getting the ratio of male odds to female odds. An odds ratio of above 1 after dividing male odds by female odds signifies male bias and the resultant value itself indicates by how much the odds of *he* before a construction are higher.

an example verb – *work* – and placing it in the simple present construction, the odds ratio to ascertain to what extent *he* or *she* collocates with the verb *work* in the simple present from 1810s GBA corpus data can be obtained from the values in Table 4.

Table 4. Absolute 1810s GBA frequencies of *he works*, *she works* and simple present tense constructions of lexical verbs preceded by *he* and *she*.

1810s		
	[he]	[she]
[he]/[she] *works*	400	20
[he]/[she] [vvz*]	222,663	20,997

Table 4 shows that there are 400 attestations of [*he*] *works* and 20 attestations of [*she*] *works* for the 1810s decade in GBA. Using the same earlier formula, the resultant odds ratio here is approximately 1.88758, indicating a male bias – in other words, the odds that the third-person singular verb form of *work* in a selection of American English texts from the 1810s in a pool of lexical verbs of the same construction takes the male subject pronoun are roughly 1.89 times higher than *she* preceding it.

As can be seen from the examples, the ratio of male odds to female odds results in an odds ratio where female bias is indicated when $0 < OR < 1$ while $1 < OR$ signals male bias. As there is theoretically no upper bound to the odds ratio when it is calculated in this manner, it may seem that male biases are sometimes grossly overrepresented or amplified compared to female biases; Ritt et al. (2024: 6) observe that this issue is particularly conspicuous in graphs as the odds ratio values tend to visually skew male and female ratio relations. To address this problem, Ritt et al. (2024: 6) transform the odds ratios with the common logarithm – logarithm with base 10 or $\log_{10}$ – which has the benefit of setting the baseline of the lack of gender bias associated with a given construction at 0 and not 1.[11] Thus, the consequence of logarithmically transforming the odds ratios is that values below 0 – negative values – indicate female bias while those above 0 – positive values – indicate male bias. Hence, as Ritt et al. (2024: 6) point out, "[equally strong] male and female biases […] show up at equal distances from the line indicating neutrality", greatly improving and simplifying interpretation and visualisation of the data. As examples, the logarithmically-transformed odds ratio (LOR) of the odds ratio garnered from the data in Table 3 is approximately -0.08656, signifying female bias, while the LOR of the value calculated from Table 4 is

[11] The LOG function is used for this purpose in Excel.

roughly 0.27591, signalling male bias. Both LORs obviously convey what is already known, but the relation of gender bias is made clearer and more evident through the converted odds ratios as the skew in the data is corrected while negative and positive values conveniently correspond to female and male bias, respectively.

Ritt et al. (2024: 7) astutely note, however, that LOR values are not all too useful in determining "the overall strength of gender bias", i.e. how strong the biases are when taking the levels of both genders into account, as negative and positive scores cancel each other out when averaged. The logarithmic (or simply log) odds ratios thus demonstrate in a straightforward manner whether a construction is more likely to be preceded by male or female pronouns but do not depict the true strength of female biases. The value of collocational strength expressed by negative LORs is the reciprocal of that indicated by positive LOR values; even if, for example, Construction A has a log odds ratio of 0.30103 and Construction B a log odds ratio of -0.30103 – meaning that the actual strength of female bias is equal to that of male bias – when taking the LOR values as a measure of collocational strength, the interpretation would be that Construction B is half as likely to take the female pronoun while Construction A is twice as likely to take *he* over *she* when converting back from the LORs. That is obviously incorrect, as the level of bias exhibited by either construction is the same. For the aforementioned reasons, Ritt et al. (2024: 7) use the absolute values of the logarithmic odds ratios to assess the overall level of gender bias that constructions display.[12] By ignoring the sign of the LOR value – which concerns female bias in this case – the absolute logarithmic odds ratio (ALOR) provides information about the rate of general bias strength even for constructions with female bias by measuring the distance of the LOR from 0, and thus can be used to compute average cumulative levels of both male and female bias for a given construction by taking both collocational affinities into account. There are therefore three values of relevance here: odds ratio (OR), logarithmic odds ratio (LOR), and absolute logarithmic odds ratio (ALOR).

The three odds ratio values of all queries were obtained for all decades. For the individual verbs component, there were 26 constructions for each of the 36 verbs to be examined for both genders, amounting to a total of 1,872 queries run. As each query returns data for 20 decades, there were enough data to fill at least 37,440 cells in Excel for analysis. Once every construction for both genders for all verbs was accounted for, multiple summary tables

[12] This is realised through the ABS function in Excel.

consisting of the verbs, the constructions, the LORs and the ALORs were then created for easier analysis. With the data in summary tables, it is possible to get and have an overview of the average levels of bias per decade for each verb and/or construction as well as for gender in a more convenient and efficient manner.[13] Similar albeit considerably simplified steps were taken for assessing global bias across grammatical categories. All these data form the foundation and bulk of the results.

It should be emphasised here that depending on the verb, attestations of constructions in GBA can be conspicuously sparse. Indeed, it was a noticeable and frequent occurrence while running the queries that no results were returned for some verbs and their constructions taking the pronouns *he* or *she* especially for the earliest documented decades in GBA, namely those of the early to mid-1800s. Occasionally, there were absolutely no attestations whatsoever in the entire corpus of certain *he* or *she* constructions with some selected verbs; in other cases, attestations were only available for either male or female constructions but not for both. The latter point poses a significant problem; if there are attestations of male constructions but not of female constructions in the same period, calculating the odds ratio is not possible as that entails dividing by zero. If there are instances of female constructions but none of male constructions, the logarithmic odds ratio cannot be computed as the odds ratio would be zero. In either case, the lack of verb construction attestations for one gender arguably suggests bias towards the other. However, since the values were not deemed comparable in such cases and the zeroes interfered with the computation process, all cells displaying either a divide-by-zero or a number error were discarded from the summary tables. While there does exist a common corrective measure known as the Haldane-Anscombe correction which involves adding 0.5 to all cells should there be at least one "zero cell", i.e. "[a] cell with a frequency of zero", in a two-by-two table for odds ratio calculations (Weber et al. 2020: 913) – utilised, for example, in the study by Toro Isaza et al. (2023: 6513) who also use odds ratio as their main statistical measure of bias – the correction was not applied for this thesis. There was the risk that applying the correction might unduly skew the odds ratio in comparison to the ratios for frequencies actually observed. It was also deemed that there can be no absolute certainty that the verb and its gendered construction – if unattested for some decades in GBA – must have been used in other settings. The decision was therefore made to just accept that no attestations are sometimes inevitable and that zero cells are technically also findings.

[13] See Appendix for the tables.

Based on the collated frequencies and the resultant log odds ratios and absolute log odds ratios, charts were produced to visualise the data and to track the trends and trajectories of gender bias across verbs and constructions as measured through derivations of the various odds ratio values. These charts serve as the basis for the findings and results provided in the succeeding chapter.

5. Results, analysis and discussion

Before the findings are presented, it ought to be reiterated that this research utilises written language data for which only the instances of the gendered subject pronouns *he* and *she* immediately before specific verb construction in the texts comprising GBA are of interest. The authors' genders – even if such information were available or easily accessible – are thus not of relevance here. As such, the study focuses on variance in language used about – not by – men and women and is therefore not necessarily about how males and females use language but rather how and what males and females are depicted or represented as doing. One must hence bear in mind that the results of the study pertain to how men or women – be they fictional or real – have been written as doing by others and are thus specific to that context.

That being said, the research yielded fairly interesting findings which were occasionally unexpected. Some of the results, such as those regarding the verbs chosen for analysis, generally corroborate previous findings on their being more strongly associated with being female or male pursuits or actions. One can also see from the results that, as similarly pointed out by e.g. Baker (2010: 145) and Ritt et al. (2024: 16), gender bias appears to have generally decreased over time on the basis of corpus data. However, many other findings of this research are quite novel and fascinating discoveries worth exploring and elucidating further.

As the research technically consisted of two parts – one which concerned measuring overall gender bias across select verb constructions where the verbs were not important and then the other assessing the level of gender bias across select verbs and their constructions – the results and analyses are likewise divided into two broad sections corresponding to those two parts of research. The findings regarding overall gender bias as analysable from the corpus is presented first, followed by the results for specific verbs analysed according to several dimensions.

5.1. Overall or global gender bias

Gauging and characterising global or overall gender bias with select verb constructions in GBA offer valuable insights as regards the general nature and trend of bias over time. Hence, before delving immediately into examining individual verbs and their constructions, it was deemed useful to also gather information about gender bias on a wide, general scale.

Beginning with the basics, it is remarkable to see the discrepancy – in absolute terms – between the frequencies of the subject pronouns *he* and *she*. When the number of occurrences is taken cumulatively, the ratio of *he* to *she* is almost three-to-one, meaning that *he* is three times more frequent than *she* in the dataset; the male subject pronoun thus accounts for at least 75% of gendered pronoun attestations in GBA across all decades, as is visible from the data provided in Table 5. As such, the instances of *he* number around 4,732.67 per million words compared to *she*, which occurs 1,562.46 times per million words in GBA.

Table 5. Total number of *he* and *she* tokens in GBA along with the total number of tokens throughout the corpus. The queries for *he* and *she* exclude other orthographic variations of the pronouns which are attested but practically negligible.

	Total # of tokens
GBA total	156,770,740,579
he|He|HE	741,944,347
she|She|SHE	244,947,358

The pronouns are clearly not evenly distributed as this cursory look at their absolute frequencies in GBA shows. This is not surprising; in fact, this is something to be expected. Both Pearce (2008: 2) and Baker (2010: 131) note such a bias in their data despite using different corpora, with Norberg (2016: 293) likewise listing other studies which make the same observation of gender asymmetries. One can surmise, then, that males – or the subject pronoun *he* at the very least – generally receive greater representation than females in corpora. Similar ratios are found for *he* and *she* followed by a verb, as shown in Table 6.

Table 6. Total number of tokens of *he* and *she* followed by any verb in GBA.

	Total # of tokens
he|He|HE [v*]	640,870,909
she|She|SHE [v*]	213,904,512

In addition to the *he* + [VERB] construction being thrice as likely as that involving the pronoun *she*, one can see from the data that it is ever so slightly more frequent for *she* to be followed by a verb (87.3%) than for *he* to be (86.4%).[14] These are interesting findings in themselves, but as we are dealing with a diachronic corpus here, it is even more interesting to see whether there is any kind of identifiable historical trend concerning these types of gender asymmetries. For example, Figure 1 depicts the trend followed by the ratio of *he* [v*] constructions to the *she* [v*] ones.

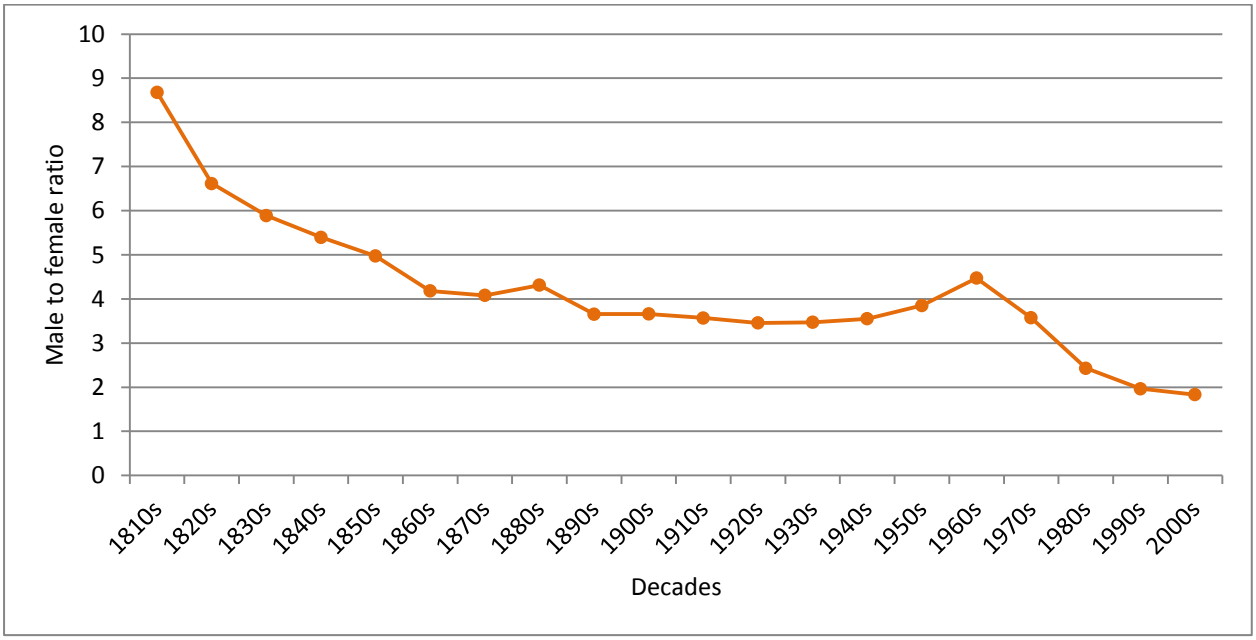

Figure 1. Trajectory of the *he* [v*] to *she* [v*] occurrence ratio over time (GBA).

What can be seen from Figure 1 is that the ratio of the male and female verb constructions has decreased over time, indicating a trend towards greater parity especially since the 1970s; a ratio of 1 would mean that the constructions are equally represented in the corpus for the decade. It should be mentioned that the male constructions outnumber the female ones in all decades. Thus, while the *he* [v*] construction is attested almost 9 times more than *she* [v*] in the 1810s going by the data in GBA, the likelihood of the male construction occurring in the 2000s decreases to such a point that it is less than twice as likely. One can ask whether this downward trend is because male constructions have become fewer or because female constructions have become more frequently attested. Figure 2 shows a logarithmically scaled chart – done to minimise distortion of the dataset – of the absolute number of attestations of the male and female constructions under investigation across 20 decades.

[14] This could be due to the use of the inclusive combination *he and she* followed by a verb, in which case *she* immediately collocates with the verb and not *he*. This is mere speculation and will not be discussed further.

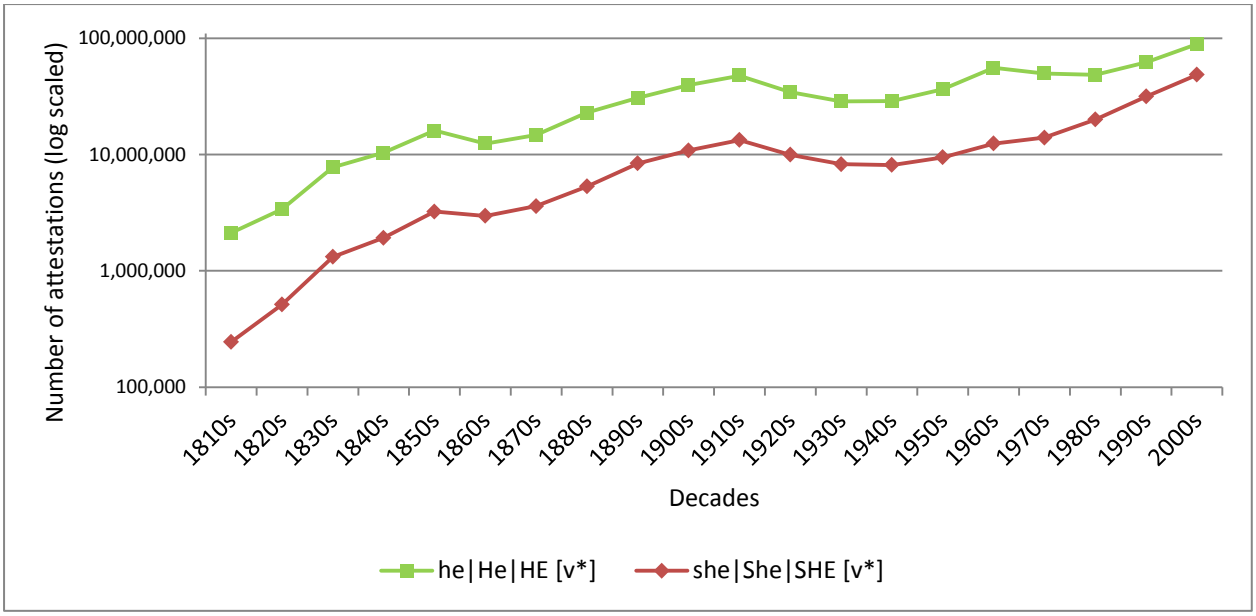

Figure 2. Logarithmically scaled line chart showing the upwards trajectory of the absolute number of male and female construction attestations in GBA over time.

As can be observed, the gap between the two constructions is closing in absolute terms; it can be said that the number of female constructions has increased quicker compared to the number of male ones in earnest since the 1980s. Indeed, since that decade, both *he* [v*] and *she* [v*] constructions see an increase, implying that female constructions have become more commonly attested relative to male constructions in the same period. In both Figures 1 and 2, a seemingly anomalous and prominent uptick can be observed for the 1960s. That could be attributed to the sudden increase in the number of tokens in GBA from the 1960s onwards; the total number of tokens in GBA for the 1950s is 8,103,600,000 while for the 1960s, the number is 13,191,800,000. Every succeeding decade then contains at least 14,000,000,000 tokens, finishing with 26,882,300,000 tokens for the 2000s alone.[15] At the same time, it is difficult to say why there was such an influx of *he* in published works relative to *she* during this time. On the other hand, somewhat easier to explain – at least speculatively – is the increase in the visibility of female verb constructions since the 1970s which could perhaps be linked to the already-mentioned rise of the second wave feminist movement which characterised the decade in matters regarding gender equality (Eckert & McConnell-Ginet 2013: 37). This information is worth keeping in mind; as later presentations of results will show, the 1970s was a seminal decade marking a turning point where gender is concerned, something also alluded to by Ritt et al. (2024: 17), and is therefore of noticeable relevance

[15] The total number of tokens per decade is available and can be retrieved from https://www.english-corpora.org/googlebooks/help/dataset-sizes.asp.

also in elucidating and interpreting the findings to follow. It should additionally be mentioned that although it appears – based on the visualisation in Figure 2 – as if the absolute frequency of female constructions is closely approaching that of male ones, the level of disparity is still remarkable: the *he* [v*] construction occurs 88,875,313 times in the 2000s compared to 48,550,900 instances of *she* [v*] in the same decade.

Moving on from merely examining absolute and relative frequencies of *he* and *she* and their verb constructions, it makes more sense to now look at how and to what degree the gendered subject pronouns collocate with select constructions globally to assess the level of possible gender bias. The absolute log odds ratios can first be surveyed to determine the overall level of gender bias across constructions in GBA. Figure 3 charts the bias trajectories of several constructions under investigation, consisting of present, past, *will*-future, progressive, perfect, modal and negative constructions as a whole.

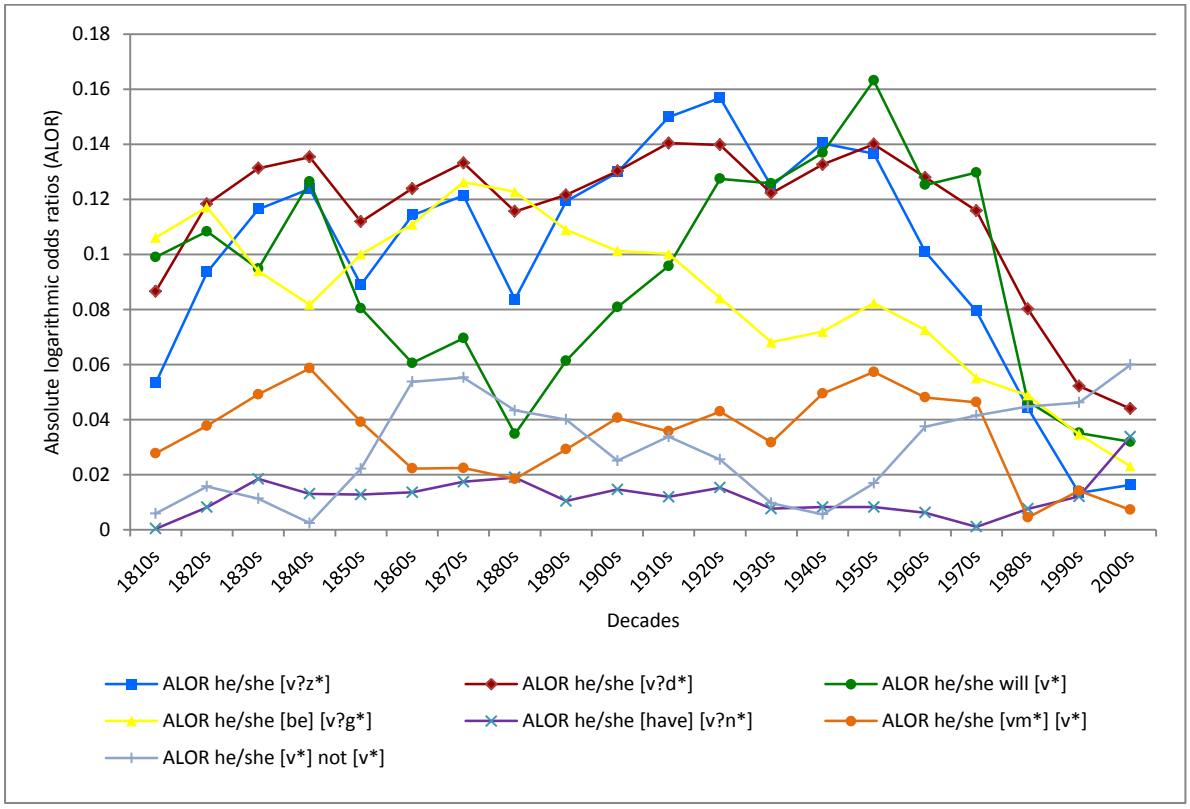

Figure 3. Absolute log odds ratio trajectories of various constructions from GBA data.

Despite the fairly erratic nature of the trajectories, a number of intriguing observations can nevertheless be made using Figure 3 as basis. Firstly, it should be noted while the chart does show gender bias – in the form of collocational strengths of *he* and *she* with select constructions – being exhibited by constructions to varying degrees, the biases are actually

quite small. Indeed, remembering that 0 is the baseline indicating no bias, the highest ALOR that occurs in the dataset is just over 0.16, expressed by the *will*-future construction in the 1950s. Ritt et al. (2024: 15), finding similar low log odds ratios, argue that these still reflect "true, albeit only statistical bias", also pointing out that a log odds ratio value of 0.16 signifies that the odds a construction follows one of the two gendered pronouns is around 44% greater than it following the other pronoun. With this reassurance, the findings here can also be considered worthy of analysis.

With that in mind, immediately conspicuous in Figure 3 is the low level of bias associated with the perfect construction: the ALOR holds steady below 0.02 up until the 1990s before almost tripling come the 2000s. The increase in the overall degree of gender bias is also noticeable with the negative constructions; the ALOR value of the construction is generally or relatively low compared to that of other constructions throughout the decades represented, at least until the 1960s. From then, the level of bias increases for the negative constructions to such a point that they constitute the most biased construction by the 2000s with an ALOR of almost 0.06. It is also especially interesting to see that these two constructions take trajectories contrary to those of other constructions from the 1970s onwards – while other constructions effectively come to display less gender bias, the perfect and the negative constructions increasingly become more biased.

Another noticeable tendency in Figure 3 concerns the bias trajectories of the progressive and negative constructions from the 1810s until the 1950s at the earliest; while they are not perfectly parallel progressions, they do rise and fall at approximately the same periods. Indeed, until the bias level of the negative construction reaches a new high in the 2000s, the progressive and negative constructions are both at their most biased in the 1870s and generally gradually lose that association of bias up to the 1930s. This contrasts with the tendency found for the present, past and *will*-future tense constructions which all register a dip – particularly conspicuous ones at that for the present and future – in the 1880s before biased constellations of the three constructions ostensibly become increasingly frequent until the 1920s. It is also the case with the present, past and future constructions that their bias trajectories are fairly similar throughout the decades represented in GBA. As with the progressive and negative constructions showing comparable ALOR trends, the three tense constructions appear to dip and peak almost simultaneously – admittedly at varying degrees but nonetheless displaying familiar trajectories. The bias trajectory of modal constructions

resembles that of the tense constructions albeit in a more irregular manner; in any case, biased use of the construction increases and decreases at roughly the same time periods as the tense constructions.

Bias trajectories of the latter half of the 20th century are also of note, made conspicuous by the rapid decrease in biased pronoun associations with most constructions after the 1950s. In fact, while it is indeed the case that especially precipitous drops in biased use can be observed with the tense and modal constructions after the 1970s, visualised data in Figure 3 do show that the downward trend towards unbiased use of the constructions appear to have already started by the 1960s. This is identifiable for all but the perfect and negative constructions; the tense, progressive and modal constructions all display gradually decreasing levels of bias from then on. However, the dramatic decrease in bias with regard to the gendered pronouns collocating with select constructions after the 1970s, a decade already established to be notable, cannot be ignored. One need only consider the steep fall in the case of the *will*-future construction – a decrease of the ALOR from 0.13 in the 1970s to under 0.05 the next decade for the future tense is quite substantial – as well as the modal construction to a lesser extent, whose ALOR values drop from roughly 0.05 to becoming almost negligible the next decade in the 1980s. It is thus also remarkable to see, for example, the reduction of subject pronominal gender bias in the present and past tenses. The ALORs for the present tense fall from around 0.08 in the 1970s to almost 0.02 – with a slight increase from the 1990s – in the 2000s while the values fall for past tense constructions from close to 0.12 to well under 0.05 in the same period. Although past tense constructions do continue to display higher levels of bias compared to other constructions in the 2000s, that the form has become less biased over several decades is still noticeable.

The trajectories in Figure 3 therefore demonstrate that gender bias has generally decreased in the case of the tense, progressive and modal constructions, at least since the second half of the 20th century after increasing considerably from the late 19th century onwards. The chart also shows that for an extended period of time, the tense and the progressive constructions displayed the highest levels of bias up until the 1980s. That the tense constructions have similar trajectories, i.e. peaks and troughs as well as increases and decreases being marked in the same or adjacent decades, is not all too surprising. What is peculiar, however, is how the trends for the progressive and negative constructions appear to be aligned with each other until the 1950s whereupon they diverge. This correlation is most likely trivial and a matter of

coincidence as there is nothing concrete which could provide and function as a suitable and satisfactory explanation. At the same time, it is admittedly rather puzzling as to why biased usage of both constructions declined in the periods 1820s-1840s as well as the 1870s-1920s and increased between the 1840s and 1870s at a quicker rate relative to other constructions, which generally registered an increase or decrease in bias in the same timeframes.

With Figure 3, the more reasonably certain interpretations are the obvious ones. For instance, the continued downward bias trajectory of the progressive construction since the 1870s suggests that depictions of incomplete, temporary, dynamic or ongoing actions and activities in American English writing have gradually become less associated with one gender over the other. The sharp increase starting from the 1880s followed by the steady post-1950s collapse of tense construction bias trajectories – especially evident for the present and future tenses – possibly hint at changing societal sensibilities and norms such that males and females in written works are represented as conducting activities in the present, past or future at greater frequencies depending on the gender. The dip in the late 20th century signifies that bias for those constructions – while still existent – has collectively reduced, indicating an approach to parity at observed rates. However, the ALOR values do not show which constructions express higher levels of bias for which gender and only offer data on how biased a construction is. On the other hand, Figure 4 presents diachronic bias trajectories of the constructions based on LORs, in which associations with gender are visible.

With LORs, negative scores indicate female bias and positive scores signify male bias, signs which are of interest for this research. Using that logic of negative and positive values, it is clear to see which constructions show higher levels of association with *he* and which with *she*. Viewing the chart, a number of constructions stand out as being distinctly biased. The past and the progressive constructions consistently display female bias throughout the 20 decades in GBA while another two constructions – the present and the future – likewise continually exhibit male bias. Indeed, the tense constructions display the highest levels of bias when viewing the data cumulatively, and it is interesting to see how the tense constructions collectively became more biased between the 1880s and 1920s.

There additionally are three constructions whose biases teeter and shift over time: the modal, the perfect and the negative constructions. The modal construction can be observed displaying male bias until the 1970s before shifting to exhibit low levels of female bias from the 1980s onwards. The perfect construction, effectively the most ‘gender-neutral’ of the

constructions as noted before, continuously shows a bias towards collocating with *she* rather than *he* at very low rates up until the 1970s when it crosses into male bias territory and exhibits considerable male bias by the 2000s in relation to other constructions. The trajectory of the negative construction is quite intriguing: the odds that *she* precedes negative constructions are greater between the 1850s and 1930s but a turn is seen from the 1940s onwards, when the construction is characterised by an increasing likelihood of male bias which continues into the 21st century. Again, Figure 4, as with Figure 3, demonstrates that gender bias has actually reduced in most cases by the 2000s in comparison to preceding decades, as evidenced by the bottleneck of the 1980s. One can see that the LORs of the tense, progressive and modal constructions – while they all still display either female or male bias – are generally approaching zero, suggesting a trend towards neutral usage.

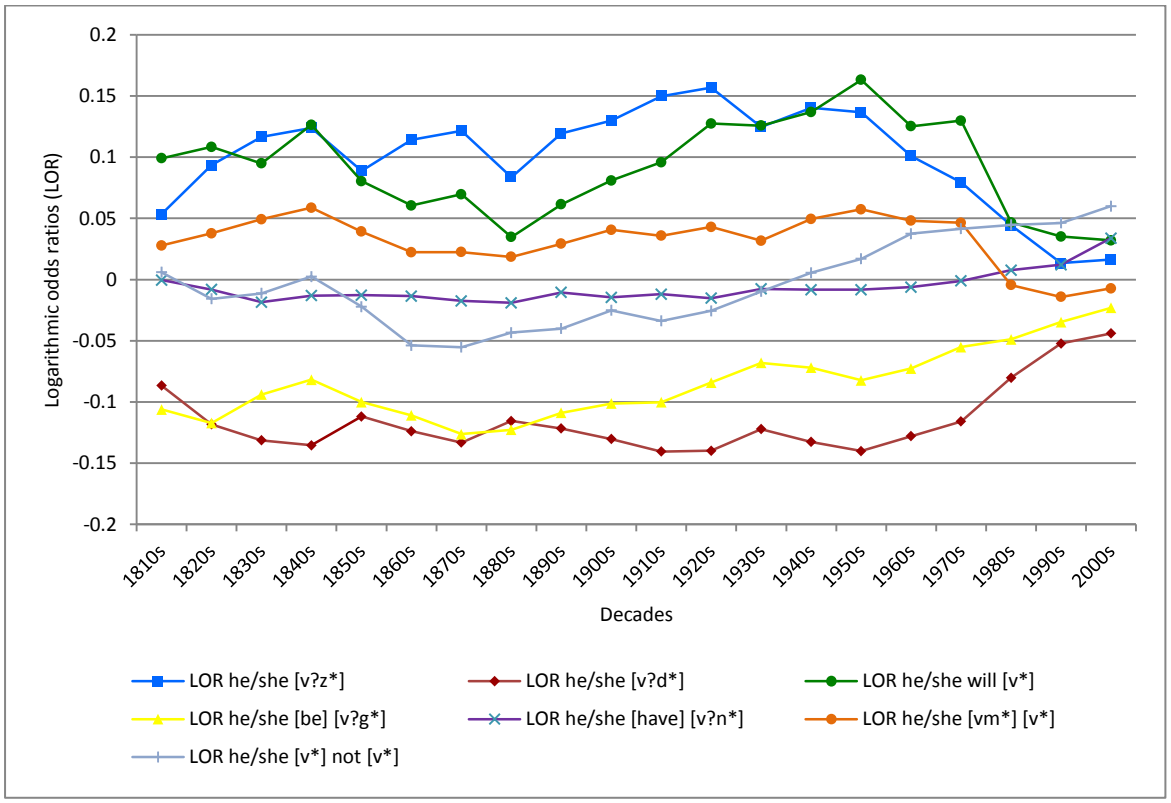

Figure 4. Log odds ratio trajectories of various constructions from GBA data.

The data as visualised in Figure 4 actually pose further questions. Why the present and future tense constructions appear as being especially biased towards taking *he* as a subject while the past tense and progressive constructions seem to particularly favour the female subject pronoun is probably not something which can be readily or easily answered. One may speculate, however, in the same way that one notices that the post-1970s decrease in bias across several constructions appears to coincide with the rise of feminist movements of the

1970s, thereafter surmising that worldwide developments towards the advancement of women's rights may have played a role in the reduction of gender bias in language as, for example, Ritt et al. (2024: 17) do.

Within the context of this research, it could perhaps be the case that there is a strong inclination for males in published written material of American English to be presented in the texts as active and current doers performing regular or habitual tasks in the present tense which imbue their actions and accomplishments with a sense of permanence and timelessness. One could also speculate that the *will*-future construction collocates strongly with *he* because men in texts may be depicted as determined beings whose future actions are certain, or it could be that they are written as lacking choice or volition as to what they are expected to do. Since it is unknown whether these tenses are used within dialogue or by a narrator describing events or states, it is impossible to make generalisations. It may be that the utilisation of the future by an omniscient narrator for the male figure could suggest that everything the character – used here in the broadest sense of the word – does is predestined; the future in connection with dialogue may signal that there is greater future predictive certainty on the part of the male as regards his actions from the perspective of other figures in the text.

With the past tense, it is known that the tense features heavily and has long been used in narrative fiction (Harvey 2006: 71-72), so for the female subject pronoun to have higher odds of appearing in connection with the past tense than *he* – the same way the male subject pronoun precedes present and simple future tense constructions at greater rates than the female pronoun – likely alludes to something. Considering the distancing effect introduced by the past tense in that it removes and detaches a situation described from reality, one wonders whether the level of female bias implies that women are more likely to be written as not just having done things such that they are not in the now, but also as being figuratively separated and maintaining a temporal and spatial distance from men who are more often inclined to be depicted as acting in the present. Ritt et al. (2024: 25) speculatively posit that such an interpretation could be possible, with female figures thought of as being "safely remote" from males so as to not be considered a threat. Of course, these are all hypotheticals and it is wise to exercise caution when interpreting any of the results without scrutinising the contexts in which the constructions appear with the gendered pronouns.

Of the other constructions, the progressive also makes for interesting viewing. The trajectory indicates that females are more likely to be represented in texts as dynamically being in the

process of doing something, durative actions in progress, for which – if the action is completed – the focus is on the steps taken to accomplishment rather than the actual culmination of the event, at least from the speaker's perspective. There additionally is the eventuality that the event may not be completed and remain unfinished too, should the action be interrupted. It could also be that since the progressive can be used to shift settings in narratives to invoke simultaneity of multiple actions (Fludernik 2012: 89) for which the progressive often provides expositions of secondary events or information, female figures in written works are relegated to secondary position, an addendum, almost akin to an afterthought, doing things but not necessarily being shown to complete the tasks. Coupled with the progressive construction, then, women can be perceived as being more susceptible to scrutiny to having their actions presented as lacking determinate endpoints through restricted viewing frames of the event described. These effects could possibly be amplified when used in combination with e.g. the present or past tense. It would therefore be of interest to see how gender bias is realised when the [he]/[she] [be] [v?g*] construction is differentiated into the present and past progressive, as visualised in Figure 5.

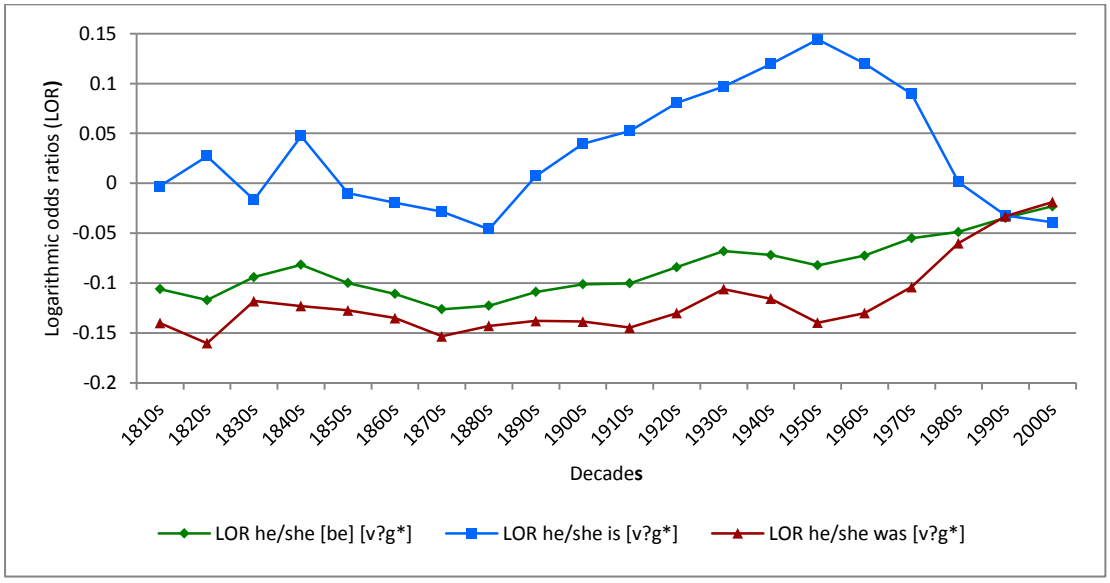

Figure 5. Log odds ratio trajectories of the present progressive and past progressive constructions from GBA data.

The two trajectories can clearly be distinguished from the undifferentiated trajectory. The odds for the past progressive construction to be preceded by *she* is consistently greater than *he* preceding it, although the likelihood has diminished considerably since the 1950s. The present progressive, on the other hand, displays an interesting tendency in which the male subject pronoun continuously collocates more strongly with the construction than the female

one between the 1890s and 1970s, whereupon a switch occurs following the 1980s occurs in that female bias is registered thereafter. The dominance of female bias with the past progressive could mean that women are more likely to be depicted as having their actions interrupted with the sense of dynamism not as apparent as it would be with the present progressive, a construction which can be said to evoke ideas of immediacy, being in the moment and taking an active role in doing something. As noted with the past tense, there is an element of distality; it could thus be interpreted that females are more often represented as taking a background role – whose actions are supplementary, but not central, to the content in written texts – than males do, and that women are therefore more frequently distanced from present events relative to the speaker's point of view.

Returning to the bias trajectories in Figure 4, the perfect construction has already been noted for its remarkably low level of bias throughout the decades – albeit with a general inclination to follow the female pronoun – compared to other constructions, at least until the 1970s whereupon the relatively unbiased construction gradually begins to exhibit increasing male bias. Taking the notion that the perfect is another realisation of the past tense in which the event discussed is either of relevance in the present or takes place further in the past than the time of reference from the speaker's perspective, one could indeed expect that the interplay of past time and present relevance results in a weak female bias tempered by the influence of the present, marked by male bias. As done with the progressive in Figure 5, it is possible to separate the perfect query into the present perfect and past perfect to better observe the effect of the tenses. The trajectory chart is given in Figure 6.

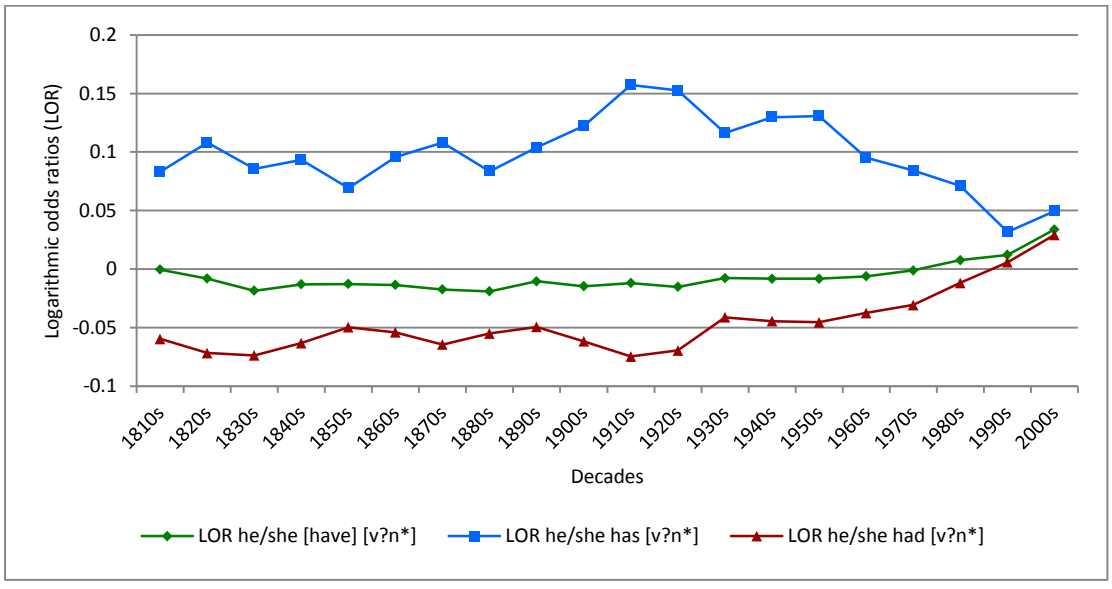

Figure 6. Log odds ratio trajectories of the present perfect and past perfect constructions from GBA data.

A somewhat similar pattern to that found for the progressive is identifiable here. With the perfect, however, the present perfect consistently displays male bias in all the decades surveyed. The past perfect, on the other hand, is inclined to take *she* rather than *he* until the 1990s when it starts to exhibit male bias at a low level; in the 2000s, the odds that *he* precedes the past perfect for the decade is about 7% greater than the female pronoun appearing before the construction. The present perfect trajectories suggest that past events of relevance to the present – from the speaker's point of view – are more likely to have been done by male figures, which relates to the proclivity of males to be depicted as doing things in the present or being relevant in the now. It could perhaps be the case, therefore, that males are more frequently made relevant than females in texts. This contrasts with the general tendency of the past perfect to take the female subject pronoun. One interpretation could be that, considering that the past tense already introduces a sense of distance from the present and that events in the past perfect are effectively doubly in the past, females are – or were – considerably more likely to be removed and distant from the present and less likely to be represented as having done things of immediate or current relevance, existing in the background. Perhaps it is the case that longstanding perceptions of women – as beings expected to be invisible or hidden, passive actors relegated to background roles and not arousing attention – within the context of a predominantly patriarchal society subconsciously determine these factors or have done so until in recent decades. These, once again, are only speculations and there must most certainly be additional factors behind these tendencies.

As for modal constructions, the fact that they collocate more strongly with the male pronoun from the 1810s to the 1970s is somewhat surprising. One might expect that since modals are said to be more associated with women and their speech, the odds that *she* precedes the constructions would be higher. This does happen starting in the 1980s, but the level of bias from then on is quite negligibly low that one could say the construction does not exactly display much bias. The finding does not counter the claim that women use modals more frequently than men as the data utilised here do not concern spoken language and the authors' genders are unknown. It was nonetheless hypothesised that data might show a higher prevalence towards *she* being associated with modal constructions due to the notion of uncertainty, non-factuality and potential reality inherent to modals which distance them from not only the present but also reality, which ties in with female figures in texts being removed from the male sphere of certainty and determination. This is certainly not – or rather has long not been – the case from what is visible in Figure 4. Furthermore, as deontic and epistemic

modalities are not distinguished from the dataset, it is also not possible to determine whether men are depicted as being obliged or able to do something. It could be that the idea of potentiality, possibility and capability – as opposed to the distancing effect – expressed by modals attracts the construction to *he*, but this is impossible to ascertain.

With constructions of primary negation, the shift from female to male bias is a peculiar one. Why negative constructions show proclivity towards *she* in the period 1880s-1930s before exhibiting constantly increasing male bias from the 1940s onwards most likely cannot be adequately explained. All one can say based on the data is that women were more likely to be depicted as not acting, doing, being allowed to or having to do something in published material in the decades immediately preceding and succeeding the turn of the 20th century, with the situation and tendency changing to encompass male rather than female figures starting from the interwar period. To wildly speculate, perhaps it has become more popular to simultaneously portray the loss of male agency and the deviation from conformity since the 1920s, when female bias in constructions of primary negation begins to diminish, maybe owing to shifts in social attitudes and sensibilities following war and economic downturn. Negation is marked, after all, and the construction queried is far less common than the positive counterparts. In that case, being written as not doing something or being contrarian in some way suggests that this aspect is deemed out of ordinary and thus something to be foregrounded. It may be that female figures acting against expectations was of greater interest to write about in the late 19th to the early 20th century before the focus shifted to men acting unusually. Of course, other more sensible determinant factors must exist for the negative constructions, but ultimately, the best one can do for this thesis is to simply present and observe the trends and trajectories while treading cautiously when attempting to make sense of global levels of bias.

Now, it should be mentioned at this point that the queries used to gather data as listed in Figures 3 and 4 are such that all instances are considered, meaning that the present, past, *will*-future, progressive and perfect constructions encompass the totality of constructions utilising the structures. The present and past queries therefore include not only the simple tenses but also all other constructions in the present and past, comprising negated constructions, progressives, perfects and passive constructions; the future query returns the latter three constructions as well. The progressive query includes cases of the going-to future as well as passive progressive constructions. The perfect query likewise covers instances of passive

constructions. A remedy would be to edit the tags to solely cover main verbs rather than all verbs, which has the unfortunate effect of removing all instances of verbs which are used as auxiliaries even when they can be used as lexical verbs as well. The verb tags in GBA are, alas, imperfect. Running the queries again after implementing the corrections to see if they cause any significant changes, the bias trajectories of the constructions with some corrections are presented in Figure 7.

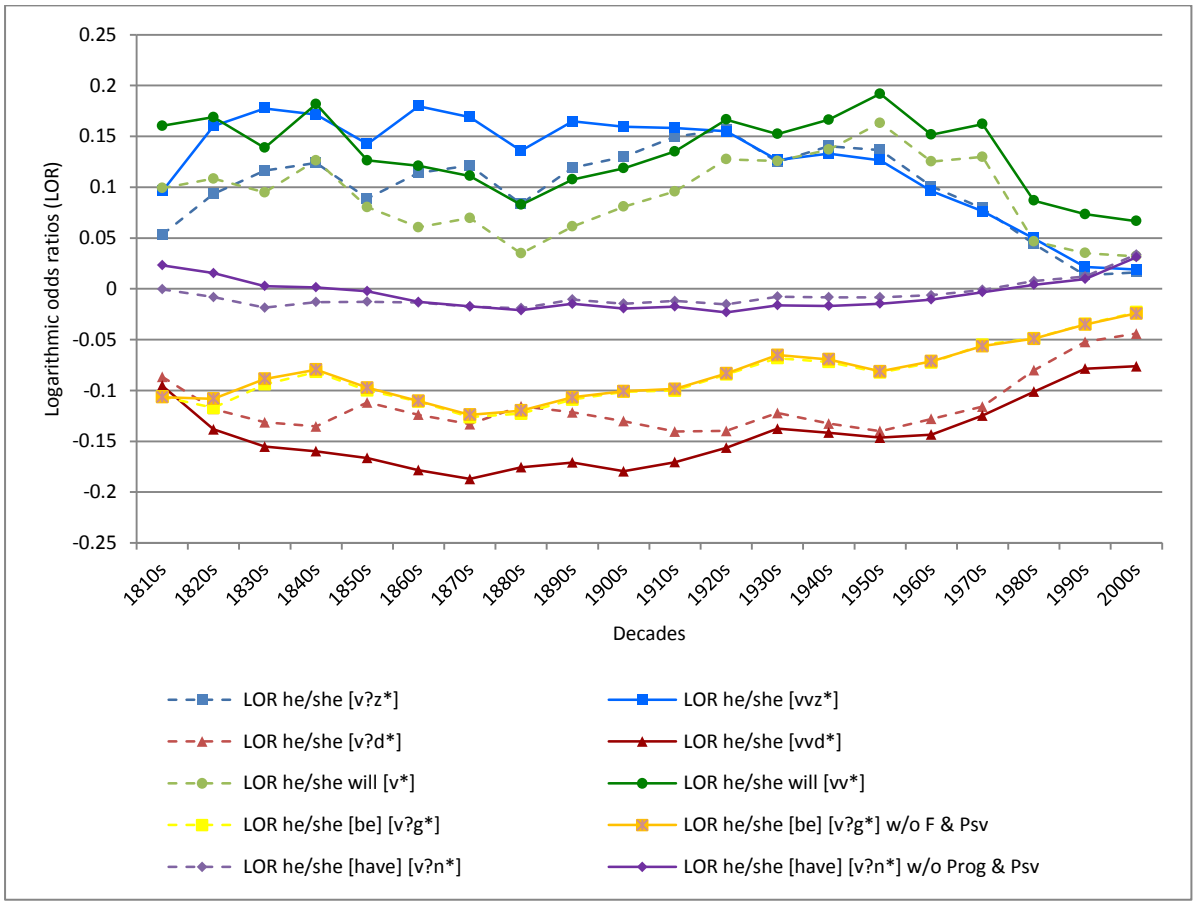

Figure 7. Log odds ratio trajectories of various constructions and corrections applied from GBA data.

What is noticeable in the trajectories of Figure 7 is that the progressive constructions – even when excluding the going-to future (F) and passive (Psv) instances – do not register much of a difference and instead follow almost identical patterns except in the 1820s and 1830s. The trends for the perfect constructions, too, are fairly similar even when accounting for perfect progressive (Prog) and passive perfect (Psv) constructions, deviating primarily in the first half of the 19th century – there is a male bias in the 1810s and 1820s with the corrected perfect constructions, followed by a female bias for about a century from the 1860s until the 1960s.

The tenses deviate quite significantly, however. Limiting the search to main verbs with the past tense shows that the simple preterite displays considerable female bias, constantly stronger than when all the past tense constructions are combined; the odds of *she* occurring with the main verb past tense construction are almost 54% greater than the odds *he* does in the 1870s, for example, compared to 36% without the correction applied. With the present tense, the gap between the levels of male bias of all constructions in the present versus lexical verbs in the present becomes bridged in the 1920s; prior to that, the simple present exhibits stronger male bias than the totality of present constructions. From the 1920s onwards, the two present tense trajectories do not show much deviance from each other. The *will*-future tense trajectories are conspicuous for the reason that the construction of *will* followed by a main verb, i.e. the construction excluding unwanted results, displays stronger male bias all throughout the decades analysed than when *will* is used in connection with all verbs. It is interesting to see that the two *will*-future constructions have superficially parallel bias trajectories; the implication is that composite constructions utilising *will* – future progressives, future perfects and future passives – are an equaliser of sorts, proportionally either exhibiting lower levels of male bias or being more inclined to take *she* over *he* as a subject in equal distribution in all decades surveyed. In any case, even with the corrections applied, the most male-biased tense construction of the 20th and 21st centuries remains the *will*-future while the past tense is more strongly associated with *she* than *he* across 20 decades. Thus, observing the trajectories, one can assume that auxiliaries introduce some kind of equalising effect on the exhibition of bias, i.e. auxiliary constructions are possibly less biased.

Since the progressives and perfects do not exhibit radically different results with the correction, the original tags for queries were retained as such for the analysis of select verbs. As previously mentioned, the tags utilised for the present and the past with the verbs were already chosen to exclude auxiliaries. For the future construction, it was decided to keep the [v*] tag rather than to use the [vv*] tag in order to avoid false negatives and to be on the safe side, considering that the future tense trajectories employing either tag share similar paths and mainly differ on the level of bias which leans towards the male pronoun in either case. The succeeding subsection deals with the findings and analyses of specific verbs and their various constructions.

5.2. Gender bias associated with specific verbs and their constructions

Continuing from the findings gathered and presented concerning overall or global gender bias displayed by select constructions, the results pertaining to gender bias associated with the verbs from Table 1 and their constructions can additionally be explored in greater detail. As with the previous subsection, the LORs and ALORs as measures of bias are of use in this part. By looking at individual verbs and constructions featuring the said verbs, one can ascertain the level of bias displayed by the verbs themselves and the effect of the constructions on the realisation of gender bias, complementing the findings for global gender bias.

5.2.1. Gender bias as expressed by verbs

To begin, it would be interesting to observe just how biased towards taking *he* or *she* the verbs themselves may be. This is admittedly neither the primary focus nor intent of the study, but it nonetheless provides ample opportunity for comparison with results from previous research works. While the constructions considered do not comprise the totality of constructions possible with the verbs, it was thought that they would nevertheless provide a general idea about gender bias seemingly 'inherent' to the verb. To do this, one would simply need to calculate the average of the resultant average LORs across all the 26 constructions selected for analysis for all 20 decades and repeat the process for every verb from the summary tables. The caveat here is that quite a number of verb constructions are not attested – most often due to the verb not being an especially frequent one – in GBA, although empty cells are still included in the computation of the averages.

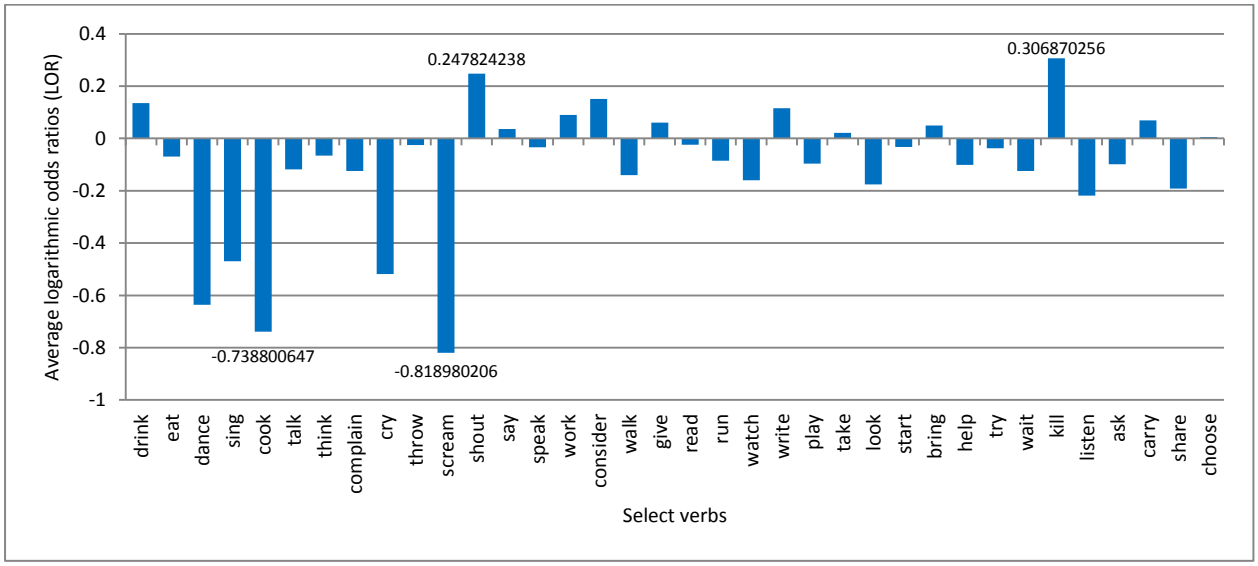

Figure 8. Total average log odds ratios for each verb obtained from the average LORs of the verbs as used in the 26 constructions under analysis across the 20 decades surveyed.

Getting the total average LOR values of each verb by averaging data from all 20 decades surveyed, one can produce a chart like in Figure 8. Once again, with negative values indicating female bias and positive values indicating male bias, certain observations can be made. The LORs in Figure 8 obtained by averaging all the ratios for constructions – both positive and negative ones – reveal what is generally already known.

Indeed, it has already been stated in previous research (e.g. Pearce 2008; Jockers & Kirilloff 2016; Norberg 2016) that certain verbs are more likely to appear with female nouns and pronouns and others with male ones; for example, *kill* collocating strongly with *he* and *dance* being more likely to occur with *she* are not surprising. The same can be observed with verbs like *scream* and *shout*, respectively exhibiting female and male bias, and others like *cry*, *sing* and *talk*, which were likewise expected to display female bias going by previous research as well as stereotyping tendencies. The idea that women are more likely to be thought of or depicted as being passive or subordinate individuals is reproduced when examining the average LORs of verbs such as *look*, *watch*, *listen*, *ask*, *wait* and *share*; whoever *she* refers to in the texts, one can speculate that they are observers more than doers; are made to listen to what is said; are more likely to ask (for) something, suggesting that the female figure requires approval or permission; are more inclined to be depicted as stationary beings; and are portrayed as being more likely to share, perhaps in both literal and metaphorical ways, having to give up a part of something of their own.

Even the disparity between the act of *reading* and *writing* is visible; men are statistically more likely to be depicted as writing – an active pursuit – whereas women, albeit at a very low level of bias, are more likely to be featured in texts as reading. A similar discrepancy can be found for *think* and *consider*; while both verbs denote internal processes, the former is associated with female bias and the latter with male bias, whereby the female bias is weaker. Speculatively, it may be that just thinking is construed as a passive act whereas considering something requires more power to accomplish and hence more apt for men to do. Both *giving* and *taking* are also seen as being something men are more inclined to do, although not strongly so. Rather intriguing is the practically negligible bias exhibited by *choose*; one hypothesises whether this is because the freedom of choice is represented as being given to everyone regardless of gender. It is also notable that bar a few exceptions, the levels of gender bias are generally fairly low – though statistically present – for the verbs surveyed.

However, in addition to corroborating what others have already found, some fairly unexpected findings are identifiable. Verbs denoting some form of physical activity – such as *run*, *walk* and *throw* – were expected to associate more strongly with *he* rather than *she*; this is not the case here, with the odds of *run* and *walk* occurring with *she* being 20% to 38% higher than the verbs being preceded by the male subject pronoun. Although the case of *throw* is more uncertain due to rather weak bias, the verb does demonstrate some proclivity towards appearing with *she* over *he* as a subject. A cursory look at GBA on what follows *throw* reveals that both men and women throw glances, stones, parties and tantrums, but that it is more frequent for males to throw a stone or a rock whereas females, when portrayed as throwing something, more typically throw glances, suggesting that throwing an item as a physical act is still more typically viewed as a male endeavour.

Of the other verbs, *say* patterning more strongly with *he* than *she* – bearing in mind that the degree of bias is rather low – is also quite interesting, considering that *talk* and *speak* show closer associations with the female subject pronoun. One supposes that since *say* is often used as a reporting verb and is common in narratives, and *he* appears in the corpus at a greater rate than *she*, it is expected for *say* – if used neutrally – to have higher odds of occurring with the male pronoun. One can also see that *eat* shows stronger preference for *she* than *he*. That *drink* collocates stronger with *he* is not unusual considering that the act of drinking in the corpus may very well include drinking in the sense of alcohol, but why *eat* should exhibit female bias is not readily explicable. In a similar vein, *start* is another example of a verb displaying female bias for which an explanation is difficult. It should be taken into consideration, however, that the levels of bias for many of these verbs are rather low and nowhere near as a significant as other verbs exhibiting high levels of bias such as *scream*, *cook*, *dance*, *cry*, *sing*, *kill* and *shout*.

We have so far looked at the average bias displayed by each verb selected for study. Since the data were collected from a diachronic corpus, the question can still be asked whether particularly biased verbs have consistently been biased or whether they have become less or more biased over the course of two centuries. For that purpose, it is sensible to produce a line chart detailing the bias trajectories of select verbs with total average LORs of at least ±0.1761.[16] The cut-off point demarcated by the value is arbitrary, but it was chosen since the score indicates that the odds are roughly 50% greater for a verb to favour one gendered

[16] In other words, ALORs of 0.1761.

pronoun over the other on the basis of all the constructions considered, making the verb sufficiently biased with regard to gender in this context. This chart is presented in Figure 9.

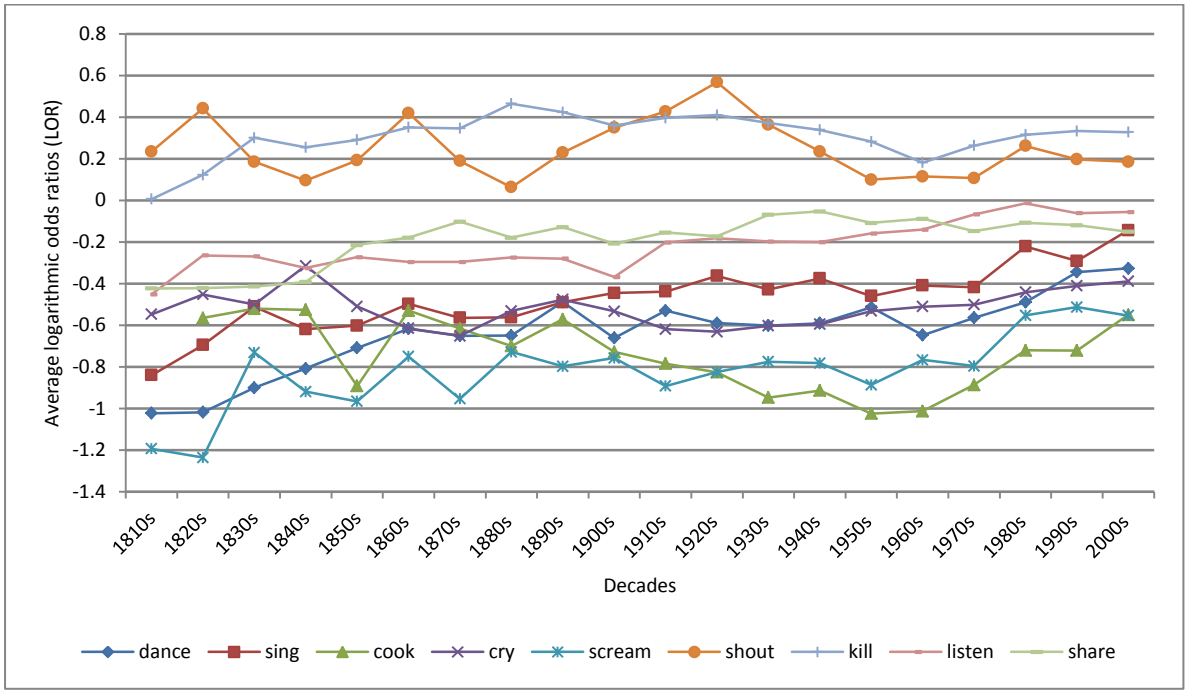

Figure 9. Average log odds ratio trajectories of select verbs taking all 26 constructions into consideration.

An observation is to be made here: of the verbs selected showing the strongest male or female bias in Figure 9, the vast majority of the ones chosen for analysis display female bias, an aspect which is also identifiable in Figure 8 when all 36 verbs are considered. In any case, female biases – for the most part – appear to have diminished in the span of two centuries going by this small selection of verbs, gradually approaching neutral use. The clearest reductions can be seen in verbs such as *dance*, *sing* and also *scream*, the latter of which is known to significantly favour the female pronoun over the male one as a subject. Other verbs like *listen* and *share*, too, have become less biased over time. With *cry*, the degree of reduction – while existent – appears to be rather minimal. The trajectory is rather erratic in the case of *cook*. After having become more biased in usage between the 1890s and 1960s – suspected to be the case due to the notion of female domesticity prevalent during that time – the level of female bias has since reduced; regardless, *cook* remains one of the most biased verbs in this selection as of the 2000s. In fact, all the verbs chosen for closer inspection here have retained their preferences for either the male or female subject pronoun; verbs with female bias and male bias continuously display their biases across the decades. The same is observable with verbs exhibiting male bias: *kill* and *shout*. For both verbs, their LORs in the 2000s are lower than they were at their highest values in the 1880s and 1920s, respectively.

Overall, however, the two verbs preferring the male pronoun do not exactly show a significant reduction in their levels of bias when viewed aggregately. That means that *killing* and *shouting* are arguably still viewed as acts more likely to be done by males, similarly to how females are more inclined to, for example, *cry*, *scream* and *cook*.

Another point worth mentioning is that in Figure 9, there is no 1810s data point for *cook*; this anomaly is because in the 26 constructions being scrutinised, there is a mismatch with the pronouns when it comes to *cook*. There actually are, in fact, attestations of *cook* with certain constructions in the 1810s; the issue is that when the construction appears with *he* or *she*, there are no attestations with the counterpart pronoun in either case, leading to errors that do not allow for LORs to be calculated. Indeed, there are numerous instances of empty data cells with some constructions, mostly with the negatives; for example, data are already quite sparse for negated present progressive constructions in early decades for GBA such that, in this research, the 1810s and 1820s cells for the construction are blank in the summary tables even when considering all 36 verbs. This is lamentable, of course, but nothing can actually be done with what is present and attested for both gendered subject pronouns.

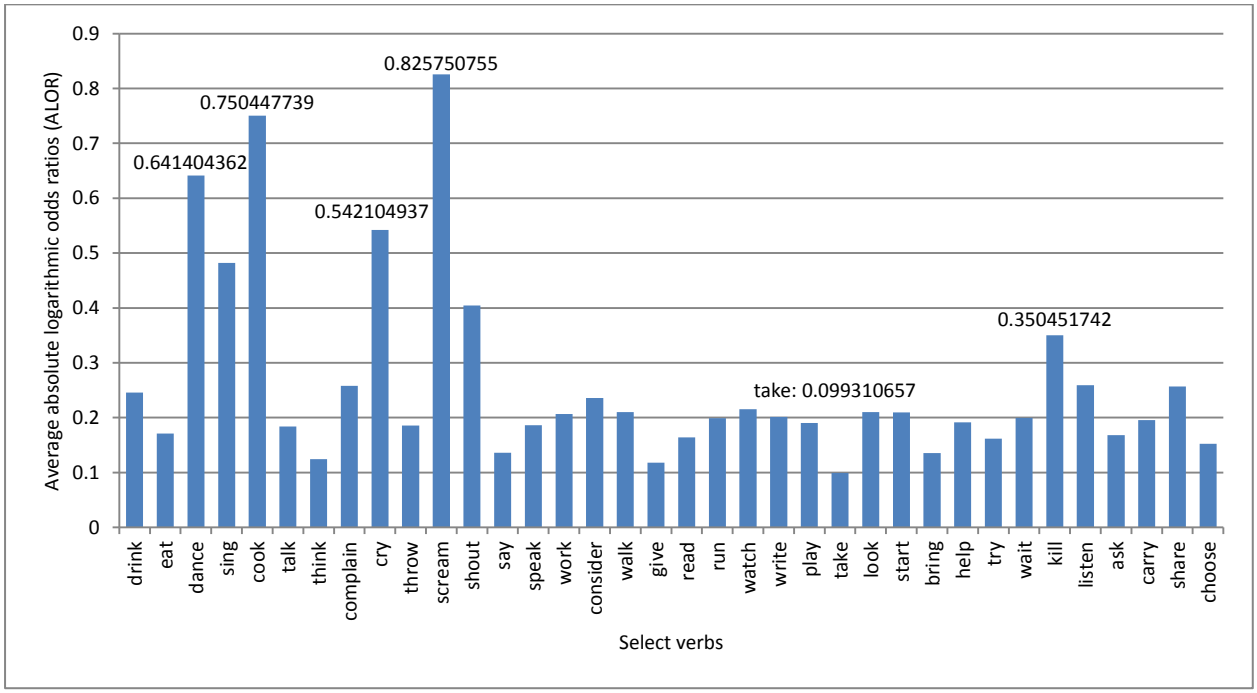

Figure 10. Total average absolute log odds ratios for each verb obtained from the average LORs of the verbs as used in the 26 constructions under analysis across the 20 decades surveyed.

While the LORs for this research are immeasurable help in determining the directionality of gender bias, i.e. whether a construction or verb favours taking the female or male pronoun as the subject, as already pointed out, averaging these ratios could lead to misleading results

when both positive and negative values are subsumed under the average. Basically, male and female biases can understandably cancel each other out. To counteract this issue, ALOR values – which technically obscure the gender aspect of bias – can instead be used for averaging to quantify bias. Replicating Figure 8 with ALORs, a different picture emerges in Figure 10.

By the ALOR metric in Figure 10, *choose* is not the least biased verb; *take* is. This indicates that *take* is generally marked by low levels of bias in either direction, which is perhaps unsurprising considering that the word is frequently utilised as a light verb. The apparent equalising effect of auxiliaries has previously been noted on the basis of Figure 7; with the understanding that light verbs share features of both main verbs and auxiliaries, light verbs would be expected to exhibit weaker bias than lexical verbs. The LOR bias trajectories of *take* and *choose* as shown in Figure 11 confirm that *take* generally displays low levels of bias, viewing the 1810s decade as an outlier.

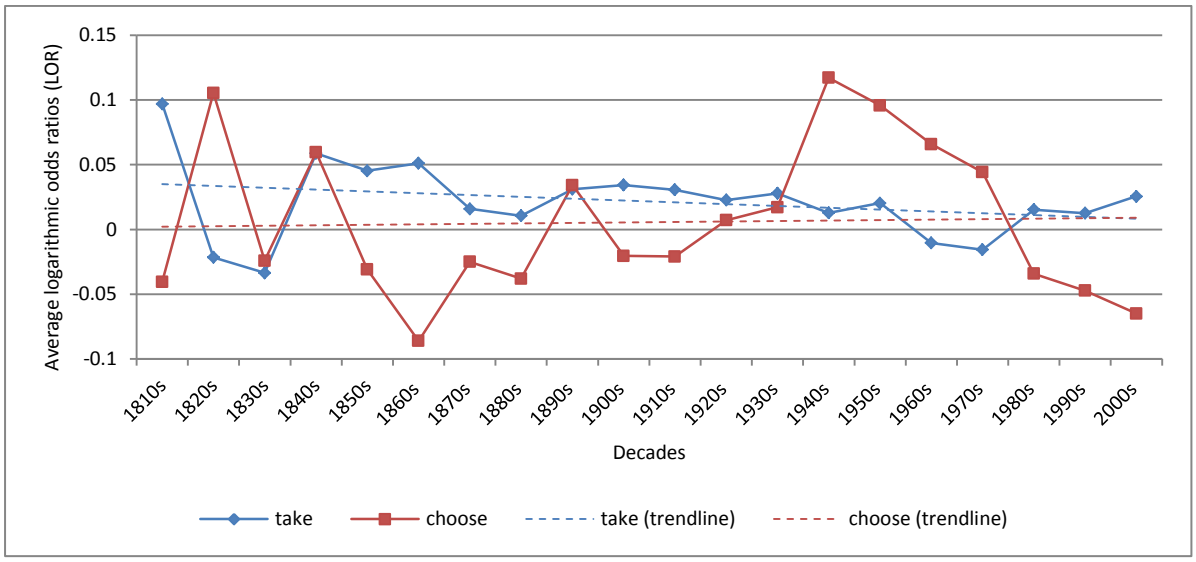

Figure 11. Average log odds ratio trajectories and accompanying trendlines for *take* and *choose* taking the chosen 26 constructions into account.

As can be seen, the trajectory of *choose* oscillates somewhat and is punctuated by peaks and troughs; the corresponding values of each data point tend to even each other out when averaged, leading to the exceptionally low average LOR score of *choose* in Figure 8. In comparison, the trajectory of *take* appears much more orderly from the 1840s onwards. Indeed, after the 1810s, the log odds ratios for *take* stay within the range of ±0.06, translating to at most 15% greater odds that one pronoun or the other precedes a construction containing the verb *take*, which indicates fairly weak bias. Also visible is how the trendline for *take*

registers a decline over 20 decades, demonstrating that (mostly male) bias has decreased overall for the verb, at least when it occurs with the constructions under consideration. In contrast, the trendline for *choose* shows a slight uptick across two centuries, indicating that the verb has become very marginally more biased.

With all this information in mind, one can repeat the process of diachronically charting the ALOR values of verbs deemed to be particularly biased as was similarly done with LOR values in Figure 9. For Figure 12, verbs whose average absolute log odds ratios across 20 decades averaged to at least 0.24304 – signifying a minimum of roughly 75% greater odds that the verb in combination with the 26 constructions occurs with one gendered pronoun or the other – were selected for inclusion. Along with those verbs, *take* and *choose* were also included to visualise their low rates of bias and, continuing on from the previous figure, to additionally provide ALOR bias trajectories of those two verbs.

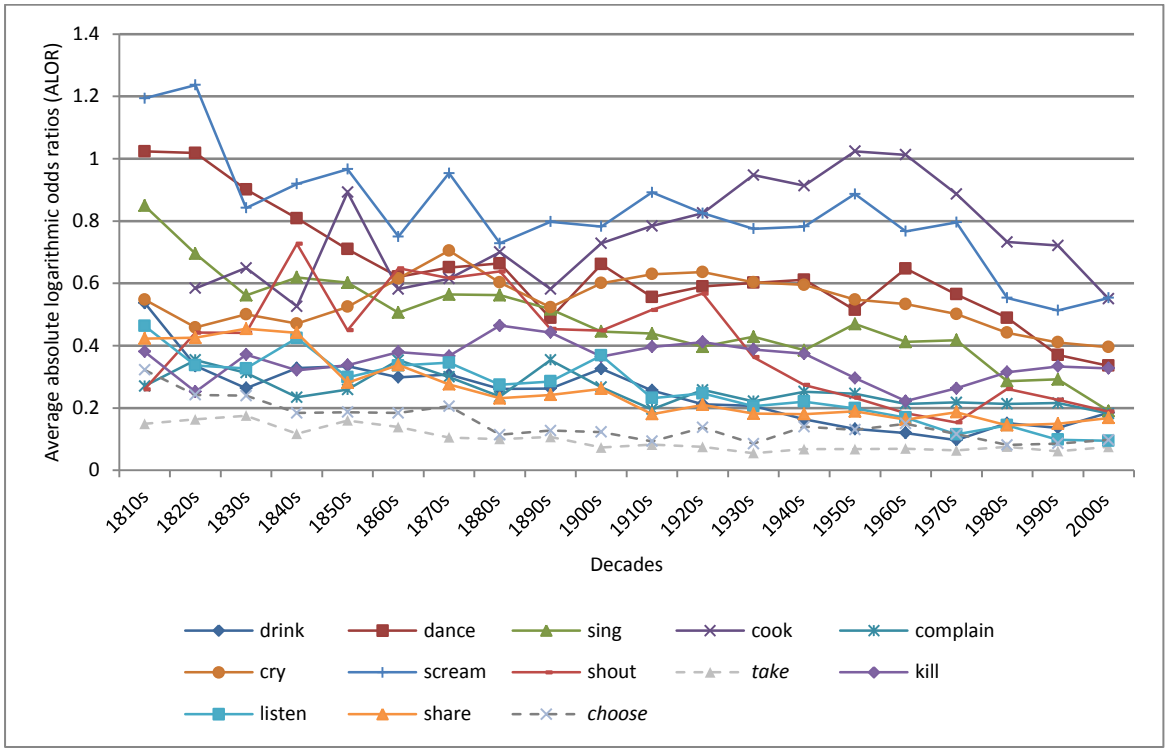

Figure 12. Average absolute log odds ratio trajectories of select verbs taking all 26 constructions into consideration.

It can be observed from Figure 12 that the general trend is for the selection of verbs to display lower levels of bias over time. As noted before with LOR values, verbs such as *dance*, *sing* and *scream* gradually register weaker biases. Some verbs, on the other hand, seem to resist the trend towards equality – *kill*, for example, demonstrates an increase in bias since the 1960s. All things considered, however, one can nevertheless see that bias has become

reduced overall for the verbs presented in Figure 12 since the beginning of the 20th century. It would be entirely unwise to generalise on the basis of a few verbs, but one could still surmise with the data provided that there must be a tendency and trend towards verb usage becoming less biased in connection with gendered subject pronouns and that such an observation can be diachronically tracked and possibly confirmed using corpus data.

Data specific to this study thus show that many verbs have consistently and continuously shown bias in the male or female dimension albeit generally at diminishing rates since the beginning of the 1900s. There are some additional facets and aspects to keep in mind, however, so as to be cautious in interpreting and understanding the results. For instance, it is the case that certain verbs – *cook*, *scream* and *shout*, to name a few – are not all too frequent in the corpus, at least in combination with the intended constructions whilst taking gender into account as well. As stated before, there were many empty cells, some of which were more pronounced with certain verbs – and constructions too, as will be seen later; consequently, the findings comprising this subsection should be treated with caution. The uncertainty is compounded by the unintended prevalence of verbs exhibiting female bias as pointed out previously. Indeed, somewhat infelicitously, the majority of the 36 verbs selected for research generally exhibit female bias. This was not exactly foreseeable and, furthermore, has repercussions with respect to drawing conclusions for bias linked to constructions based on LOR values. Ideally, there would be parity such that an equal number of verbs should show preference of occurring with either *he* or *she*, but that is clearly not the case in this component of the study. While it is less than optimal, one must manage with what is available.

5.2.2. Gender bias associated with constructions

It is now time to consider and explore the possible role of constructions – when used with select verbs – in exhibiting gender bias and investigating how such trends have changed in the span of two centuries. In this portion of research, it makes sense to observe and examine the levels of bias as displayed by the constructions under analysis over time. Arguably the simplest method of going about this is by getting the average ALORs of each of the verbs per decade for each construction. These averaged values can then be charted and presented either separately or categorised together; for example, the simple present trajectory can appear alongside other simple constructions, but can also be grouped and averaged together with the rest of the present constructions, i.e. the present progressive and present perfect, to provide an

overview of all the present tense configurations taken into account for this study. Multiple combinations are possible; the most important charts are presented below.

Let us firstly take a look at the ALORs of positive simple tense constructions consisting of the present, past and future tenses in their most basic forms for each decade garnered by averaging the ALOR scores of the 36 verbs. The trajectories of the simple tenses are presented in Figure 13.

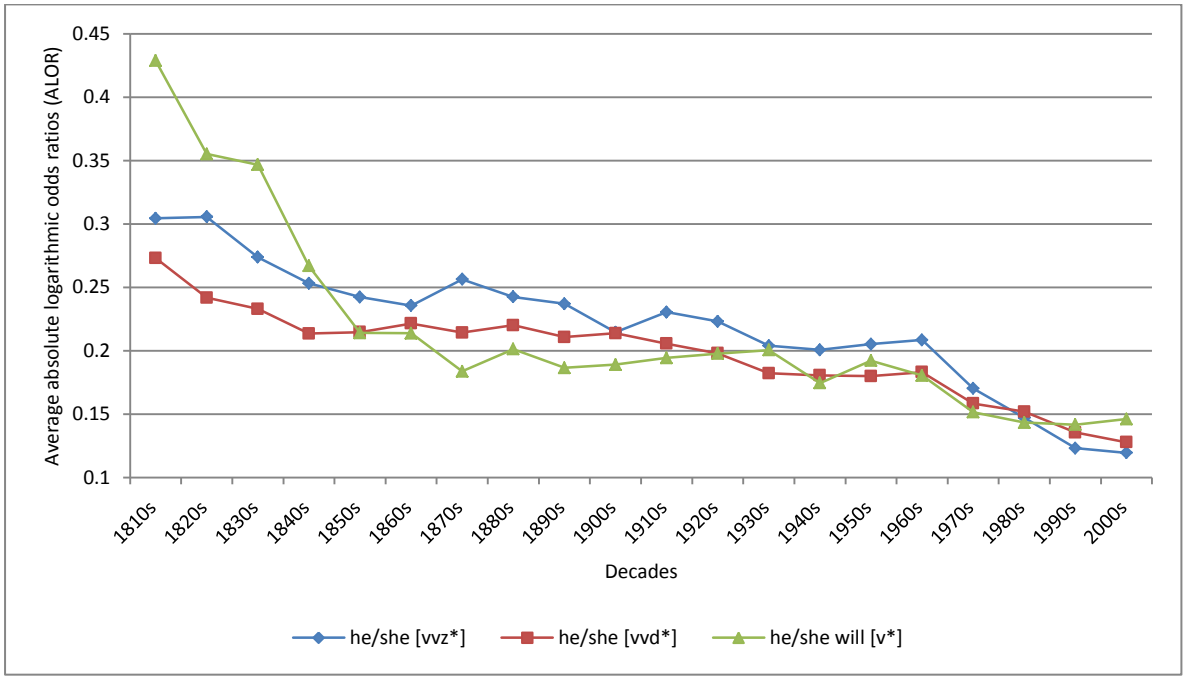

Figure 13. Average absolute log odds ratios of positive simple tense constructions with the 36 verbs over time.

It can immediately be recognised that bias has decreased for the simple tenses over the course of two centuries with a clear pattern of reduction identifiable for the present and past simple tenses after the 1970s. Another observation is that the present simple construction continuously exhibits the highest level of bias among the three constructions between the 1850s and 1970s following the precipitous drop of bias associated with the *will*-future in the span of 40 years from the 1810s to the 1850s. By the 1990s and into the 2000s, the present simple then displays the lowest bias, whereupon the future tense begins to become slightly more biased. It is also worth noting that the degrees of bias between each simple construction have diminished over time as well; the ALOR values can be seen somehow coalescing from the 1970s onwards, meaning that the constructions show roughly similar levels of bias as each other. Comparing the 2000s ALOR scores to those from the 1900s, it is evident that biases – while still existent – have become considerably weaker; a decrease from a ratio value

of around 0.215 to 0.120 for the present simple corresponds to bias being almost halved within a century. Nevertheless, an ALOR of approximately 0.120 still means that the odds the construction prefers one gendered pronoun over the other are over 30% greater with one subject pronoun.

The information above holds true specifically for positive simple tense constructions utilising the 36 verbs under investigation. There also are forms of primary negation to be considered. So, repeating what was done for Figure 13 with the negated counterparts of the simple tense constructions hitherto discussed, one produces a similar chart as shown in Figure 14.

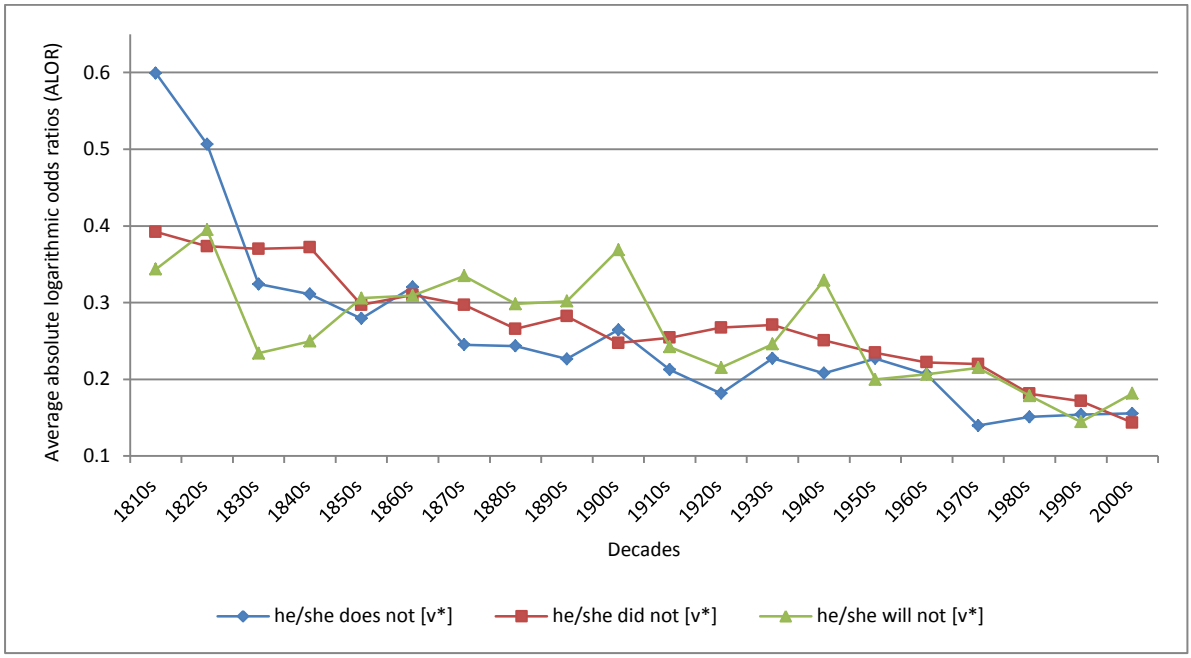

Figure 14. Average absolute log odds ratios of negative simple tense constructions with the 36 verbs over time.

With the negated variants of simple tense constructions, a similar decreasing trend for biases to that found in positive constructions is identifiable; at the same time, the negative constructions tend to display greater bias than their positive counterparts on the whole, suggesting that negation being marked may amplify gender bias in constructions. A notable difference between the positive and negative trajectories is that negative past simple constructions appear to be generally more biased than negative present simple ones; there are only a few decades in which the reverse is true. Conspicuous too is the progression of bias levels for the negated future construction which contrasts with that for the positive forms in the earliest decades; the negated present – not the negated future – is most biased in the 1810s and 1820s. One must remember, however, that negated constructions are less frequent than positive ones – for instance, out of 36 cells which can be filled for each decade (as there are

36 verbs for study), there were 34 filled cells for the positive present simple construction as opposed to 16 cells for the negated counterpart for the 1810s. Indeed, taking into account empty cells which may also explain the more erratic trajectories for constructions of primary negation, it must be explicitly stated that the findings for negative variants are less definitive, conclusive and arguably less meaningful than those for positive constructions. Still, they are worth examining, hence their inclusion. The point is that, with 7 counts to 16, fewer cells were filled for negated future constructions than for negated present forms – especially in the earlier decades – making it difficult to properly compare the trajectories between the positive and negative constructions this way.

Regardless, although caution is warranted when looking at and interpreting the trajectories in Figure 14, one may assume that the higher general prevalence of bias for *did not* constructions compared to the negative present simple has something do with the markedness of the negative construction in the past tense. It could be that, for example, the negated present states a perceived truth or common knowledge that is expected and unsurprising whereas the negated past draws attention to actions not done or accomplished which may be construed as unusual, thus making such acts more prone to being reported. This discrepancy between the two uses could then result in more visible bias for the negated past as the gender is foregrounded, demonstrating that it is unusual or unexpected for either males or females not to have done something. The unstable and teetering path of the negative future, on the other hand, is difficult to categorise. The one thing that can be surmised is that since the negated *will*-future appears – for the most part – more biased than its positive form, biased representations of a figure marked by gender determined, predicted, made or certain not to do something are more likely than those figures being depicted as being certain or determined to do the same acts.

It seems to be the case, then, that gender bias has diminished over time for both positive and negative simple tense constructions. It has been already mentioned that negative simple forms appear to be more biased than positive simple constructions. Simplifying the visualisation by averaging positive present, past and future simple ALORs together and doing the same for the negated variants, two trajectories are created as seen in Figure 15, from which it is clear that when using the verbs selected for this study, negative simple tense constructions are more biased than the unmarked positive constructions.

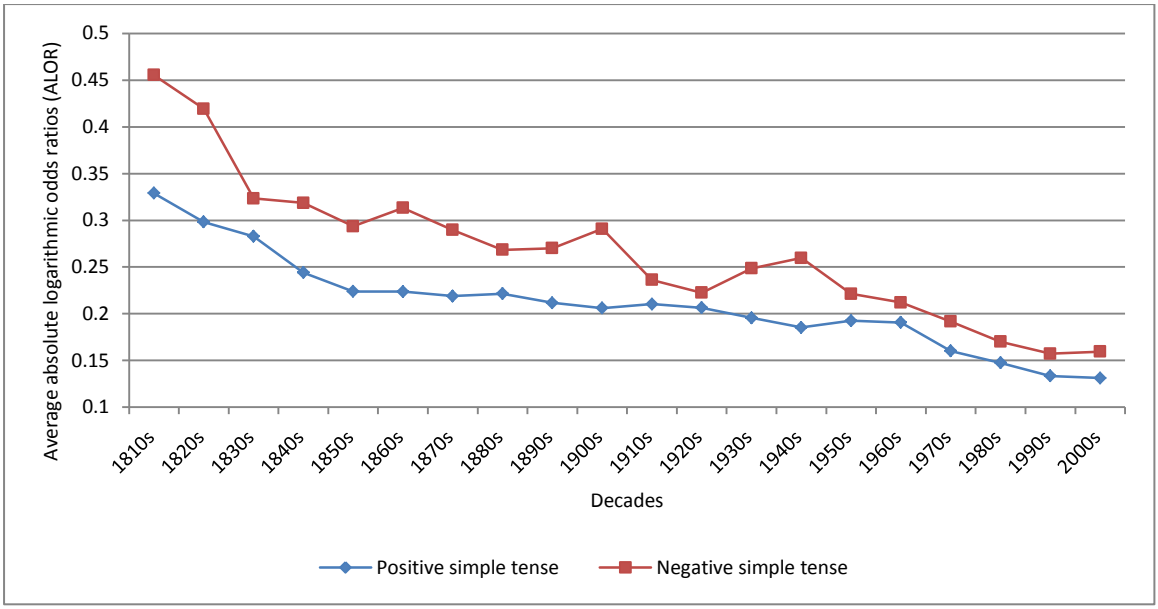

Figure 15. Average absolute log odds ratios of positive and negative simple tense constructions over time.

Another interesting visualisation would be the bias trajectories of the present, past and future simple tenses when both positive and negative forms are taken into account for each tense. This involves grouping and averaging the ALOR scores concerning polarity, i.e. positive or negative constructions. That way, one has an overview of the total degree of bias exhibited by the simple tenses or any other construction for that matter. The trajectories based on the averaged ALORs of both positive and negative simple tense constructions for each decade are given in Figure 16.

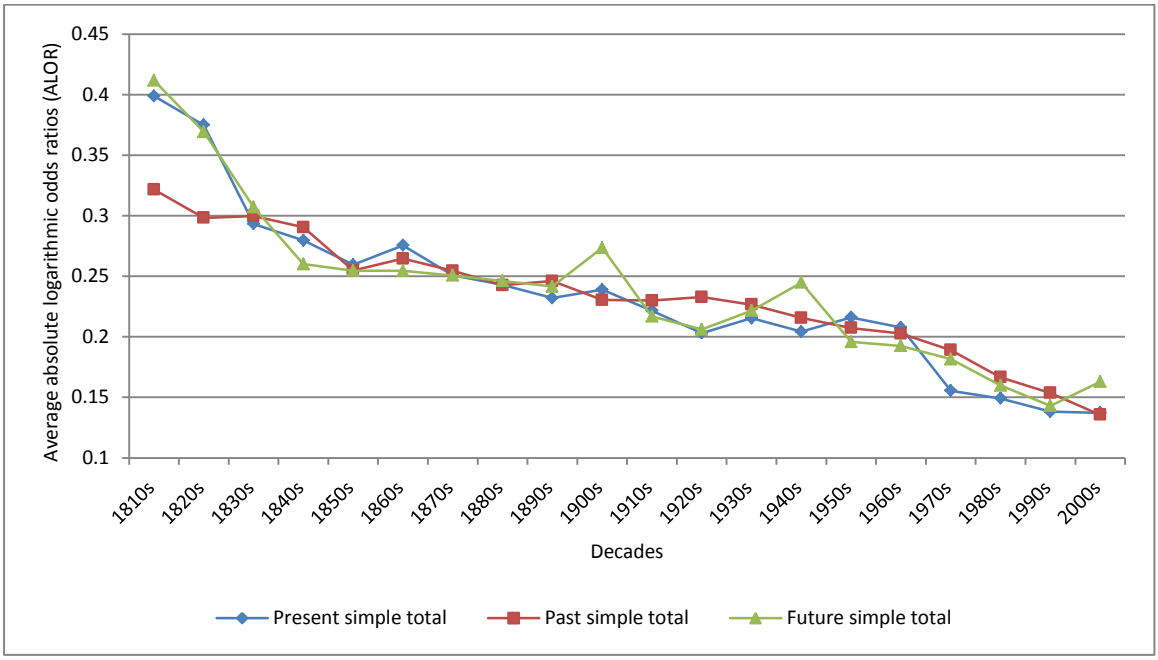

Figure 16. Average absolute log odds ratios of simple tense constructions over time.

Taking a holistic view of the simple tense constructions, it is evident that the simple tenses are similar to each other, which is unsurprising. Excepting the data for the 1810s and 1820s as outliers, there is considerable overlapping and approximating occurring with the trends displayed by the simple tenses, with roughly similar trajectories being followed. All charts show a decrease in bias over time, indicating that the tendency for simple tense constructions – regardless of tense or polarity – is towards gender-neutral usage of the forms.

There are thus three variables to be considered here: tense, negation and, technically, the construction itself. These factors are applicable to the progressive and perfect constructions to follow; the modal constructions are only examined from the viewpoint of polarity. In essence, one can ask which tense exhibits stronger bias, whether negated constructions are more biased than positive ones, and which construction is most gender-biased.

Continuing onwards, the process taken for the simple tenses can be applied to the progressive constructions in the same manner. One first produces a chart to compare the trajectories of the present and past progressive constructions with positive polarity as in Figure 17.

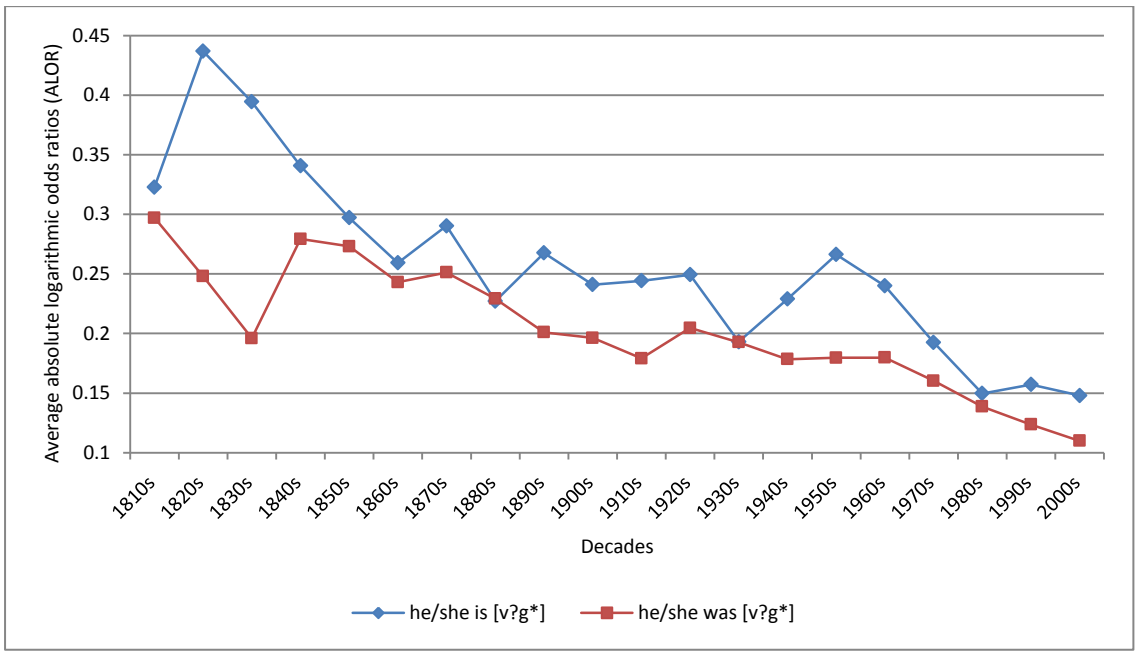

Figure 17. Average absolute log odds ratios of positive progressive constructions with the 36 verbs over time.

A gradual decrease of bias is clearly detectable also with the positive progressive constructions in this study. With the exception of the 1880s, the present progressive exhibits greater bias than its past counterpart in the decades surveyed. That the positive present tense construction is more biased than the past one has already been noted for the simple tenses as

well. It is worth remembering that, in general, past tense constructions are much better represented in GBA than those in the present tense; supposing that an equal number of tokens exist for both the present and the past, bias is heightened and amplified with constructions in the present tense. Discounting that, the other reading is that there just is a higher chance of biased depictions of figures, whether male or female, in written language as doing one of the 36 sample verb activities in the now and in progress from the perspective of the speaker. Also conspicuous is the rise and fall of the rate of bias associated with the present progressive between the 1930s and 1980s which contrasts sharply with the trend exhibited by the past progressive in the same time period; what the underlying cause may be is uncertain and will not be pursued further.

Now, to see whether the present tense bias holds true for negative progressive constructions as well, another line chart can be produced as provided in Figure 18.

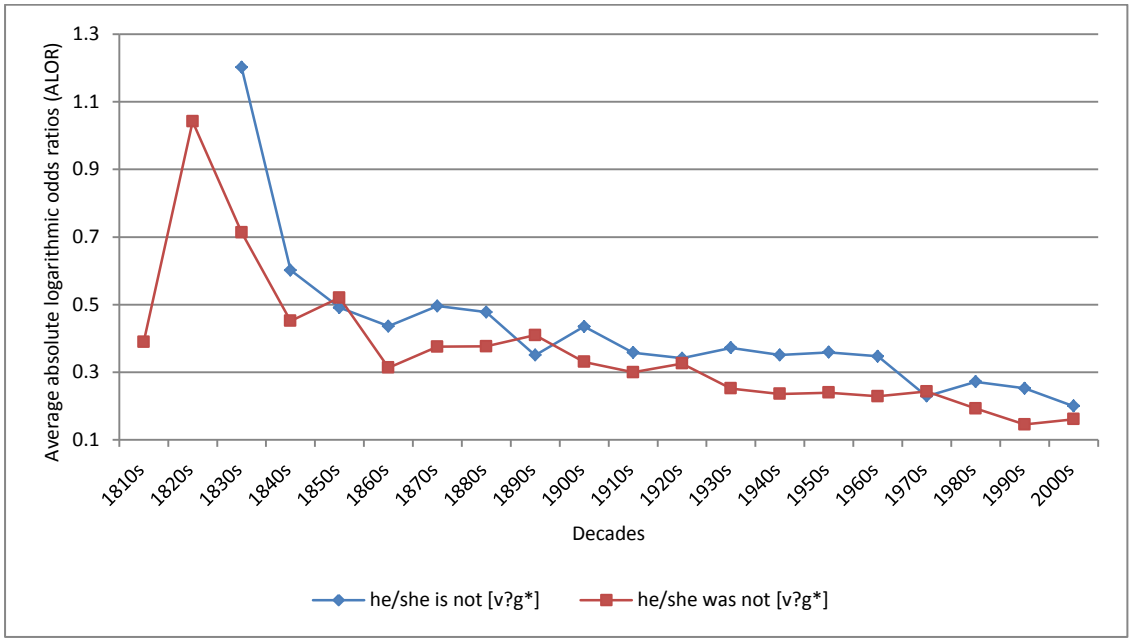

Figure 18. Average absolute log odds ratios of negative progressive constructions with the 36 verbs over time.

The aforementioned issue with the relative paucity of negative constructions is laid bare here; due to a lack of either male or female construction data – or in some cases, due to a lack of both, like the construction *he*/*she is not shouting* which is not attested at all in GBA – there actually are no data points for the negative present progressive in 1810s and 1820s. There is also only one negative past progressive verb construction in the 1810s based on the sample of 36 verbs which is attested for both genders: *he*/*she was not speaking*. Needless to say, the

findings for earlier decades must be viewed and treated with extreme caution; only from the 1850s onwards do the data and results become clearer and more readily interpretable.

While it is difficult to derive and pass judgements owing to the dearth of data availability, one can nevertheless see that the present progressive constructions with negative polarity are still generally more biased than their past counterparts. A peculiarity can also be observed for the 20th century: the gap between the present and past progressives of negative polarity closes before widening again in the 1920s and 1970s. This is interesting insofar the trends are compared to the positive progressives – in Figure 17, one can see that the gaps close and reopen with a delay of a decade, i.e. in the 1930s and 1980s. The assumption is that male and female figures became more stereotypically portrayed as actively doing – or not doing – things in narrative real-time in that timespan, but once again, the why and what associations the present progressive has with the uptick remains elusive.

The gravamen here is that trends for the progressive constructions indicate a reduction in the level of bias over the course of two centuries. This is the case for forms of both positive and negative polarity, as can be seen visually summarised in Figure 19.

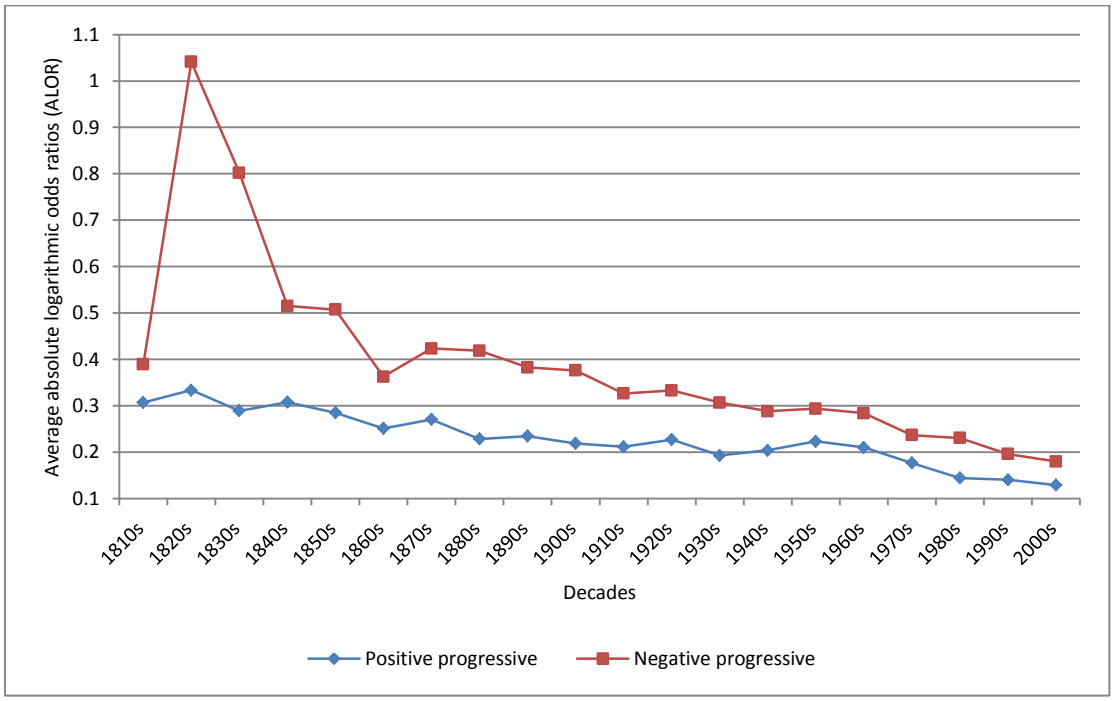

Figure 19. Average absolute log odds ratios of positive and negative progressive constructions over time.

Even when ignoring the earliest data points – since, for example, the data point for the negative progressive in the 1810s consists of just one verb construction – and instead tracking the change between the 1850s and 2000s, it is readily apparent that negative progressive

constructions are indeed more biased than positive ones. The tendency is commensurate with what has been noticed for the simple tenses as well. The discrepancy between the two construction types – the positive and negative progressives – is, however, greater than what is found in the simple tenses; there is a minimum ALOR difference between the two forms of at least 0.05, translating to an approximately 12% minimum difference in the level of bias. These values suggest that biased or stereotypical depictions of gendered figures in texts not in the midst of doing or having done something are more likely as opposed to their being portrayed as being or having been in the process of conducting an action.

This time, grouping together the positive and negative constructions for the present and past progressives separately, further trajectories can be visualised; these are presented in Figure 20.

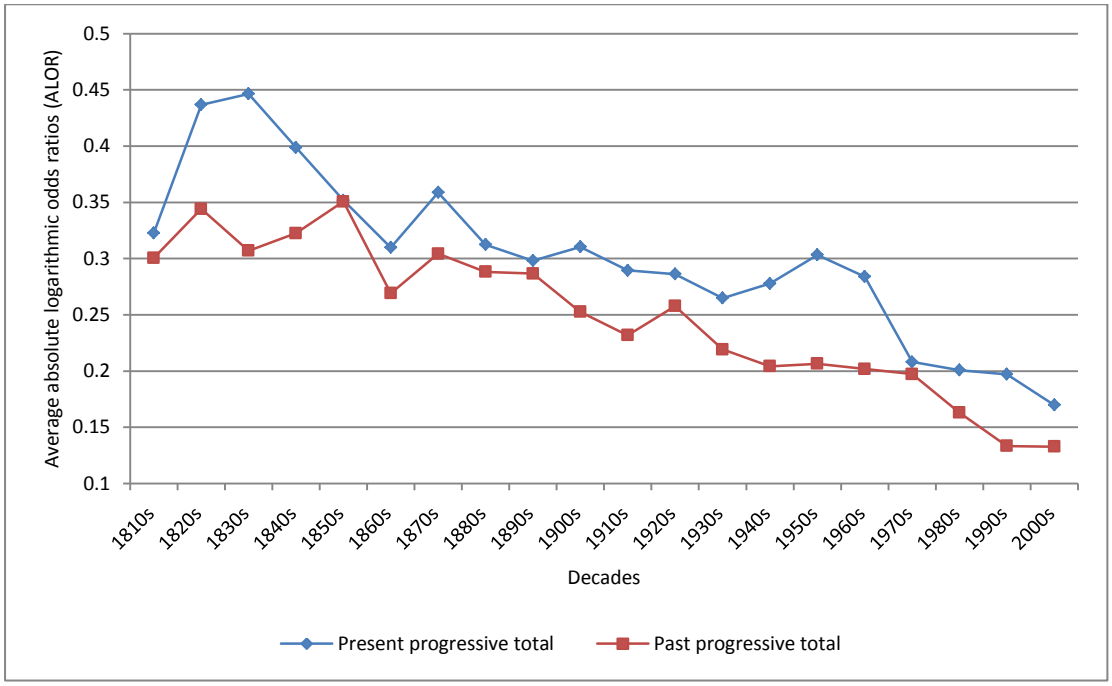

Figure 20. Average absolute log odds ratios of present and past progressive constructions over time.

Unsurprisingly, one can see that the present progressive construction generally exhibits greater bias than the past progressive. The difference between the two trajectories is somewhat more marked and noteworthy than was noticed for the simple tenses; the gaps are, for the most part, larger. The discrepancy indicates that the tense does leave an imprint on the construction with regard to the level of bias, at least in the case of the progressives.

In this study, both tense and negation are of relevance not only to progressive constructions but also to perfect constructions; bias trajectories of the present and past perfects are thus also

of interest. Identically to what has been done before, the trajectories of perfect constructions with positive polarity are presented first in Figure 21.

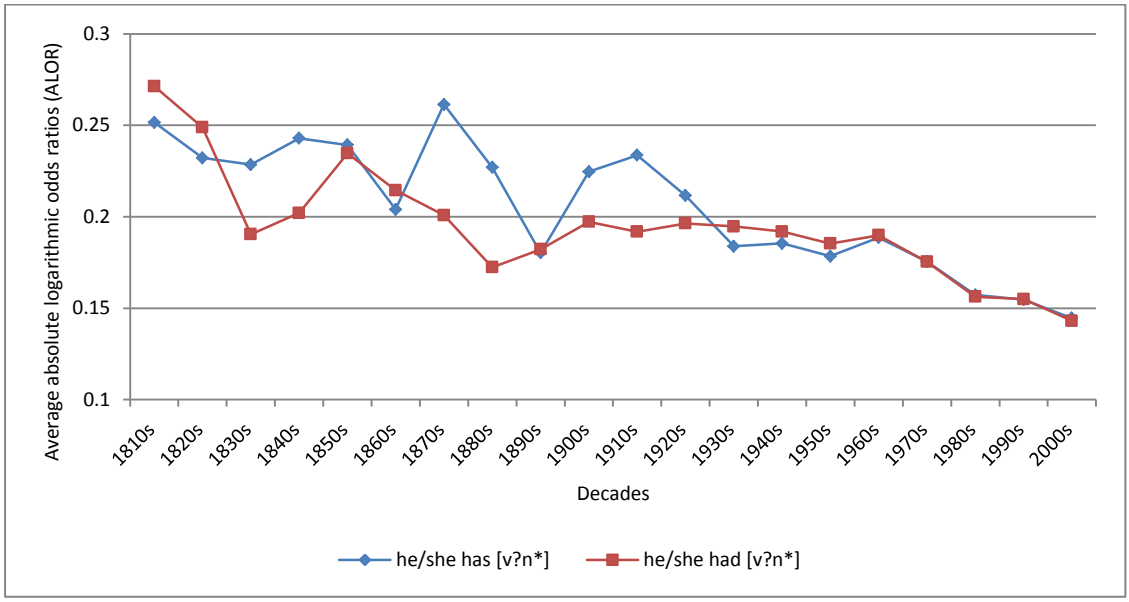

Figure 21. Average absolute log odds ratios of positive perfect constructions with the 36 verbs over time.

As with the other constructions examined previously, bias has diminished over time for both perfect constructions. The trends are fairly irregular in the first hundred or so years represented in the corpus, marked by several peaks and troughs. It almost appears as if the present and past tense perfects are competing to be more biased than the other; however, the present perfect generally happens to exhibit higher levels of bias in decades where it registers greater bias over the past perfect than when the past construction is more biased than the present one. What is especially noteworthy about Figure 21 is the fascinating phenomenon visible starting in the 1930s and made extremely conspicuous from the 1960s onwards. Simply put, the gap between the rates of bias displayed by the two perfect constructions with positive polarity steadily begins to narrow between the 1930s and 1950s before it almost entirely vanishes after the 1960s, evidenced by the overlap of the trajectories in the decades 1960s to 2000s. With nearly identical post-1950s bias trajectories, the difference in bias – as measured by the ALORs – between the two constructions is on the degree of thousandths at most; effectively, neither construction is more strongly biased than the other in the time period, at least when surveyed with the chosen 36 verbs. This striking feature is not found in negated perfect constructions as shown in Figure 22.

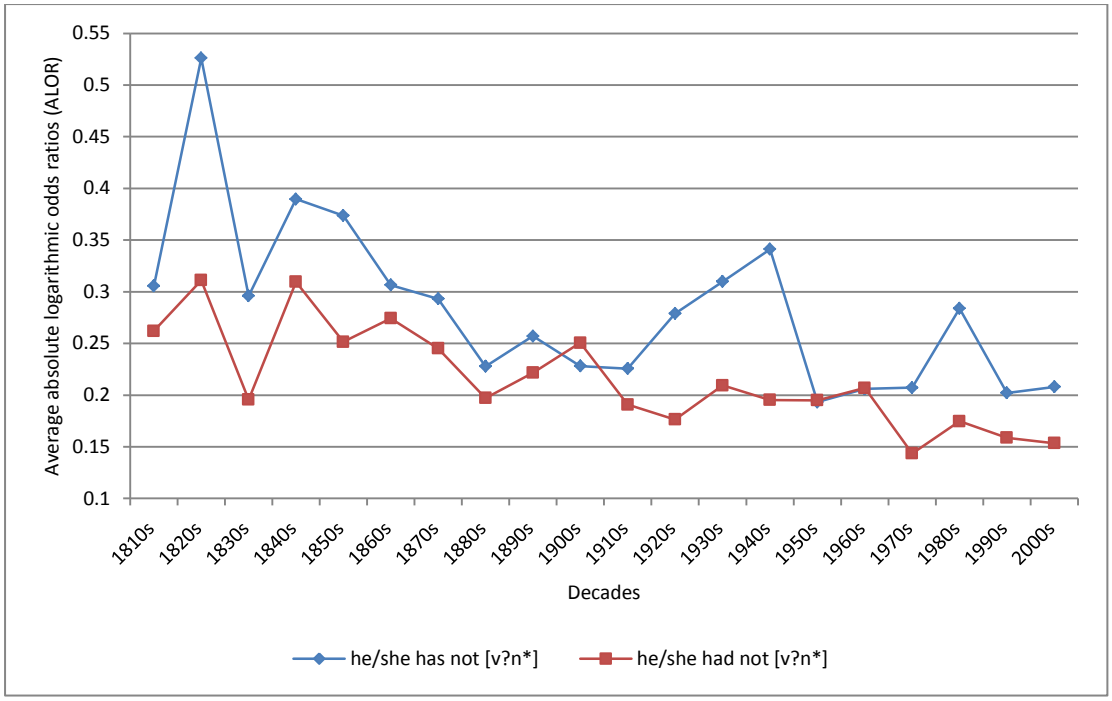

Figure 22. Average absolute log odds ratios of negative perfect constructions with the 36 verbs over time.

Indeed, except for a short period of coalescence in the 1950s and 1960s, the two bias trajectories for the negated perfects are clearly distinguishable. On the whole, the negative present perfect construction is conspicuously more biased than the past tense form. Additionally, while both the negative present and past perfects have collectively become less biased, it would seem that the former is more resistant to unbiased use, judging by the prominence of the peaks in the 1940s and the 1980s relative to the ones found for the negative past perfect construction.

Once again segregating the factor of polarity, the historical trends for positive and negative perfect variants are shown in Figure 23, where one can see that negative constructions tend to more biased than their positive counterparts.

That having been said, there are several points in which there is little difference between the positive and negative perfect constructions in terms of bias displayed, such as in the 1880s, 1910s and 1970s; in the latter two decades, the positive constructions even register statistically greater bias than the negative ones. These findings stand in contrast to what has been observed in both simple and progressive constructions for which the negative variants always appear more biased than the positive forms. Likewise conspicuous is the fairly unstable bias trajectory of the negative perfect constructions compared to the one for the positive perfects; one can see that bias has decreased at a slower and more gradual rate in the

case of positive perfects as opposed to perfects with negative polarity. This appears to be a feature common to the unmarked positive variants across simple, progressive and perfect constructions.

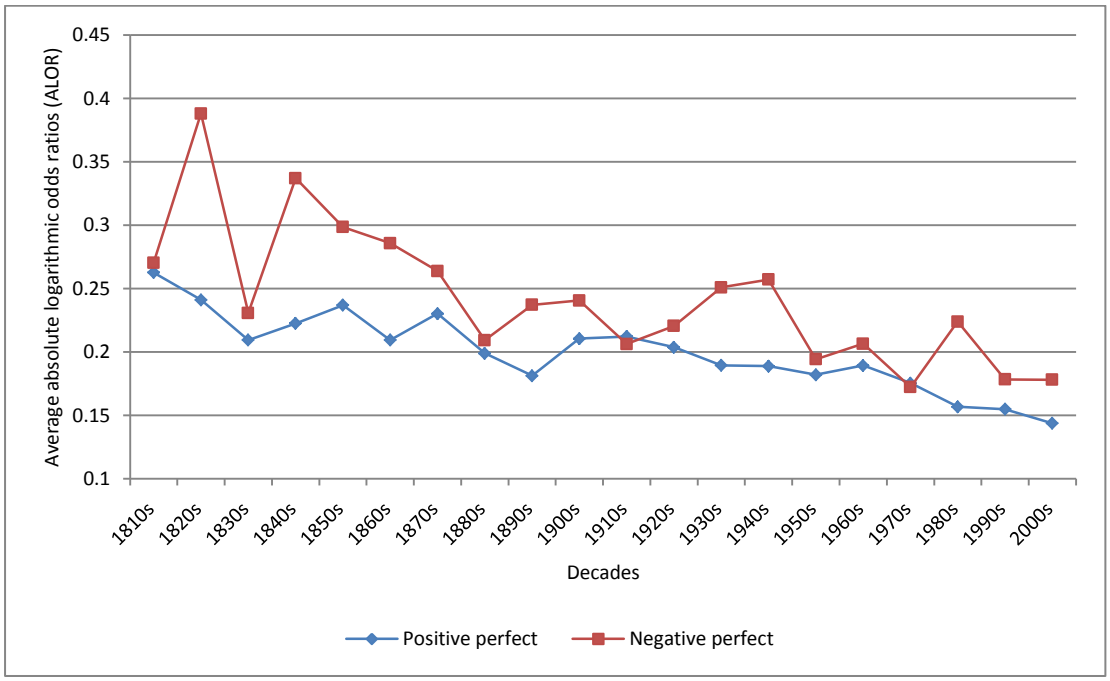

Figure 23. Average absolute log odds ratios of positive and negative perfect constructions over time.

Turning now to the tense summary of the perfect constructions, it is possible to observe how the bias trajectories of the present and past perfects have progressed over 200 years in Figure 24, taking both positive and negative forms into consideration.

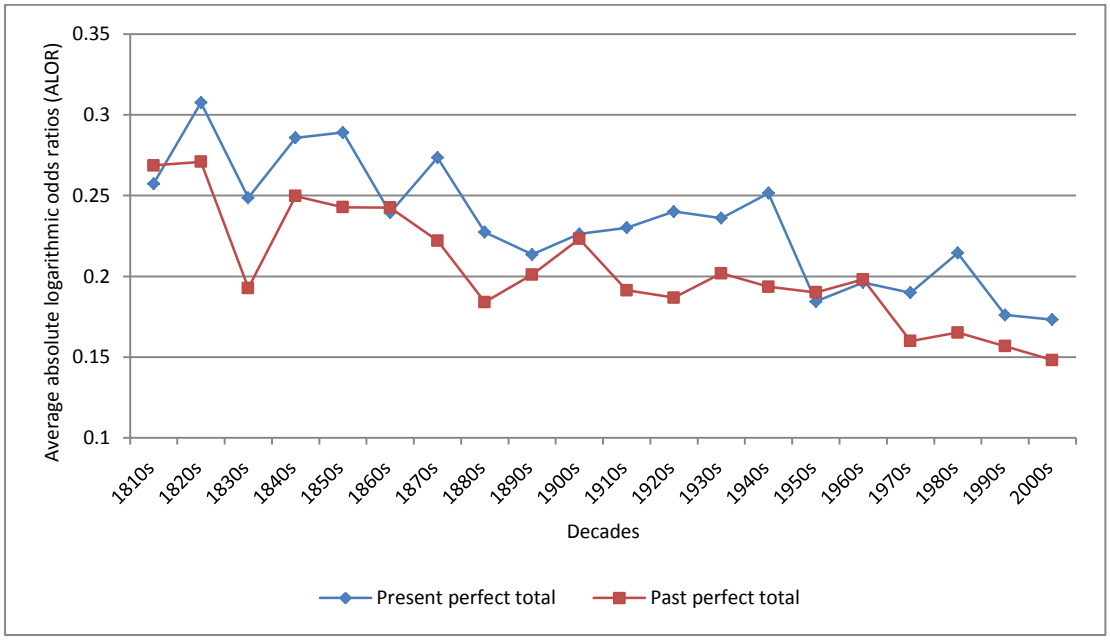

Figure 24. Average absolute log odds ratios of present and past perfect constructions over time.

Once again, the construction in the present tense tends to display more bias than in the past tense. There are certain periods in which the difference is noticeable, such as during the late 1800s, the first half of the 20th century and then from the 1970s onwards. The aspect of the post-1970s decline in bias has been mentioned previously, and indeed, the decrease is visible in many constellations where it is generally the case that the level of bias associated with constructions towards the end of the 1900s does not exceed 1960s levels. The perfect tense constructions – specifically the negative formulations – are therefore somewhat of an oddity.

The positive forms do not register a bias rebound but the negated variants do, and this is especially strong for the negative present perfect as was seen in Figure 22 and also observable in Figure 24 with the present tense bias trajectory; in comparison, the past perfect sees a steeper reduction in bias by the 1970s. Considering that the perfect trajectories have been remarked to be rather erratic, this may reflect a kind of sampling issue. It could be that a number of texts comprising the GBA for certain decades – the 1980s, for example – have, for whatever reason, an abundance of instances in which the 36 sample verbs (a selection thereof, more probably) appearing as part of the perfect constructions collocate particularly strongly with either *he* or *she* when negated and in present tense. Whether this is in fact the case for all decades and whether this points to actual variation in language with respect to gender bias or is simply a result of authorial propensity towards biased forms cannot be determined, but these aspects are something worth thinking about when utilising written language corpora.

So far, we have seen and examined bias trajectories of tensed constructions. With the modals, tense is not a determinant factor in this study – only the difference in polarity is of actual interest. Technically, of course, 4 of the 6 selected modals are already in their preterite forms: *would*, *could*, *should* and *might*. Of the two remaining ones, *can* is the only present tense counterpart represented since *must* stands alone, *shall* and *may* were excluded, and *will* is included as part of the simple tense constructions. For the most part, the auxiliaries here are therefore grouped and averaged together in lieu of being explored individually, with attention being paid to the negation aspect above all else. Figure 25 thus shows the bias trajectories of the positive and negative modal constructions, where it can once again be observed that negated constructions appear to be more biased even for the modals.

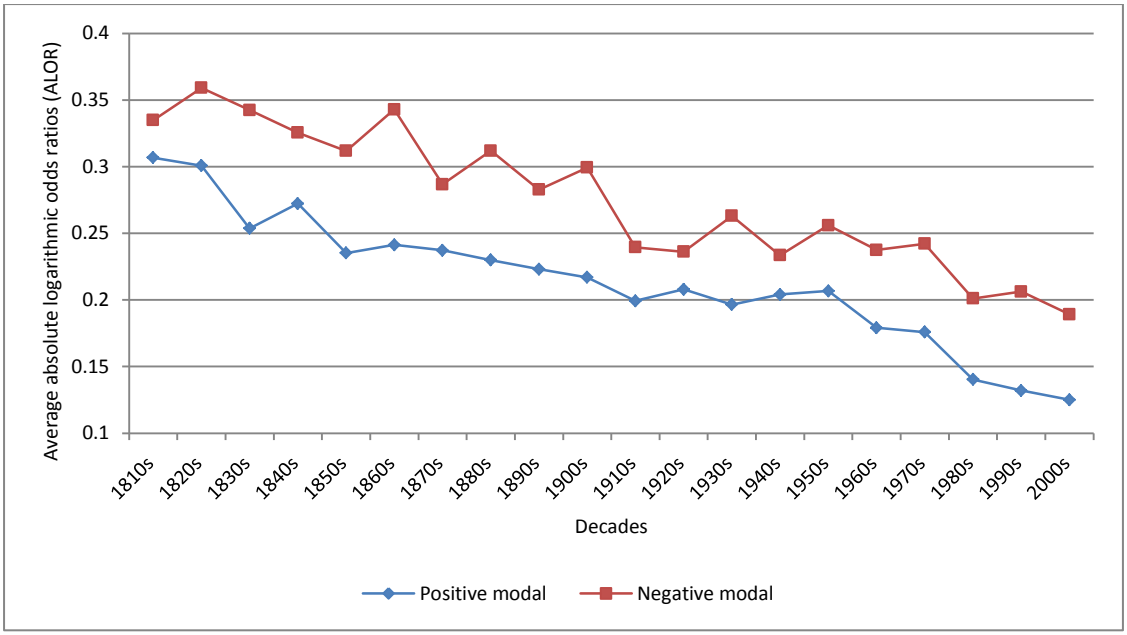

Figure 25. Average absolute log odds ratios of positive and negative modal constructions over time.

On average, there is an ALOR difference between the two of approximately 0.061, corresponding to at least a 15% difference; even in the 2000s, the ALOR value of the negative modal construction is 0.189 – the odds of the construction being biased are almost 55% greater than if it were not. When considering that negated modal constructions in combination with the verbs selected typically indicate prohibition or proscription in the deontic sense (*must not*, *should not*, *cannot*), and inability (*cannot*, *could not*), uncertainty (*might not*) or unwillingness (*would not*) in the epistemic sense, it would not be remiss of one to suspect that such conditions display stronger gender bias than forms of positive polarity owing to societal norms and expectations. A clearer picture would probably emerge if more verbs and modal auxiliaries were examined and the constructions segregated according to their uses in the deontic and epistemic senses wherever possible; perhaps in the future, it would also help to investigate the modals individually in finer detail since they all have their special functions.[17]

Also noticeable in Figure 25 is how gender bias has decreased overall for both constructions, a trend that is seemingly common to all constructions examined in this study. One notices that the 1970s drop occurs a decade earlier in the case of the modals, with the decline accelerating between the 1970s and 1980s; the level of bias does not reach pre-1970s levels thereafter. This may suggest that modal constructions retain their biases slightly longer, something which can also apply to e.g. negative progressives and other such constructions. In

[17] Several charts involving the modal verbs are included in the Appendix.

any case, there is a clear distinction in most cases – barring some perfect tense constructions – between the discernible levels of gender bias associated with constructions in the period lasting from the 1810s to the 1950s and the latter half of the 20th century to the beginning of the 2000s. In fact, now that all the relevant charts have been presented, it is possible to make more concrete generalisations with regard to the dimensions of interest starting with tense and negation.

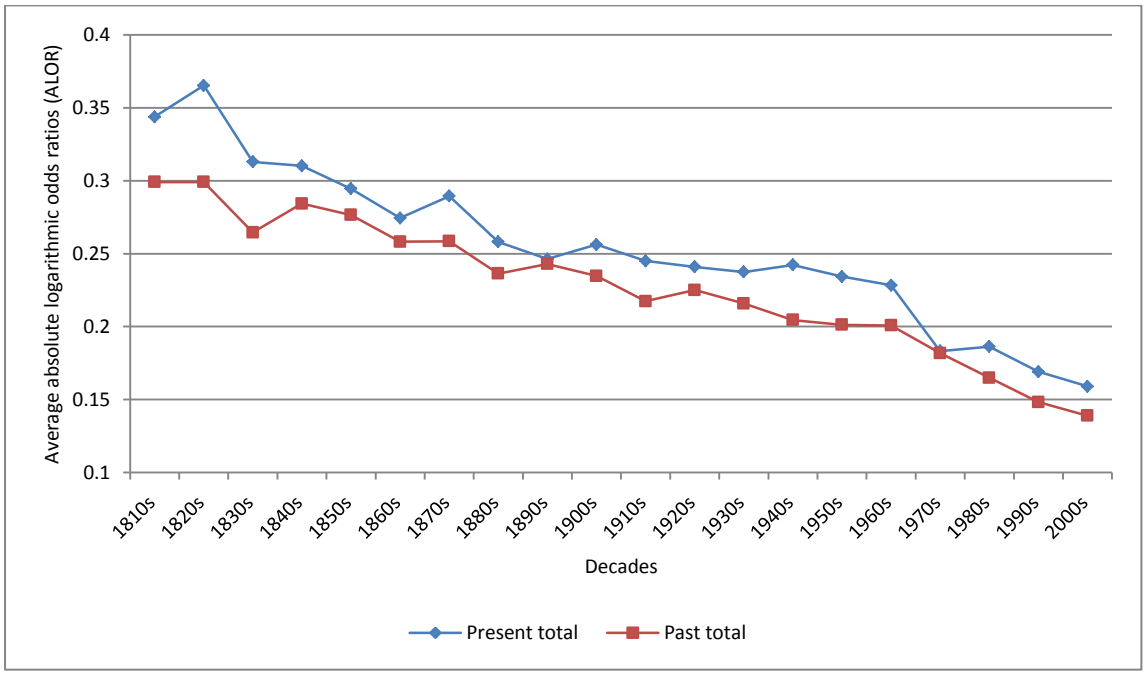

Figure 26. Average absolute log odds ratios of all present and past tense constructions over time.

Aggregating the data collected, one can compare the overall bias trajectories of the tensed constructions – as well as all other constructions according to the aspects of interest. For the tenses, only the present and past tense constructions are considered; modal constructions are therefore included from this collective dataset. Combining and averaging all the ALOR values for each decade resulting from relevant tensed constructions, a chart with trajectories as shown in Figure 26 is produced. Comprising constructions in the present tense are the present simple, progressive and perfect; those in the past tense are of the same types.

The total tense averages confirm both the gradual reduction of bias and, more importantly, the nature of the present tense to exhibit greater bias than the past tense. What the bias trajectories indicate is that the odds for present tense constructions being more strongly inclined to take *he* or *she* as subjects for certain verbs are, on average, about 6% higher than those for the past tense. The assumption is that the present tense in published written American English is marked, judging by the tendency noted. This is unsurprising for the past

tense is much more frequently used and found in published written material even in absolute terms; while the frequencies or the markedness of the tense by themselves admittedly do not explain the strength of bias, they are factors which probably should not be ignored. The reading is that it is more likely – albeit not extremely so – for male and female figures in texts to be depicted in more stereotypical manners with the present tense than the past tense, i.e. as doing things regularly or habitually, actively doing things contemporarily within the speaker's perspective, or having done something of relevance to the speaker's present. The reason behind this phenomenon is probably contingent on multiple factors, but speculating, perhaps the sense of immediacy expressed by the present tense introduces greater subjectivity on the part of the speaker – be it character, narrator or writer – and conditions them to resort to biased portrayals corresponding to societal gender norms and expectations as opposed to the past tense which creates temporal distancing that allows time for reflection. With that in mind, the question going unanswered is whether the authors themselves were and are aware of this in their own writing.

Furthermore, although present tense constructions used for this research have continuously been more biased than the ones in past tense, there are two distinct decades in which the difference in the degree of bias as represented by the ratios is virtually negligible, merely measurable in the thousandths: 1890s and 1970s. The 1890s may seem an anomaly, but the decade did see the two competing women's suffrage organisations in the United States at the time merge into one as well as the recognition of women's right to vote in some states (Frost-Knappman & Cullen-DuPont 2005: 243, 271); whether these events had an effect in somehow bridging the gap between the present and past tense constructions is purely conjectural. What is certain, however, is that the aforementioned 1970s drop visible in both trajectories occurs stronger with the present tense – the decrease in bias for the present tense from the 1960s to the 1970s is more than twice that registered for past tense constructions. One suspects that there was a realisation of such biases, hence the noticeable reduction. Perhaps if the effort towards gender-neutral use of present tense constructions had not stalled in the 1980s – seeing as how there is minimal change in bias displayed between the 1970s and 1980s based on the trajectory – the present and past tense formulations would exhibit roughly equal rates of bias into the 21st century, considering the similar post-1980s trends of the constructions.

Trends pertinent to polarity too can be compiled and presented in aggregated format. In the case of positive and negative constructions, all available ALOR scores of the 13 constructions each comprising the two categories, i.e. 13 positives and 13 negatives, were averaged per decade to produce the trajectories in Figure 27.

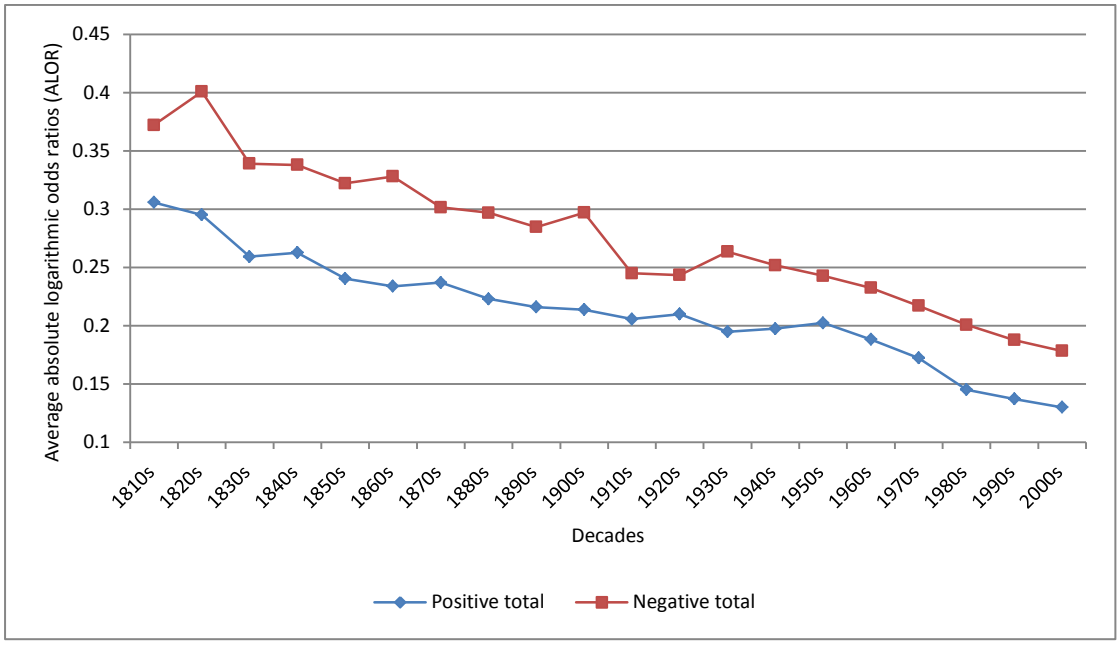

Figure 27. Average absolute log odds ratios of all positive and negative constructions over time.

As already observed, Figure 27 provides further confirmation of the trend of diminishing gender bias associated with constructions, this time with negation. Also immediately conspicuous is that negative constructions are indubitably and unambiguously more biased than positive ones; on average, there is an ALOR difference of around 0.064 between the positives and negatives or a minimum 15% discrepancy in the degree of bias associated with the constructions. Much like present tense constructions, negative constructions – or at the very least, those featuring the negator *not* – are considerably fewer than positive ones in terms of token count. That and the fact that negation is understood to be a marked concept may, in this context, be relevant aspects in making sense of the higher level of bias. One posits that portrayals of males and females not doing something in texts for whatever reason – owing to unwillingness, inability, incapacity, etc. – have greater odds of being biased when speakers remark on the unusualness of the figure not doing the thing. For instance, it may be that in connection with verbs already exhibiting significant gender bias, men and women not acting in stereotypical ways is more worth noting and writing about. If, for the sake of argument, males and females are respectively more likely to be depicted writing and reading, men not writing and women not reading can be construed as being more 'remarkable' than

their doing the things; this would be especially prominent with verbs like *shout* and *scream* which are clearly biased, effectively reinforcing the bias. Adding to that, these biases could then be amplified due to lower frequencies of negative constructions.

Also interesting upon examining the positive and negative construction trajectories is the trend of decline in the 20th century. Negative constructions decreased in bias between the quite sharply between the 1900s and 1910s before rising again in the 1930s. Thereafter, however, negative constructions have seen a continuous and stable decrease in associative gender bias. This differs from what is observable for the positive constructions, characterised by small falls and rises in the first half of the 20th century before the reduction in bias continues definitively from the 1960s onwards. The chart thus shows that diminishing bias trends in the dimension of polarity actually began quite early – surprisingly so for the negative forms, which register a definitive shift two decades before the positive variants do. Moreover, the oft-noticed 1970s drop is not visible here; there is a somewhat discernible steeper fall for the positive constructions occurring between the 1970s and 1980s, but there is no readily analogous identifiable drop for the negated forms. This post-1930s regularity in the trajectory indicates that bias decreases are present and assessable but also hints at some kind of resistance on the part of negative constructions as well as pointing to the stability of the form. Still, remembering that negative construction ALORs are for the most part incomplete in the sense that there are lacking attestations, caution is warranted when making any statements and interpreting these findings on the matter of bias associated with negation.

We have hitherto looked at the constructions separately on the dimensions of tense (present and past) and polarity (positive and negative). It would make perfect sense now to average the biases of the construction types. That entails examining the constructions in their totality; the bias trajectories of the simple, progressive, perfect and modal constructions – the tensed and negated forms of each type combined – can also be charted to see what effect, if there is any, the type of construction itself has on the degree of bias displayed on the basis of the 36 verbs selected for study. The results complement findings already presented on the overall gender bias exhibited by the constructions; here, the constellations are slightly different as tense and negation are rendered irrelevant and subsumed into the category of type, and specific verbs were used to compile the datasets. The bias trajectories as determined by construction type are given in Figure 28.

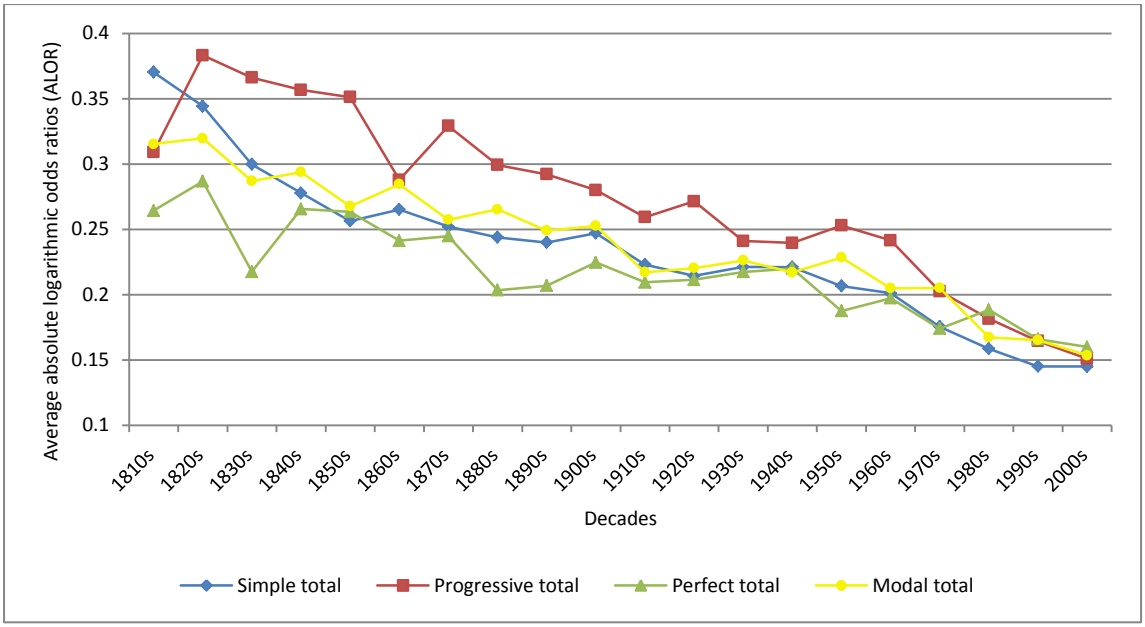

Figure 28. Average absolute log odds ratios of construction types over time.

Predictably, similar trends of bias reduction are visible for each of the constructions, albeit obviously occurring at varying rates; some register quicker and steeper falls while others take longer in their efforts to minimise bias. The trajectories in Figure 28 are effectively summaries of all that has been previously presented; nevertheless, presenting them in one chart allows one to compare the tendencies conveniently. Indeed, it is evident that progressive constructions have exhibited higher levels of bias than other constructions for a long time, specifically between the 1820s and 1960s. At the other end of the range, the perfects constitute the least biased construction for most of the decades studied, although – as noted before – perfect construction trajectories have a tendency to be somewhat erratic. Meanwhile, the simple and modal constructions can be seen competing for the middle ground; in a remarkable turn of events, the simple constructions go from being the most biased construction type in the 1810s to displaying the least bias of the four types under investigation starting from the 1980s and into the 21st century. Progressive forms being more biased compared to other constructions means that, with the 36 verbs analysed, gender-biased or stereotyped depictions of imperfective actions – here, actions in progress – are more likely than they would be, for example, for the perfects and simple forms. One wonders whether this is so due to the focus on the duration of a verb in the progressive; bias could potentially be amplified when an event is presented as being stretched out over a period of time.

One can also observe that there are several points in which the levels of aggregate bias expressed by the constructions converge prominently in 20th century data. Between the 1910s

and 1940s as well as in the 1960s, the average ALORs of the simple, perfect and modal constructions do not deviate much from each other, indicating that although there exists tangible gender asymmetries associated with the constructions, one construction does not display significantly more bias than others. The progressive construction, then, seems to be a kind of outlier; its data points appear to generally hover above others up until the 1960s. Indeed, the trends of the second half of the 20th century make for interesting findings. After the 1950s, the falling tendency towards biased usage becomes entrenched; the downward trajectory does not reverse for the simple and progressive constructions, and the modals also do not return to their pre-1950s level of bias. The perfects are anomalous in that regard; previously among the least biased constructions, usually in conjunction with the simple constructions, perfect forms somehow become more biased relative to others and end up displaying the most bias come the 1980s. From then on, the constructions all register a continued decrease in bias and seemingly strive towards converging in the 2000s – perhaps reflecting greater awareness in gender-neutral language use – where the differences between the degrees of bias expressed by constructions become minimal as to be negligible. The 1970s drop, then, is somewhat displaced; it is noticeable strongest in the case of the progressives but absent for the modals. Thus, with the construction types, it seems that the reduction in bias either began sooner already in the 1960s or later, in earnest, in the 1980s. Whatever the case, there are clearly conspicuous measurable decreases for all the constructions, indicating the presence of a universal trend towards gender equality in written American English use with regard to verb constructions.

These results corroborate findings from other research works such as by Ritt et al. (2024: 16) who likewise note gender asymmetries as having reduced in both written American and British English in the span of 200 years. How much of the decrease in bias in various constructions can actually be attributed to social factors and shifting norms and perceptions cannot be adequately assessed with certainty, but going by general identifiable trends coinciding with several historical developments, e.g. feminist calls for advancing women's rights in the 1970s and beyond, it is plausible that increased awareness towards gender issues must have played at least some role in how males and females are depicted in written works. This view is taken by Ritt et al. (2024: 17-18), who remark that there are "conscious efforts […] to avoid stereotyping" in the post-1960s period by those in publishing; in fact, style guides did begin to draw attention to gender bias by the 1970s, also explicitly addressing writers to avoid biased or stereotyped formulations in the 1980s (Curzan 2014: 121, 123).

However, these instances of prescriptivism do not explain the gradual decrease in bias noted from the 19th century onwards. It could be that females simply becoming more visible and represented in texts accounts for the reduction in bias associated with constructions. Perhaps it is the case that male bias has reduced while female bias has increased or vice-versa, or that they both demonstrate diminishing levels of bias.

5.2.3. Male versus female bias on the level of constructions

The previous point raises an intriguing question about the findings for the individual verbs component of this research: whether and how both male and female biases have reduced or increased at varying rates when taking the 36 verbs and the 26 constructions into consideration. It could be that the overall level of bias decrease is due to both male and female constructions showing a decrease in bias, one increasing and the other decreasing, or one decreasing while the other remains stable. The rate at which the female and male biases increase or decrease may also affect general trends; a more rapid decline in female bias as opposed to a steadier decrease of male bias suggests an unequal effort distribution as regards bias reduction. For these reasons, it makes sense to briefly examine the difference between male and female bias trajectories of the constructions over two centuries.

To compare male and female bias on the level of constructions, a slightly different approach to the one hitherto employed is taken. It is first worth bearing in mind what data are available for study: for each construction, 36 verbs were examined across 20 decades. The odds ratios – more specifically the logarithmic and absolute logarithmic odds ratios – for each construction, itself containing data for each decade for each verb, could then be compiled in tabular format to store the data. There are thus 720 possible cells for each construction (36×20); however, none of the 26 tables have all 720 cells filled with the ratios owing to the previously mentioned issue of empty cells. Now, the LOR summary table is once again used here to average the ratios for each construction, but rather than averaging the data for the whole decade, i.e. all available LORs per decade for which there may be a maximum of 36 counts, the verbs are first distinguished as to whether they exhibit female or male bias. Only then are the verbs selected for inclusion into the calculation of averages for female and male bias per decade for every construction. Determining which verbs actually exhibit gender bias – with regard to how consistently they demonstrate biases – requires a standard measure of categorisation.

In this case, the LOR counts of each verb for a particular construction are relevant in ascertaining gender bias. For each construction, there can be a maximum of 20 counts for each verb; with certain constructions, the LORs cannot be computed for a number of verbs in several or all decades, in which case these verbs are definitely excluded. The LORs, as we know, can be either negative or positive, respectively indicating female or male bias. It is quite often the case that a verb used in a construction does not consistently exhibit one type of bias; the ratios can fluctuate between negative and positive values across the decades in which the ratios can be calculated. With that in mind, the idea is basically that if a verb in any given construction has more negative value counts than positive counts, it is deemed to display female bias; more positive value counts then indicates male bias. Should a verb show parity in its bias count – e.g. 10 negatives and 10 positives – it is not considered biased in the context of what is to be accomplished here and is therefore excluded from the bias averages of the construction. The issue with this procedure concerns instances of minimal discrepancy, such as counts of 11-to-9 or low counts of 9-to-5. It was thus decided to include in the averages only those verbs which meet a certain threshold to be considered consistently demonstrating strong male or female biases, strength in this case taken to mean whether the verb in a given construction consistently displays gender bias.

For that purpose, the proportions of the counts are the determinant factor. Two proportions are obtainable for each verb in every construction: the quotient of the number of negative value counts and the sum of negative and positive counts on one hand and the quotient of the number of positive value counts and the sum of negative and positive counts on the other. The interest lies in comparing the proportions to determine whether a verb is consistently biased. To do this, Cohen's h – a measure of difference between proportions (Cohen 1988: 181) – is used whereby the two proportions firstly undergo arcsine transformation and the difference is then obtained between the transformed values; the absolute value of the resulting difference is used in this case.[18] Cohen (1988: 184-185), while cautioning against using interpretation conventions, describes h = 0.2 as showing small effect, 0.5 medium effect, and 0.8 large effect. The threshold for verb inclusion into the averages is thus set, rather arbitrarily, to $h \geq 0.8$ for the sake of simplicity and uniformity. Thus, if a verb for a particular construction has 14 negative counts and 6 positive counts or vice-versa ($h \approx 0.823$),

[18] Cohen (1988: 181) provides the formula $\varphi = 2 \arcsin \sqrt{p}$, whereby p is the proportion, for transforming; the nondirectional effect size measure h is obtained thus: $h = |\varphi_1 - \varphi_2|$. Absolute values are used mainly to facilitate the procedure in Excel, for which the following formula is used: ABS(2*(ASIN(SQRT(p_1))-ASIN(SQRT(p_2)))).

it is included; if, however, a verb has 13 negative counts and 6 positive counts or vice-versa ($h \approx 0.755$), it is excluded.

With the effect size measure *h* as basis, the LORs of verbs with more negative counts are factored into female bias averages and those with more positive counts into male bias averages for each construction. By doing so, the risk of biases cancelling each other out is diminished with negative and positive values generally – but not always – kept separately. The male and female bias averages garnered for each construction are then themselves aggregated and averaged for different dimensions of interest. One could, of course, look at the bias trajectories individually construction by construction. However, the fact that certain constructions have such a paucity of odds ratio data complicates matters in that regard; it is the case, for example, that female or male decade averages – often in the 1800s – are sometimes blank, and that occasionally – due to the limited selection – very few verbs exhibit consistent male or female bias in some construction types (especially egregious with *should not*, which only has one verb meeting the criteria for consistent male bias). Furthermore, the verbs included in the female and male bias averages differ from one construction to another. Due to these factors, it was presumed more meaningful to compile and average the averages rather than viewing the constructions separately for the matter of bias consistency.

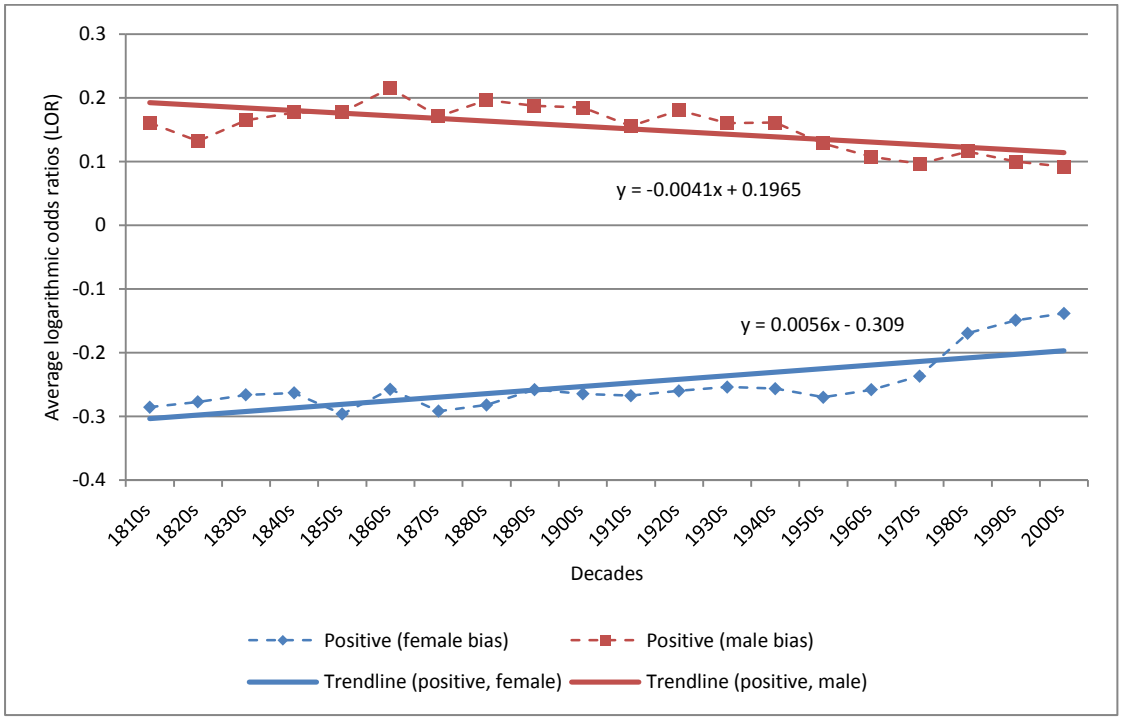

Figure 29. Average log odds ratio trajectories and accompanying trendlines for positive constructions.

To provide a sample of how the rates of male or female bias have progressed between the 1810s and 2000s, Figure 29 presents the trajectories and trendlines of male and female bias associated with positive constructions; the linear trendlines are somewhat more salient than the trajectories in this part.

In the case of positive constructions in Figure 29, one can see that both male and female biases have reduced over time but at slightly different rates. Since this rate is of interest, the trendlines are especially well-suited for such a purpose. Taking the equation of a straight line as $y = mx + c$, where m is the slope or gradient and c the y-intercept, the value of m – the slope – provides a general means for comparing the rates at which the bias decreases or increases. Thus, in Figure 29, the female-biased verbs in positive constructions reduced in bias more quickly than male-biased ones did in the same timespan. The interpretation is that verbs more inclined to take *he* over *she* in positive constructions retain the tendency longer than those which exhibit female bias.

Repeating this process to check if the same trends are perceptible with negative constructions, a chart as visualised in Figure 30 can be produced, where a different picture emerges.

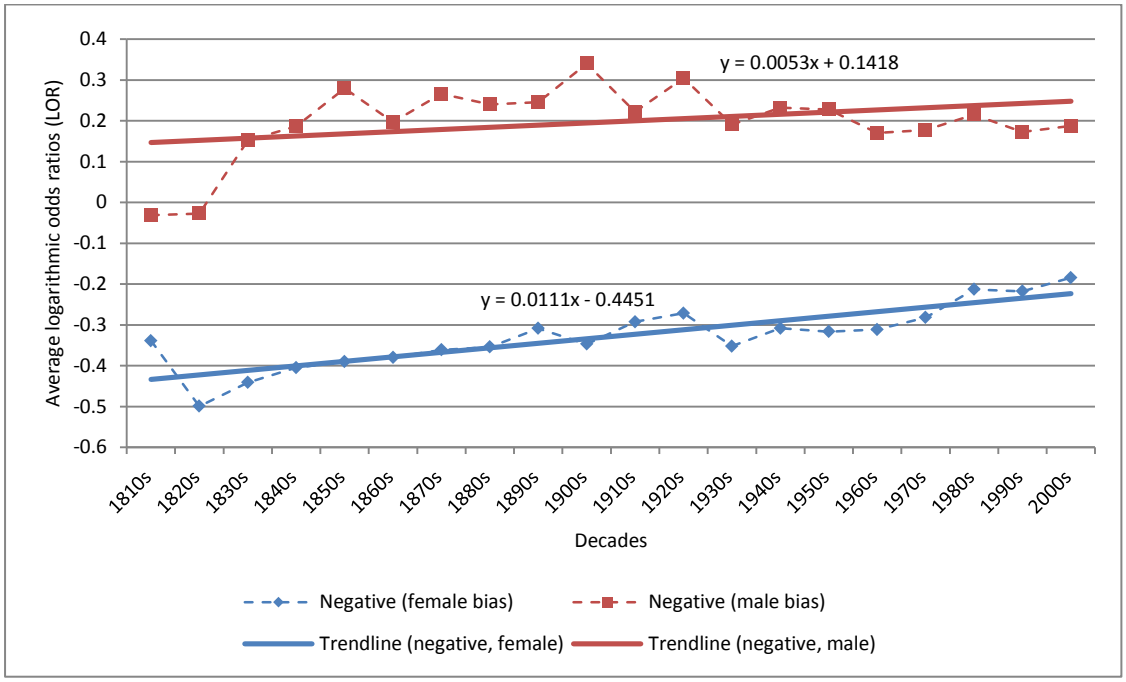

Figure 30. Average log odds ratio trajectories and accompanying trendlines for negative constructions.

With negative constructions, it is still evident that female bias has been decreasing; the same cannot be said for verbs displaying consistent male bias, which appear to have become more biased in the period of two centuries. If one were to take a holistic view and see that, for

example, the level of bias associated with negative constructions has diminished over time on the basis of the verbs chosen for the study which consistently display gender bias, the reasoning behind that would not be because both male and female biases see reductions, but is rather thanks to the reduction in the strength of female bias. Indeed, the gradation of female bias reduction is more than double that of male bias increase, which masks the tendency of the latter. In other words, over the course of 200 years, verbs exhibiting male bias in negative constructions have somehow become more biased while female-biased verbs generally have gradually tended towards gender neutrality. Decrease in overall negative construction bias would then be seen as being more or less instigated by a reduction in female bias.

There are, however, important points for consideration: firstly, from the selection of 36 verbs, those displaying male bias are fewer, and secondly, negated constructions are frequently marked by zero cells when it comes to odds ratios. Data are also quite unreliable for earlier decades; the aggregate average LORs of strongly male-biased verbs in constructions of negative polarity – determined based on the number of positive value counts – are below zero for the 1810s and 1820s, effectively indicating female bias in those decades. Additionally, the tendencies depicted are specific to the verbs considered for inclusion in the male and female bias averages for each construction. Still, the discrepancy – at least for negative constructions – does shed light on the metaphorical division of effort and labour with regard to bias reduction over time, with female bias leading the way towards gender neutral language use.

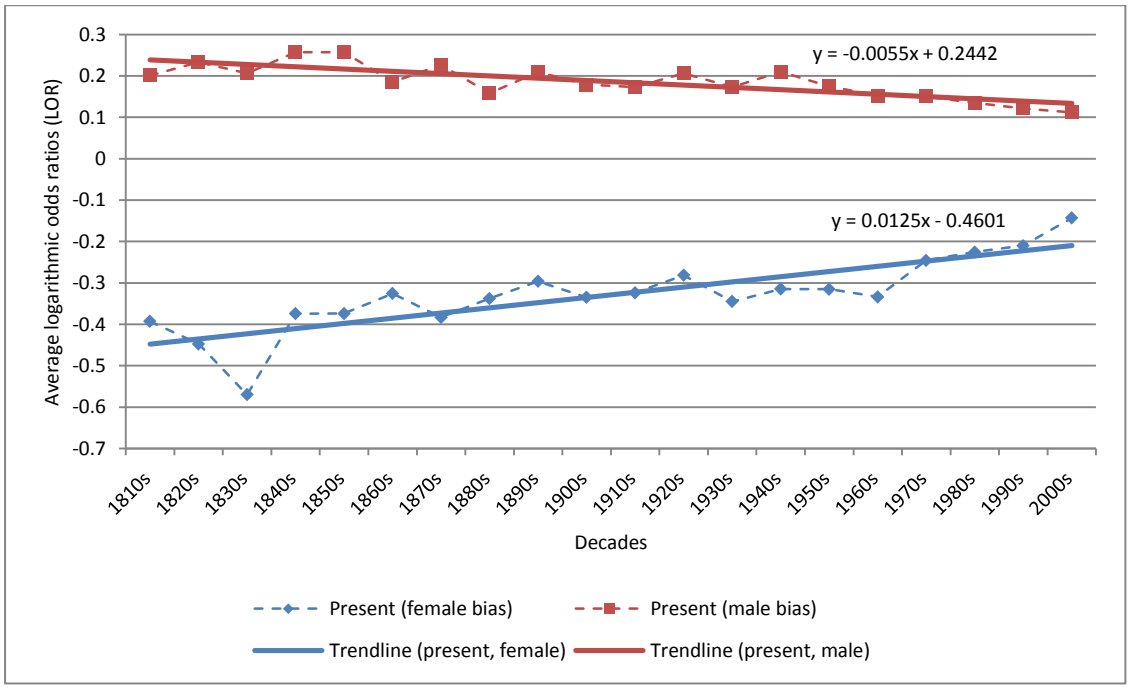

Figure 31. Average log odds ratio trajectories and accompanying trendlines for present tense constructions.

In fact, the tendency for female bias to reduce compared to male bias can be observed for other constructions as well. Looking at the tensed constructions starting with those in the present tense in Figure 31, the rate of female bias decrease is immediately conspicuous; once again, the trends show that the reduction in female bias has been more prominent than for male-biased verbs in present tense constructions over two centuries. At the same time, one sees that there is a discernible decrease in male bias at the very least, evidenced by a negative slope. Notable is that the rate of female bias decrease is more than twice as strong as the decrease in male bias, indicating that the changes in the level of female bias accelerates the changes marked for the construction at large. Since bias reduction is observed for both genders, the combined trend towards unbiased use with the present tense would possibly amplify the apparent decrease in overall bias when the construction is viewed as a whole.

It is somewhat interesting that different trends for male bias are seen for the negative and present tense constructions, considering that these two have been noted to be marked in published written language. The intransigence displayed by male-biased verbs in negative constructions as regards bias reduction contrasts with the steady decline in biased use of certain verbs in the present tense favouring the male subject pronoun; the positive and negative slopes are almost of the same gradation in absolute terms. It is thus plausible that markedness is not a distinguishing factor in the matter of bias reduction, although it should be reiterated that a considerably limited selection of verbs was taken into account to produce trendlines visible here, understandably far fewer than the number of verbs used for the investigations in previous subsections.

Examining next the past tense construction bias trends differentiated by gender as shown in Figure 32, familiar tendencies can be observed. The trendlines for past tense bias progression bear some similarity to the ones found for negative constructions in terms of outlook, namely that female bias – yet again – is seen as having weakened whereas male bias registers an increase, indicated by the positive slopes. The decrease in female bias is, at this point, something which is to be expected. The reduction in male bias, on the other hand, is not a guarantee – taking a timeframe of 200 years, male bias may appear to be weakened or strengthened, depending on the construction. With past tense constructions, female bias decreases at a quicker rate than male bias rises, concealing the fact that gender bias has not, in fact, diminished equally in both male and female dimensions over time; this is a trait likewise visible in bias trends for negative constructions.

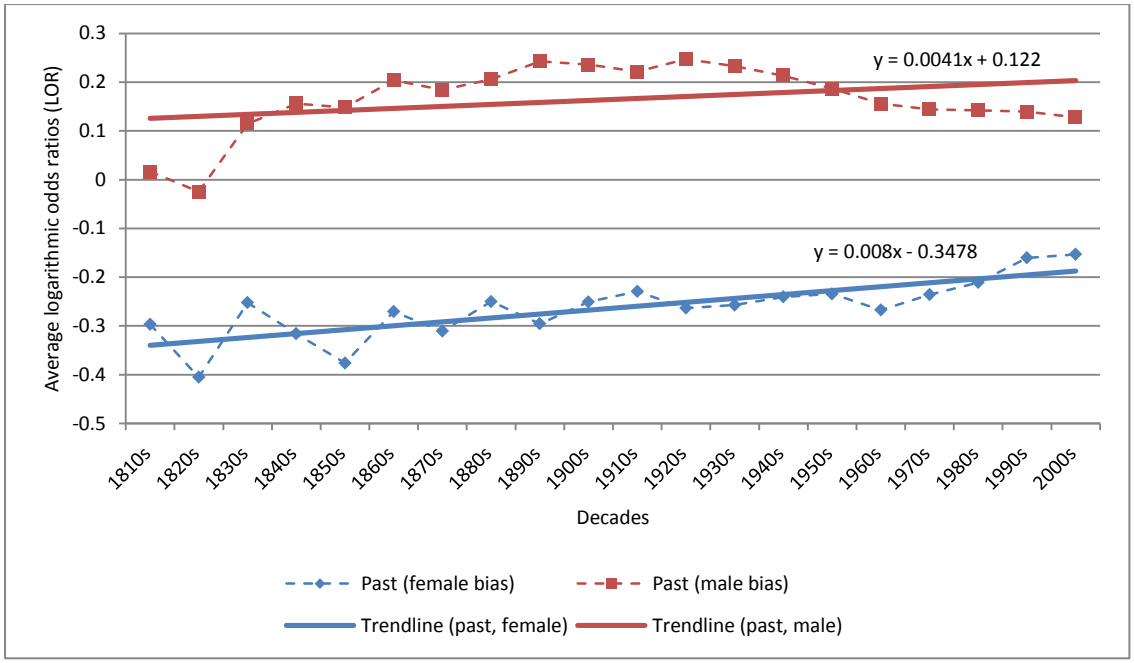

Figure 32. Average log odds ratio trajectories and accompanying trendlines for past tense constructions.

The trendlines for past tense bias progression bear some similarity to the ones found for negative constructions in terms of outlook, namely that female bias – yet again – is seen as having weakened whereas male bias registers an increase, indicated by the positive slopes. The decrease in female bias is, at this point, something which is to be expected. The reduction in male bias, on the other hand, is not a guarantee – taking a timeframe of 200 years, male bias may appear to be weakened or strengthened, depending on the construction. With past tense constructions, female bias decreases at a quicker rate than male bias rises, concealing the fact that gender bias has not, in fact, diminished equally in both male and female dimensions over time; this is a trait likewise visible in bias trends for negative constructions.

It is worth pointing out, however, that the rise in male bias is primarily due to the low average LORs in the earlier decades with uneven data. If one were to start observations from the 1900s, then the trendline would clearly show a decrease in male bias, at least for the past tense constructions. This is a limitation of linear trendlines as bends and turns obviously cannot be accounted for, straight lines being generated for the purpose of identifying best-fit linear trends to the data points. In actuality, to visualise the male bias trend in Figure 32, it would be much more apt to use, for example, a polynomial trendline – even one just of the

second degree – rather than a linear trendline which is a remarkably poor fit to the dataset.[19] Nevertheless, linear trendlines are used in the examples here for their simplicity and also to maintain uniformity.

Most importantly, both female and male consistent bias averages of all 26 constructions can themselves be averaged to procure a grand picture of gender bias progressions between the 1810s and 2000s. The trends based on the average of averages are visualised in Figure 33.

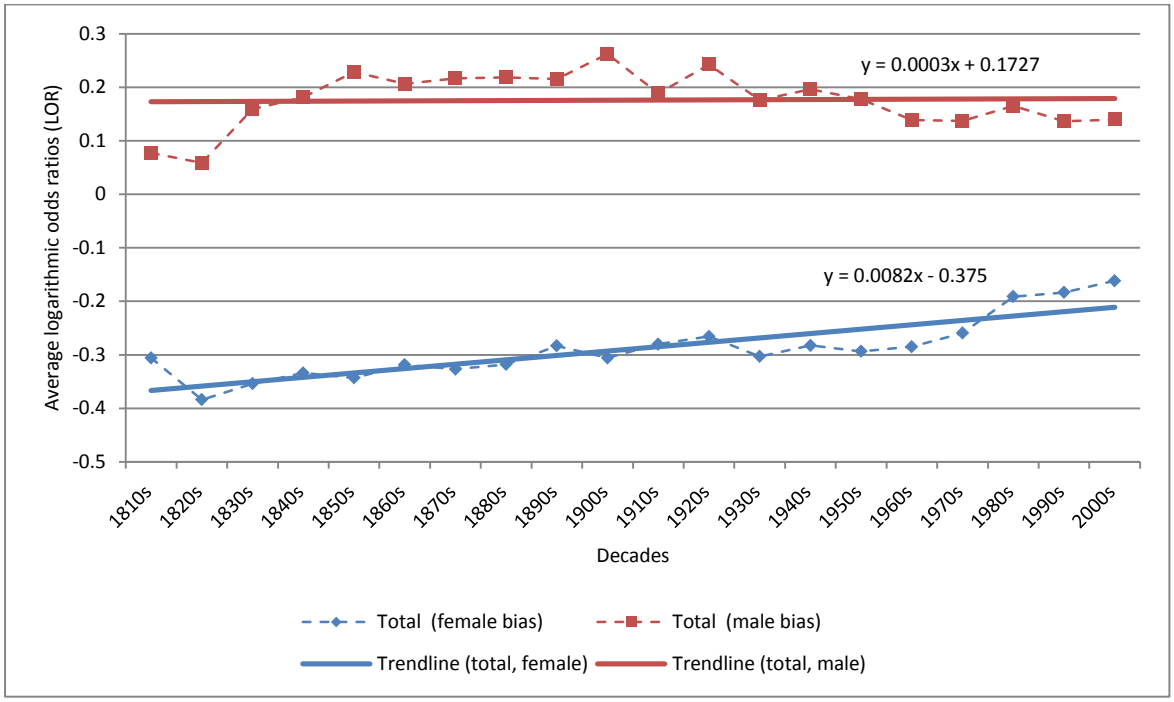

Figure 33. Average log odds ratio trajectories and accompanying trendlines taking all 26 constructions into account.

The two most salient observations to be made here concern the gradual reduction of female bias and relative constancy of male bias over the same time period. Both linear trendlines have a positive gradient, meaning that while female bias steadily weakens or decreases, male bias sees a very slight, almost negligible increase which makes it look as if there have not been significant shifts in male bias progression between the 19th and 21st century. In that case, the trend of falling female bias along with the fairly stable nature of male bias indicate and confirm that the decrease in overall bias can most likely be attributed to changing patterns vis-à-vis female bias rather than to efforts made for male-biased verbs; in a way, female constructions are the driving force behind bias reduction. The interpretation here is that verbs

[19] Values arising from the coefficient of determination R^2 measure, not provided here but which Excel can calculate, are used to determine the goodness-of-fit of the data to the trendlines, whereby a value of 1 indicates a perfect fit (Chicco, Warrens & Jurman 2021: 3).

having greater chances of the tendency to exhibit male bias have managed to retain their biases across constructions in the span of two centuries, generally resisting attempts towards gender neutral usage, whereas female-biased verbs have become less associated with the female subject pronoun in the same timeframe.

That female bias is much more prominently affected hints at a kind of masculinising trend with verbs deemed to be more biased towards females on the account of how consistently they have displayed female bias across 20 decades. One could surmise that the results here point to the continued presence of male dominance and bias in society, with men infiltrating female spaces. Alternatively, it may be that the reduction in female bias actually suggests a feminising trend from the male perspective; essentially, as time goes by, it becomes more acceptable and common for men to do what women tend to do. One assumes that due to male norms and standards, it was perhaps not deemed too unusual for women to act like men by approximating male behaviour while men acting like women might have been received negatively. Thus, it could be that male bias sees little change as maleness is viewed as an immutable natural standard whereas the strength of female bias weakens over time owing to shifts in perceptions of acceptable behaviour where men are concerned. Indeed, considering that societal norms and expectations have long espoused gender conformity – dictating how one should act and behave – it does not seem too surprising to see varying changes in how males and females are depicted in texts. All this is admittedly mere speculation, but these plausible concepts and ideas do provide food for thought.

5.3. General discussion of findings

This research has produced some interesting and intriguing results as regards gender bias linked to various constructions and constellations thereof for which certain observations can be made across multiple dimensions. The findings, on the whole, satisfy the purpose of the study which was to investigate and inspect the constructions and then to demonstrate that they do exhibit different degrees of gender asymmetries and inequalities.

There are several recurring themes and patterns identifiable from examining gender bias trends diachronically. Most crucial is the fact that bias associated with verb constructions has decreased overall in published written American English material between the 1810s and 2000s based on GBA data. The findings also show that bias is prevalent and pervasive; even if the biases are small and seemingly negligible, they are nevertheless statistically measurable.

Indeed, gender bias has reduced but has not been eliminated – it persists at low rates even into the 21st century.

Regardless, that bias reduction appears to be the norm – with visible efforts towards gender neutral language use with certain verbs and even verb constructions – is something to welcome. While there are a few exceptions to this tendency – in that some constructions see an increase in bias into the 2000s – the general bias progression is still one of decline. This is especially conspicuous with global bias trends which show a marked and continuous decrease in bias since not just the beginning of the 20th century but even the 1950s. In fact, the drop in gender bias in the latter half of the 1900s is documented across multiple constructions. In line with that, it is also interesting to see that the post-1950s decline appears to be precipitated by an increase in bias in the first half of the 20th century, however punctuated – in the case of simple constructions, for instance – by a fall in the 1930s.

With that information, one wonders whether wars have an effect on shifting social trends, including issues concerning gender bias. Perhaps it is the case that the want of and return to normalcy in the decade following a major war account for the stability of or increase in bias, against which a countermovement arises to challenge these norms in the decade succeeding the immediate post-war decade, leading to a decline in bias in the two decades after the war. The reaction against established expectations could thus explain the drops of the 1930s and 1960s in the aftermath of the First and Second World Wars, respectively. In the latter case, one could further speculate that the counterculture of the time coupled with efforts against gender stereotyping and bias would have contributed to overall bias reduction which has evidently continued into the 2000s.

In addition to commonalities, the differences in the levels of bias expressed by constructions are notable, for which multiple dimensions can be factored into consideration. The results of this research show that constructions display bias at varying rates, meaning that some constructions appear as being more biased than others. For example, taking a global view across two centuries, one can see that positive simple constructions consisting of present, past and *will*-future tenses are collectively among the most gender-biased constructions in terms of type; the progressive is also considered to be quite biased compared to other forms in connection with select verbs. The perfect construction, on the other hand, appears to be considerably unbiased in comparison to the rest, suggesting that the construction is – or has generally been – more likely to be used in gender-neutral ways. Modals also seem to register

bias less prominently than the simple and progressive constructions when the entire corpus is taken into account.

The findings additionally reveal gender differences on the part of constructions; depending on the type, a construction may exhibit greater female or male bias. It seems that progressive forms are more inclined to take *she* over *he* as subject, present and future tense constructions are more associated with males than females, constructions in the past tense are more likely to display female bias, and that modals – in their totality – have, up until the 1970s, shown greater affinity towards the male subject pronoun over the female one. While perfect constructions have tended to teeter between the boundary of male and female bias, they nonetheless showed – for a century between the 1860s and 1960s – a slight but definitive preference towards the female pronoun. Constructions with the negator *not* see an interesting shift from female to male bias around the 1930s. Of these findings, the gender relationship between the tensed constructions is particularly noteworthy: the present tense (as does the *will*-future) favours *he* while the past tense prefers *she*. Seeing as how these collocational patterns which only become readily perceivable when utilising large datasets do exist – such that the odds of present tense constructions taking the male subject pronoun are greater than it would *she* and the reverse being true for past tense forms – the link between tense and gender may be something worth bearing in mind and exploring further in computational linguistics research, for instance, with regard to bias mitigation in natural language processing tasks and the like.

Another discovery involving tense concerns the generally higher level of bias associated with the present tense over the past tense, a finding obtained from investigating select verbs and various constructions employing those verbs. It cannot be said with certainty whether this is true for global bias as well, considering that negative constructions were treated as a separate category in measuring overall bias. Indeed, the same is the case for negated constructions, comprising the other finding specific to the choice of 36 verbs utilised in this study; constructions of negative polarity appear as consistently expressing greater bias than their positive counterparts. These findings are true insofar the verbs and constructions built around those verbs are concerned. The immediate thought is that the present tense and negation are marked concepts in written language – in comparison to the past tense and positive polarity, which are plausibly the norm – and so are seen as being more biased on account of their markedness. To reiterate, however, these findings should be viewed with extreme caution not

solely due to the limited selection of verbs but also since the data on negation, as has already been emphasised, are not as reliable as one would hope owing to the infrequency of attestations.

The research also found that, once again on the basis of the selected verbs, the reduction in overall gender bias from the 19th to the 21st century is arguably possibly more attributable to the weakening of female bias rather than the joint decrease in both male and female bias, i.e. trends towards gender neutrality on both sides relating to gender-biased verbs and the constructions utilising them. While consistent and strong male bias may have indeed decreased for e.g. positive and present tense forms, the opposite trend is identifiable for constructions with negative polarity as well as those in the past tense: male bias appears to have actually strengthened in those cases. Conversely, female bias is demonstrated as having gradually lost its strength over the course of two centuries in the all dimensions of interest examined. Taking an aggregate view of all the constructions and the resultant averages, one indeed sees that male bias is stable while female bias shows a trend indicating increasing gender-neutral language use with certain verbs and accompanying verb constructions which are characterised by persistent female bias. It could be that, in various texts, male-biased acts are constantly depicted as being done by males while portrayal of acts associated with frequent female bias – thus viewed as being female pursuits – have become less differentiated by gender; this phenomenon suggests male safety as opposed to female volatility as well as an approximation of male behaviour on the part of females or, on the other hand, the increasing acceptability of males to engage in endeavours deemed feminine.

Regarding individual selected verbs, they too were explored with respect to gender to ascertain whether results from previous research can be replicated. The findings largely do corroborate what has been discovered before in several studies (e.g. Pearce 2008; Jockers & Kirilloff 2016; Norberg 2016): certain verbs tied to introspective, emotive and passive acts and of cognition are more likely to collocate with *she* while verbs denoting physical activities and manual labour occur with a higher chance of taking *he* as subject. The results thus confirm that verbs are prone to exhibiting gender bias themselves, which is wholly unsurprising considering societal norms. It is nevertheless interesting to see that some synonymous verbs – such as *scream* and *shout* – are clearly distinguished by gender, which has also been pointed out in previous research, and that other sets of semantically related words – e.g. *talk*, *say* and *speak* – are of ambiguous persuasion. The identification of gender

bias as expressed by near-synonymous terms thus opens an avenue for potential feature research which may be worth exploring. Moreover, as with the constructions, many of the verbs studied – most of them generally female-biased – also demonstrate a gradual trend towards gender-neutral use judging by the weakening of bias over time.

All in all, this research was able to offer insight into gender bias tied to verbs and constructions, the latter aspect having been frequently overlooked by most research on gender bias in language. The results prove that bias is measurable and existent in various constructions and that the biases and the degrees of bias are comparable on the basis of diachronically organised language data. From the findings, one can crucially observe that there is an unmistakable trend towards gender-neutral language use in written American English, evidenced by decreasing levels of bias. The influence of verbs and constructions in the exhibition of bias can also be seen; indeed, it is not only verbs which display male or female bias – constructions do too.

That the odds are higher for males to be portrayed as doing things in the present and future tense and for females being depicted as having done things in the past or doing things which are in progress; that the present tense and negated constructions appear to be more biased than past tense and positive ones in certain contexts; that simple and progressive constructions display greater bias than perfects do for the most part; that female bias seems to have generally reduced at a faster rate relative to male bias – all these are findings of social relevance, putting things into perspective. We arguably do not exactly presume that constructions themselves are biased; certain verbs and adjectives being more likely to collocate with male or female nouns and pronouns is more obvious than constructions demonstrating those tendencies. It is admittedly not immediately apparent why tense and negation – along with aspect and modality – have an effect in the reproduction of gender asymmetries, however small and barely perceptible it may be. One can speculate, as was done for this study, but a more thorough investigation which could possibly address the reasons for such displays of bias is absolutely warranted for future research.

5.4. Issues and limitations

Perhaps the most significant issue with the research concerns generalisability, brought on by the fairly narrow scope of definitions and datasets. It has been routinely emphasised that the findings are valid for specific contexts and rightfully so; the scope and feasibility

understandably had to be considered for this study, which led to limitations regarding the datasets.

To start, only the gendered subject pronouns *he* and *she* were taken into account to measure gender bias, excluding *him* and *her* as well as common nouns like *man* and *woman*. Thus, the results of measurable bias are only true for *he* and *she* constructions and should not be taken to represent or apply to the entirety of the gender binary. The rigidness of the constructions queried – all having the same form of *he* or *she* immediately followed by a verb – meant that cases of adverbs or any other elements separating the pronoun and verb were completely ignored. In a similar vein, the constructions for negation only included instances where the full negator *not* is present; contracted – or rather, inflected – negation forms were not considered and neither were constructions making use of other negative words such as *never*, *no* or *nothing*. The latter two do pose somewhat of a problem as constructions in the form of *he*/*she* [VERB] *no* [NOUN] or *he*/*she* [VERB] *nothing* are not infrequent and would have been included as belonging not to the set of negative constructions but of positive ones. This issue is the result of all right-hand elements following the constructions simply not being considered; it was never the main point to look at the actual collocates of the constructions, although doing so would have allowed for better and more comprehensive analyses of at least the verbs displaying gender asymmetries if not the constructions themselves – after all, context is undeniably important, and this is something that could be addressed in the future.

Moreover, the set of verbs for closer inspection could have been both larger and more structured. In addition, the 36 verbs were, as already mentioned, chosen fairly haphazardly; it would perhaps have made more sense to create a randomised list of the most common lexical verbs in GBA to ensure attestation and to eliminate preconceived bias in the selection procedure. Although unintended, there were more verbs with female inclination than there were those likely to be used with *he*; randomising the selection might have reduced this bias. It should also be said that the choice of constructions was also rather arbitrary, especially so in the case of the modal constructions. They were included for the sake of completeness, but the problem of not being able to conveniently disambiguate between deontic and epistemic senses of some of the modals beyond perusing each attested clause was known from the very beginning, as was the fact that, similarly, polysemous words would not be distinguished by meaning either. It is therefore difficult to generalise let alone draw definitive conclusions on the basis of such uncurated data; more verbs and constructions being explored could have

allowed for more robust statements and elucidations. It should be noted, however, that while a larger verb dataset would have been more welcome, the number of queries to be run with 'only' 36 verbs was already quite substantial when considering the two genders and the number of constructions – 26 in total – as well. Also, had the GBA not had a limitation of 100 queries in 24 hours, more verbs could have been taken into account in order to reduce possible noise and to introduce greater variation.

The issues facing GBA have previously been noted. Beyond the question of representativeness surrounding the corpus, the fact that the findings are – in actuality – applicable to only published written American English means that the scope is very limited and generalisations are impossible. Caution is also advised when dealing with written language data for matters pertaining to gender inequalities as written language, especially when intended for publishing, is arguably more polished and carefully considered. Thus, while it can be surmised that social changes may have had an effect on how males and females have been depicted in texts over several decades – such that the reduction of overall bias can be explained – the findings, on their own, cannot say if gender-neutral language use has actually become the norm outside of the textual form. It would thus be most ideal to use spoken language corpora to examine gender bias diachronically, but spoken data would only stretch back a century or so.

Another problem concerning the corpus is technically external to it: the occasional scarcity or outright absence of data. The presence of zero cells meant that the odds ratios could not be calculated in numerous instances, which leads one to ask how trustworthy, reliable and meaningful the findings are when taking into account the lack of data in computing the averages. There is no true answer to that; we do not simply accept the results at face value as we ought to engage critically with them in light of the known limitations, but they are still valid findings obtained by operating within the parameters set by sound methodology and so should not be discounted. It just means that one must be cognisant as to how and whence the results were derived. In the case of no attestations, it is important to acknowledge that a corpus is not a repository of the whole of a given language but only an assumed representative subsample thereof; one must work within the set of constraints and limitations.

So, it is undoubtedly unwise – not to mention impossible – to generalise findings for such a complex topic; making any strong and conclusive claims in accordance with this context is unwarranted. Nevertheless, the study could be improved and made more meaningful through

alterations in the research design: the incorporation of more verbs and constructions; better screening of the constructions to detect and omit potential false positives; inclusion of collocates preceding and following the constructions; and a more comprehensive selection of gendered terms as well as more varied queries for construction types. That way, hopefully more convincing and definitive trends and tendencies can be ascertained and biases better assessed. Future research work on a larger scale along these lines would ideally provide more opportunities for analysis and could perhaps provide answers to questions going unanswered in this thesis.

6. Conclusion

Measuring gender bias associated with verbs and verb constructions as the ratio of the odds that *he* precedes a given construction and the odds for *she* to appear in the same position and then diachronically tracking the bias trends, the exploratory study comprising this thesis endeavoured to demonstrate that different grammatical features – through constructions accounting for tense, aspect, modality and negation – do exhibit gender bias and asymmetries in unobvious yet comparable ways, beyond the level of lexis, in published American English material. However small, it has been shown that gender bias is statistically measurable and also persistent across the decades surveyed, albeit at varying and changing rates. It can be said, therefore, that there are discernible differences in the representation and portrayal of males and females in written texts.

The aggregated historical bias trends and trajectories importantly show that gender bias has gradually reduced between the 1810s and 2000s in published written texts of American English. While it is not possible to make sweeping generalisations, it is suspected that the decrease possibly reflects and correlates to the increasing social awareness regarding the issue of gender bias and inequality. It is unclear whether conscious efforts are behind the trend of bias reduction and greater gender neutrality in language use(d), but a decrease in bias is a positive sign and something to be applauded in any case.

The research also found that certain constructions display higher levels of bias than others, further indicating that constructions are not – or more specifically, have historically not been – used entirely neutrally in relation to gender. This is most likely not an intentional occurrence on the part of writers and language users in general – if anything, the results could reflect subconscious bias judging by the size and degree of bias expressed, but deriving

psychological or cognitive conclusions solely based on the results is an impossible task. Also considering that the data are only taken from a subsample of written language, it could very well be the case that corpus data are themselves skewed.

Despite all these conjectures, however, the fact remains that there is quantifiable bias visible. Particularly striking are findings which demonstrate that past tense constructions seem to prefer the female subject pronoun over the male one while constructions in the present tense, conversely, are more inclined to take the male pronoun. It is also the case that in terms of aspectuality, the progressive constructions collectively show female bias, i.e. the odds of *she* occurring before progressive forms are higher than *he* does, and that the progressive tends to be one of the more biased constructions in general terms with certain verbs. Modal constructions viewed in totality appear to have exhibited male bias consistently up until the 1970s, whereupon the construction becomes much more gender-neutral in usage. In contrast to all these, the perfect forms have long seen remarkably unbiased use, oscillating between noticeably low levels of male and female bias. Negation, at least when represented by negative constructions featuring *not*, was found to display stronger bias than constructions of positive polarity on the basis of select lexical verbs; it was additionally discovered that *he* began to be more associated with such negative constructions from the 1940s onwards over the female subject pronoun, resulting in male bias. Furthermore, according to the findings as regards the levels of bias expressed, both negative and present tense constructions could also be construed as being marked concepts at least in published American English texts, in opposition to positive and past tense forms, respectively.

The study also corroborated the tendency of certain verbs to collocate more strongly with either the male or female subject pronoun. Although exceptions do exist, verbs of emotion, cognition and introspection are more likely to occur with the female pronoun whereas the likelihood is greater for verbs denoting physical activity and manual labour to take *he* as subject. Bias trends for the verbs also show that many of the verbs are characterised by weak bias in either direction, but the ones which particularly favour the female pronoun appear to have steadily lost their bias potency, meaning that verbs associated with femininity and female behaviour have become more less biased with respect to gender; this phenomenon affects the female-leaning verbs far heavily than it does verbs displaying male bias. Indeed, it would appear by extension that the reduction in bias noted for constructions may be the result of female bias weakening rather than both male and female biases registering equal rates of

decrease in the span of two centuries. On the other hand, the level of male bias putatively remains stable when taking the entire twenty decades as a timeframe; it would be different if one were to set the timespan as beginning with the 1900s, in which case a reduction in male bias would clearly be observed. Since there are fewer verbs that display consistent or strong male bias, more verbs could be explored to see whether it continues to be the case that male bias is resistant to fluctuations.

With gender bias in language having become a renewed topic of interest with the advent of and advancements in machine learning, the findings have implications for that relevant domain and related fields. If machines can perceive biases that humans cannot – and considering that they also have the documented potential to amplify biases with limited human-generated data – certain grammatical features may also have to be taken into consideration when dealing with matters pertinent to gender bias and bias mitigation strategies, seeing as how constructions do have the capacity to be biased. It could of course be that the bias expressed by the grammatical categories comprising the constructions is so minimal and that elements following constructions are better predictors of bias and thus of greater salience, but this research nonetheless provides novel and plausibly important information with regard to the potential influence of grammatical elements on gender bias, an aspect which has hitherto been seemingly overlooked.

This research is thus a worthy and necessary contribution to the body of work concerned with linguistic gender bias. It is hoped that this study provides some direction for future research on the topic of bias in general, and especially in relation to linguistic elements which are not usually considered or expected to display overt gender asymmetries. Later works could – based on large-scale studies involving proper qualitative analyses – ideally eventually provide satisfactory and convincing answers to the remaining undetermined questions and also resolve existing problems by addressing the limitations. Further research could then perhaps illuminate the reasons behind the bias tendencies, whether similar inequalities and trends of diminishing overall bias are also observable in other varieties of English, different text genres as well as for spoken language, what possible effects even minimal biases may have, and many other such curiosities and intricacies.

7. References

Babaeianjelodar, Marzieh; Lorenz, Stephen; Gordon, Josh; Matthews, Jeanna; Freitag, Evan. 2020. "Quantifying gender bias in different corpora". In El Fallah Seghrouchni, Amal; Sukthankar, Gita; Liu, Tie-Yan; van Steen, Maarten (eds.). *WWW '20: Companion proceedings of the web conference 2020*. New York, NY: Association for Computing Machinery, 752-759. https: //doi.org/10.1145/3366424.3383559.

Baker, Paul. 2010. "Will Ms ever be as frequent as Mr? A corpus-based comparison of gendered terms across four diachronic corpora of British English". *Gender and Language* 4(1), 125-149.

Baker, Paul. 2014. *Using corpora to analyze gender*. London: Bloomsbury.

Basta, Christine; Costa-jussà, Marta R.; Casas, Noe. 2019. "Evaluating the underlying gender bias in contextualized word embeddings". In Costa-jussà, Marta; Hardmeier, Christian; Radford, Will; Webster, Kellie (eds.). *Proceedings of the first workshop on gender bias in natural language processing*. Association for Computational Linguistics, 33-39. https://doi.org/10.18653/v1/W19-3805.

Bender, Emily; Gebru, Timnit; McMillan-Major, Angelina; Shmitchell, Shmargaret. 2021. "On the dangers of stochastic parrots: can language models be too big?". In *FAccT '21: Proceedings of the 2021 ACM conference on fairness, accountability, and transparency*. New York, NY: Association for Computing Machinery, 610-623. https://doi.org/10.1145/3442188.3445922.

Berglind Söderqvist, Erika. 2017. "Evidentiality across age and gender: a corpus-based study of variation in spoken British English". *Research in Corpus Linguistics* 5, 17-33.

Biber, Douglas; Burges, Jená. 2000. "Historical change in the language use of women and men: gender differences in dramatic dialogue". *Journal of English Linguistics* 28(1), 21-37.

Binnick, Robert I. 2006. "Aspect and aspectuality". In Aarts, Bas; McMahon, April. (eds.). *The handbook of English linguistics*. Oxford: Blackwell, 244-268.

Bodine, Ann. 1975. "Androcentrism in prescriptive grammar: singular 'they', sex-indefinite 'he', and 'he or she'". *Language in Society* 4(2), 129-146.

Bolukbasi, Tolga; Chang, Kai-wei; Zou, James; Saligrama, Venkatesh; Kalai, Adam. 2016. "Man is to computer programmer as woman is to homemaker? Debiasing word embeddings". In Lee, Daniel D.; von Luxburg, Ulrike; Garnett, Roman; Sugiyama, Masashi; Guyon, Isabelle (eds.). *NIPS'16: Proceedings of the 30th international conference on neural information processing systems*. Red Hook, NY: Curran Associates, 4356-4364. https://doi.org/10.5555/3157382.3157584.

Butt, Miriam. 2010. "The light verb jungle: still hacking away". In Amberber, Mengistu; Baker, Brett; Harvey, Mark (eds.). *Complex predicates: cross-linguistic perspectives on event structure*. Cambridge: Cambridge University Press, 48-78.

Cameron, Deborah. 2005. "Language, gender, and sexuality: current issues and new directions". *Applied Linguistics* 26(4), 482-502.

Cameron, Deborah. 2007. *The myth of Mars and Venus: do men and women really speak different languages?*. Oxford: Oxford University Press.

Cameron, Deborah. 2023. *Language, sexism and misogyny*. London: Routledge.

Čermáková, Anna; Fárová, Lenka. 2017. "His eyes narrowed — her eyes downcast: contrastive corpus-stylistic analysis of female and male writing". *Linguistica Pragensia* 27(2), 7-34.

Chicco, Davide; Warrens, Matthijs J.; Jurman, Giuseppe. 2021. "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and

RMSE in regression analysis evaluation". *PeerJ Computer Science* 7: e623, 1-24. https://doi.org/10.7717/peerj-cs.623.
Cho, Won Ik; Kim, Ji Won; Kim, Seok Min; Kim, Nam Soo. 2019. "On measuring gender bias in translation of gender-neutral pronouns". In Costa-jussà, Marta; Hardmeier, Christian; Radford, Will; Webster, Kellie (eds.). *Proceedings of the first workshop on gender bias in natural language processing*. Association for Computational Linguistics, 173-181. https://doi.org/10.18653/v1/W19-3824.
Coates, Jennifer. 2016. *Women, men and language: a sociolinguistic account of gender differences in language*. (3rd edition). London: Routledge.
Cohen, Jacob. 1988. *Statistical power analysis for the behavioral sciences*. (2nd edition). Hillsdale, NJ: Lawrence Erlbaum Associates.
Crawley, Michael J. 2015. *Statistics: an introduction using R*. (2nd edition). Chichester: Wiley.
Curzan, Anne. 2014. *Fixing English: prescriptivism and language history*. Cambridge: Cambridge University Press.
Czerwionka, Lori; Staszkiewicz, Bruno; Shamloo, Farzin. 2023. "Contextual variables as predictors of verb form: an analysis of gender and stance in Peninsular Spanish requests". *Languages* 8: 202, 1-26. https://doi.org/10.3390/ languages8030202.
Davies, Mark. 2011-. *Google Books (American English) Corpus: 155 billion words, 1810-2009*. https://www.english-corpora.org/googlebooks/x.asp (10 July 2024).
DeFranza, David; Mishra, Himanshu; Mishra, Arul. 2020. "How language shapes prejudice against women: an examination across 45 world languages". *Journal of Personality and Social Psychology* 119(1), 7-22.
Doughman, Jad; Khreich, Wael; El Gharib, Maya; Wiss, Maha; Berjawi, Zahraa. 2021. "Gender bias in text: origin, taxonomy, and implications". In Costa-jussà, Marta; Gonen, Hila; Hardmeier, Christian; Webster, Kellie (eds.). *Proceedings of the 3rd workshop on gender bias in natural language processing*. Association for Computational Linguistics, 34-44. https://doi.org/10.18653/v1/2021.gebnlp-1.5.
Dubois, Betty Lou; Crouch, Isabel. 1975. "The question of tag questions in women's speech: they don't really use more of them, do they?". *Language in Society* 4(3), 289-294.
Eberhardt, Maeve. 2017. "Gendered representations through speech: the case of the Harry Potter series". *Language and Literature* 26(3), 227-246.
Eckert, Penelope; McConnell-Ginet, Sally. 2013. *Language and gender*. (2nd edition). Cambridge: Cambridge University Press.
Egbert, Jesse; Larsson, Tove; Biber, Douglas. 2020. *Doing linguistics with a corpus: methodological considerations for the everyday user*. Cambridge: Cambridge University Press.
Einersen, Anna Franciska; Villesèche, Florence; Huopalainen, Astrid. 2021. "A Lacanian perspective on bias in language: how women can(not) ever make it in academia". *Women, Gender & Research* 30(3), 43-55.
Endriga, Divine Angeli; Rosario Jr., Francisco. 2023. "Gender bias in machine translation: the case of Filipino-English translation in Google Translate". In Moratto, Riccardo; Bacolod, Mary Ann G. (eds.). *Translation studies in the Philippines: navigating a multilingual archipelago*. London: Routledge, 83-99.
Ferstl, Evelyn C.; Garnham, Alan; Manouilidou, Christina. 2011. "Implicit causality bias in English: a corpus of 300 verbs". *Behavior Research Methods* 43, 124-135.
Fludernik, Monika. 2012. "Narratology and literary linguistics". In Binnick, Robert I. (ed.). *The Oxford handbook of tense and aspect*. Oxford: Oxford University Press, 75-101.
Frost-Knappman, Elizabeth; Cullen-DuPont, Kathryn. 2005. *Women's suffrage in America*. (Updated edition). New York, NY: Facts on File.

Gardelle, Laure. 2015. “Let her rain, she’s snowing pretty good: the use of feminine pronouns with weather verbs in colloquial English”. *Folia Linguistica* 49(2), 353-379.
Garg, Nikhil; Schiebinger, Londa; Jurafsky, Dan; Zou, James. 2018. “Word embeddings quantify 100 years of gender and ethnic stereotypes”. *Proceedings of the National Academy of Sciences* 115(16), 3635-3644. https://doi.org/10.1073/pnas.1720347115.
Garnham, Alan; Vorthmann, Svenja; Kaplanova, Karolina. 2021. “Implicit consequentiality bias in English: a corpus of 300+ verbs”. *Behavior Research Methods* 53, 1530-1550.
Ginsburg, Harvey J.; Ogletree, Shirley M.; Silakowski, Tammy D. 2003. “Vulgar language: review of sex differences in usage, attributions, and pathologies”. *North American Journal of Psychology* 5(1), 105-115.
Harrigan, Jinni A.; Lucic, Karen S. 1988. “Attitudes about gender bias in language: a reevaluation”. *Sex Roles* 19(3/4), 129-140.
Harvey, John. 2006. “Fiction in the present tense”. *Textual Practice* 20(1), 71-98.
Hauser, David J.; Schwarz, Norbert. 2022. “Implicit bias reflects the company that words keep”. *Frontiers in Psychology* 13: 871221, 1-11. https://doi.org/10.3389/fpsyg.2022.871221.
Hepburn, Alexa; Potter, Jonathan. 2011. “Recipients redesigned: tag questions and gender”. In Speer, Susan A.; Stokoe, Elizabeth (eds.). *Conversation and gender*. Cambridge: Cambridge University Press, 135-152.
Hilpert, Martin. 2020. “The great temptation: what diachronic corpora do and do not reveal about social change”. In Rautionaho, Paula; Nurmi, Arja; Klemola, Juhani (eds.). *Corpora and the changing society: studies in the evolution of English*. Amsterdam: John Benjamins, 3-28.
Holmes, Janet. 1986. “Functions of you know in women’s and men’s speech”. *Language in Society* 15(1), 1-21.
Holmes, Janet. 2007. “Social constructionism, postmodernism and feminist sociolinguistics”. *Gender and Language* 1(1), 51-65.
Huddleston, Rodney. 2002. “The verb”. In Huddleston, Rodney; Pullum, Geoffrey K. (eds.). *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press, 71-212.
Jockers, Matthew; Kirilloff, Gabi. 2016. “Understanding gender and character agency in the 19th century novel”. *Journal of Cultural Analytics* 2(2), 1-26.
Kuznetsova, Julia. 2015. “Are verbs politically correct? A corpus study of gender in Russian verbs”. *Gender and Language* 9(3), 391-434.
Lakoff, Robin. 1973. “Language and woman’s place”. *Language in Society* 2(1), 45-80.
Leaper, Campbell; Robnett, Rachael D. 2011. “Women are more likely than men to use tentative language, aren’t they? A meta-analysis testing for gender differences and moderators”. *Psychology of Women Quarterly* 35(1), 129-142.
Leone, Ljubica. 2022. “A corpus-based investigation of modals in spoken British English: gender variation and change in the years 1994 and 2014”. *English Studies* 103(8), 1318-1333.
Lewis, Molly; Lupyan, Gary. 2020. “Gender stereotypes are reflected in the distributional structure of 25 languages”. *Nature Human Behaviour* 4, 1021-1028.
Limisiewicz, Tomasz; Mareček, David. 2022. “Don’t forget about pronouns: removing gender bias in language models without losing factual gender information”. In Hardmeier, Christian; Basta, Christine; Costa-jussà, Marta R.; Stanovsky, Gabriel; Gonen, Hila (eds.). *Proceedings of the 4th workshop on gender bias in natural language processing (GeBNLP)*, 17-29. https://doi.org/10.18653/v1/2022.gebnlp-1.3.
Lindquist, Hans; Levin, Magnus. 2018. *Corpus linguistics and the description of English*. (2nd edition). Edinburgh: Edinburgh University Press.

Lindqvist, Anna; Renström, Emma Aurora; Gustafsson Sendén, Marie. 2019. "Reducing a male bias in language? Establishing the efficiency of three different gender-fair language strategies". *Sex Roles* 81, 109-117.
Linell, Per. 2005. *The written language bias in linguistics: its nature, origins and transformations*. London: Routledge.
Livia, Anna. 2003. ""One man in two is a woman": linguistic approaches to gender in literary texts". In Holmes, Janet; Meyerhoff, Miriam (eds.). *The handbook of language and gender*. Oxford: Blackwell, 142-158.
Lu, Kaiji; Mardziel, Piotr; Wu, Fangjing; Amancharla, Preetam; Datta, Anupam. 2020. "Gender bias in neural natural language processing". In Nigam, Vivek; Ban Kirigin, Tajana; Talcott, Carolyn; Guttman, Joshua; Kuznetsov, Stepan; Loo, Boon Thau; Okada, Mitsuhiro (eds.). *Logic, language, and security: essays dedicated to Andre Scedrov on the occasion of his 65th birthday*. Cham: Springer, 189-202.
MacArthur, Heather J.; Cundiff, Jessica L.; Mehl, Matthias R. 2020. "Estimating the prevalence of gender-biased language in undergraduates' everyday speech". *Sex Roles* 82, 81-93.
Michaelis, Laura A. 2006. "Tense in English". In Aarts, Bas; McMahon, April. (eds.). *The handbook of English linguistics*. Oxford: Blackwell, 220-243.
Mondorf, Britta. 2004. *Gender differences in English syntax*. Tübingen: Max Niemeyer.
Mulac, Anthony; Lundell, Torborg Louisa. 1994. "Effects of gender-linked language differences in adults' written discourse: multivariate tests of language effects". *Language & Communication* 14(3), 299-309.
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. 2002. "Language and woman's place in earlier English". *Journal of English Linguistics* 30(2), 181-199.
Newman, Matthew L.; Groom, Carla J.; Handelman, Lori D.; Pennebaker, James W. 2008. "Gender differences in language use: an analysis of 14,000 text samples." *Discourse Processes* 45(3), 211-236.
Norberg, Cathrine. 2016. "Naughty boys and sexy girls: the representation of young individuals in a web-based corpus of English". *Journal of English Linguistics* 44(4), 291-317.
Orgeira-Crespo, Pedro; Míguez-Álvarez, Carla; Cuevas-Alonso, Miguel; Rivo-López, Elena. 2021. "An analysis of unconscious gender bias in academic texts by means of a decision algorithm". *PLoS ONE* 16(9): e0257903, 1-20. https://doi.org/10.1371/journal.pone.0257903.
Pauwels, Anne. 2003. "Linguistic sexism and feminist linguistic activism". In Holmes, Janet; Meyerhoff, Miriam (eds.). *The handbook of language and gender*. Oxford: Blackwell, 550-570.
Pearce, Michael. 2008. "Investigating the collocational behaviour of man and woman in the BNC using Sketch Engine". *Corpora* 3(1), 1-29.
Pechenick, Eitan Adam; Danforth, Christopher M.; Dodds, Peter Sheridan. 2015. "Characterizing the Google Books corpus: strong limits to inferences of socio-cultural and linguistic evolution". *PLoS ONE* 10(10): e0137041, 1-24. https://doi.org/10.1371/journal.pone.0137041.
Piits, Liisi. 2016. "Man's gender and age as based on the collocations of the Estonian word *mees* 'man'". *Folklore* 64, 143-158.

Prates, Marcelo O. R.; Avelar, Pedro H.; Lamb, Luís C. 2020. "Assessing gender bias in machine translation: a case study with Google Translate". *Neural Computing and Applications* 32, 6363-6381.
Preisler, Bent. 1986. *Linguistic sex roles in conversation: social variation in the expression of tentativeness in English*. Berlin: Mouton de Gruyter.
Pullum, Geoffrey K.; Huddleston, Rodney. 2012. "Negation". In Huddleston, Rodney; Pullum, Geoffrey K. (eds.). *The Cambridge grammar of the English language*. Cambridge: Cambridge University Press, 785-849.
Radden, Günter; Dirven, René. 2007. *Cognitive English grammar*. Amsterdam: John Benjamins.
Raphael, Esaya Britto. 2023. "Gendered representations in language: a corpus-based comparative study of adjective-noun collocations for marital relationships". *Theory and Practice in Language Studies* 13(5), 1191-1196.
Ritt, Nikolaus; Alderighi, Viola; Amato, Ylenia; Battistelli, Serena; Bertini, Giorgia; Canneva, Chiara; Capelli, Sabrina; Carli, Caterina; Castelli, Martina; Di Chiaro, Benedetta; Giannattaso, Giulia; Guarducci, Ilaria; Lincu, Iulia; Lodovichi, Federica; Manfra, Alessandro; Mantia, Ylenia; Pacini, Letizia; Pauletto, Chiara; Pecchioli, Leonardo; Perez, Matthew; Ramundo, Gloria; She, Yan; Simoncini, Elisa; Suma, Silvia; Adametz, Hannah; Azikhanova, Alua; Bicker, Martina; Fleck, Johanna; Fritz, Anna-Lea; Gründler, Isabella; Gschwendt, Claudia; Huseinovic, Sendi; Kalt, Viktoria; Kieronska, Regina Luiza; Lang, Sarah; Nedeltschew, Nicole; Paleczny, Victoria; Wallner, Anita; Wonnebauer, Tabea; Zang, Shan. 2024. "Collocational reflexes of gender binarism in published books have become less pronounced during the last 200 years". *VIEWS* (*Vienna English Working Papers*) 31, 1-31.
Romaine, Suzanne. 2003. "Variation in language and gender". In Holmes, Janet; Meyerhoff, Miriam (eds.). *The handbook of language and gender*. Oxford: Blackwell, 98-118.
Rubini, Monica; Menegatti, Michela. 2014. "Hindering women's careers in academia: gender linguistic bias in personnel selection". *Journal of Language and Social Psychology* 33(6), 632-650.
Schmauss, Lindsay Susannah; Kilian, Kelly. 2023. "Hedging with modal auxiliary verbs in scientific discourse and women's language". *Open Linguistics* 9: 20220229, 1-27. https://doi.org/10.1515/opli-2022-0229.
Schmid, Hans-Jörg; Küchenhoff, Helmut. 2013. "Collostructional analysis and other ways of measuring lexicogrammatical attraction: theoretical premises, practical problems and cognitive underpinnings". *Cognitive Linguistics* 24(3), 531-577.
Sheng, Emily; Chang, Kai-Wei; Natarajan, Premkumar; Peng, Nanyun. 2019. "The woman worked as a babysitter: on biases in language generation". In Inui, Kentaro; Jiang, Jing; Ng, Vincent; Wan, Xiaojun (eds.). *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP)*. Association for Computational Linguistics, 3407-3412. https://doi.org/10.18653/v1/D19-1339.
Sigley, Robert; Holmes, Janet. 2002. "Looking at girls in corpora of English". *Journal of English Linguistics* 30(2), 138-157.
Stańczak, Karolina; Augenstein, Isabelle. 2021. "A survey on gender bias in natural language processing". *arXiv*: 2112.14168, 1-35. https://doi.org/10.48550/arXiv.2112.14168.
Stefanowitsch, Anatol; Gries, Stefan Th. 2003. "Collostructions: investigating the interaction of words and constructions". *International Journal of Corpus Linguistics* 8(2), 209-243.
Sun, Tony; Gaut, Andrew; Tang, Shirlyn; Huang, Yuxin; ElSherief, Mai; Zhao, Jieyu; Mirza, Diba; Belding, Elizabeth; Chang, Kai-Wei; Wang, William Yang. 2019. "Mitigating gender bias in natural language processing: literature review". In Korhonen, Anna;

Traum, David; Màrquez, Lluís (eds.). *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, 1630-1640. https://doi.org/10.18653/v1/P19-1159.
Szuba, Agnieszka; Redl, Theresa; de Hoop, Helen. 2022. "Are second person masculine generics easier to process for men than for women? Evidence from Polish". *Journal of Psycholinguistic Research* 51, 819-845.
Talbot, Mary. 2003. "Gender stereotypes: reproduction and challenge". In Holmes, Janet; Meyerhoff, Miriam (eds.). *The handbook of language and gender*. Oxford: Blackwell, 468-486.
Taylor, John R. 2002. *Cognitive grammar*. Oxford: Oxford University Press.
Toro Isaza, Paulina; Xu, Guangxuan; Oloko, Akintoye; Hou, Yufang; Peng, Nanyun; Wang, Dakuo. 2023. "Are fairy tales fair? Analyzing gender bias in temporal narrative event chains of children's fairy tales". In Rogers, Anna; Boyd-Graber, Jordan; Okazaki, Naoaki (eds.). *Proceedings of the 61st annual meeting of the Association for Computational Linguistics (volume 1: long papers)*. Association for Computational Linguistics, 6509-6531. https://doi.org/10.18653/v1/2023.acl-long.359.
Touileb, Samia. 2022. "Exploring the effects of negation and grammatical tense on bias probes". In He, Yulan; Ji, Heng; Li, Sujian; Liu, Yang; Chang, Chua-Hui (eds.). *Proceedings of the 2nd conference of the Asia-Pacific chapter of the Association for Computational Linguistics and the 12th international joint conference on natural language processing (volume 2: short papers)*. Association for Computational Linguistics, 423-429. https://aclanthology.org/2022.aacl-short.53.pdf (10 July 2024).
Trudgill, Peter. 1972. "Sex, covert prestige and linguistic change in the urban British English of Norwich". *Language in Society* 1(2), 179-195.
Weber, Frank; Knapp; Guido; Ickstadt, Katja; Kundt, Günther; Glass, Änne. 2020. "Zero-cell corrections in random-effects meta-analyses". *Research Synthesis Methods* 11(6), 913-919.
Ye, Shenglu; Cai, Simin; Chen, Chuansheng; Wan, Qun; Qian, Xiuying. 2018. "How have males and females been described over the past two centuries? An analysis of Big-Five personality-related adjectives in the Google English Books". *Journal of Research in Personality* 76, 6-16.
Zhao, Jieyu; Wang, Tianlu; Yatskar, Mark; Cotterell, Ryan; Ordonez, Vicente; Chang, Kai-Wei. 2019. "Gender bias in contextualized word embeddings". In Burstein, Jill; Doran, Christy; Solorio, Thamar (eds.). *Proceedings of the 2019 conference of the North American chapter of the Association for Computational Linguistics: human language technologies, volume 1 (long and short papers)*. Association for Computational Linguistics. 629-634. https://doi.org/10.18653/v1/N19-1064.
Zhou, Pei; Shi, Weijia; Zhao, Jieyu; Huang, Kuan-Hao; Chen, Muhao; Cotterell, Ryan; Chang, Kai-Wei. 2019. "Examining gender bias in languages with grammatical gender". In Inui, Kentaro; Jiang, Jing; Ng, Vincent; Wan, Xiaojun (eds.). *Proceedings of the 2019 conference on empirical methods in natural language processing and the 9th international joint conference on natural language processing (EMNLP-IJCNLP)*. Association for Computational Linguistics, 5276-5284. https://doi.org/10.18653/v1/D19-1531.
Zwicky, Arnold M.; Pullum, Geoffrey K. 1983. "Cliticization vs. inflection: English n't". *Language* 59(3), 502-513.

8. Appendix

Additional charts and tables not included in the main text are presented here in this chapter. Unless otherwise stated, the logarithmic odds ratio (LOR) values are provided for the summary tables. The odds ratios (OR) have been explained as being limiting towards female bias, and absolute logarithmic odds ratios (ALOR) can very easily be derived from the LORs by simply dropping the negative sign whenever they appear; no treatments are required for positive values.

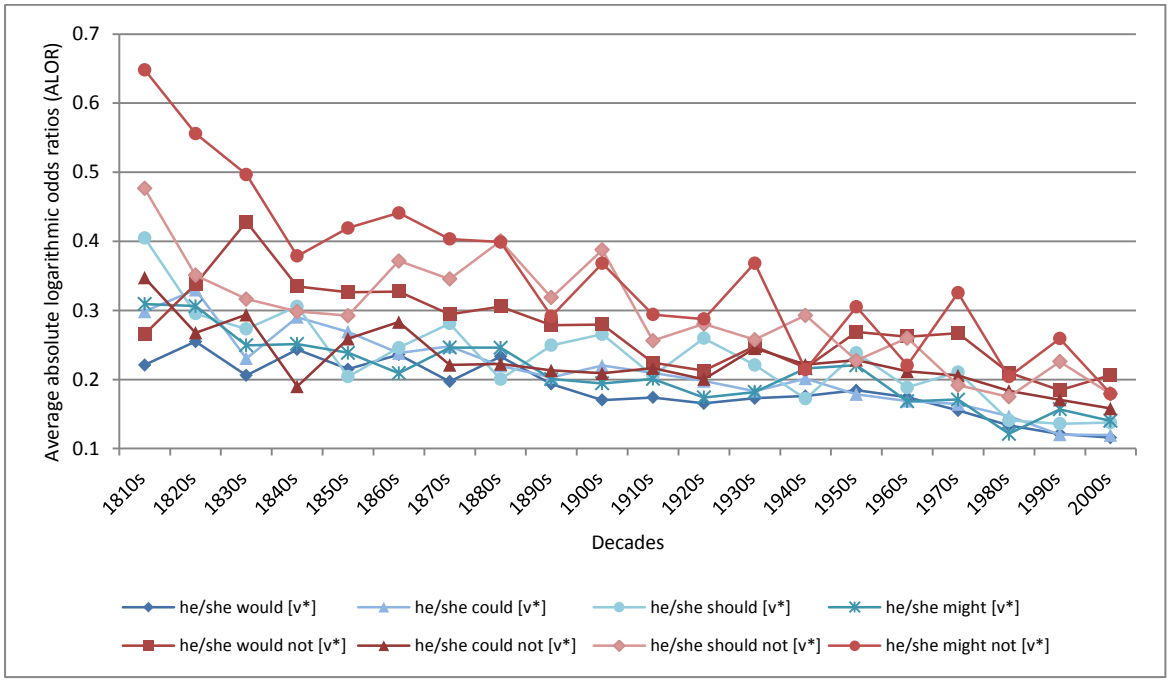

Figure 34. Absolute log odds ratios of positive and negative preterite modal constructions over time.

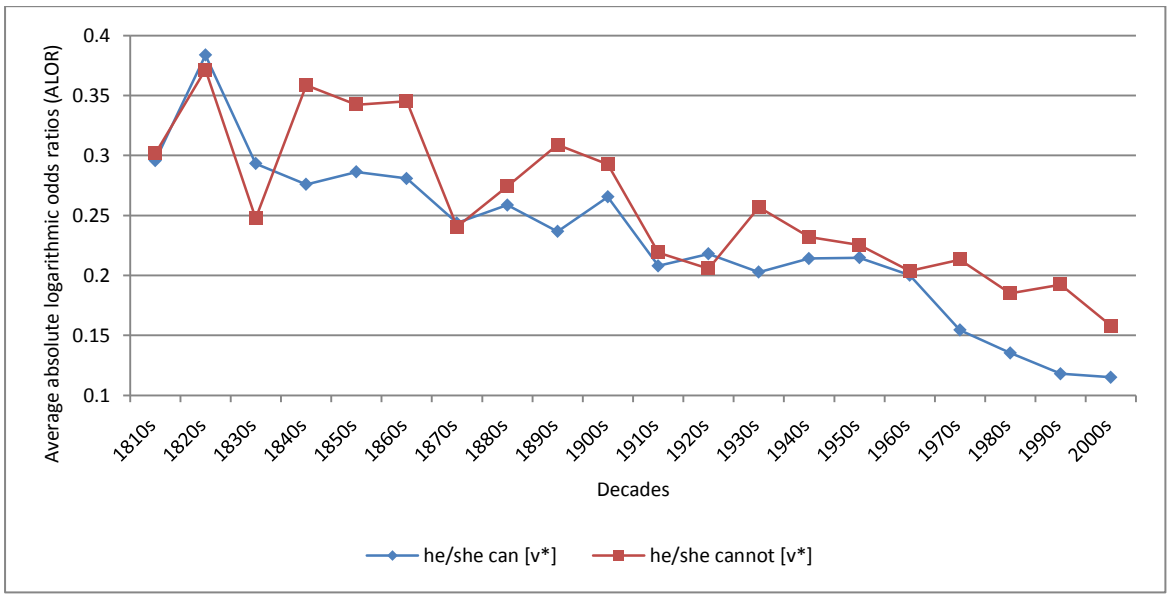

Figure 35. Absolute log odds ratios of *can* and *cannot* constructions over time.

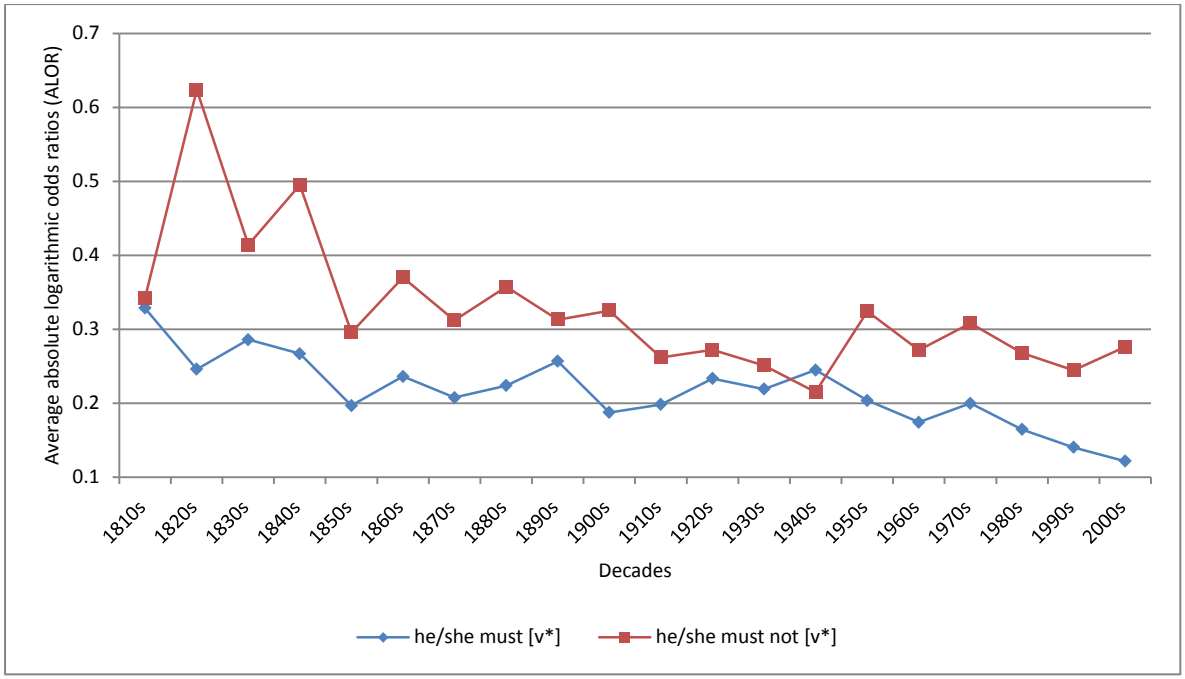

Figure 36. Absolute log odds ratios of *must* and *must not* constructions over time.

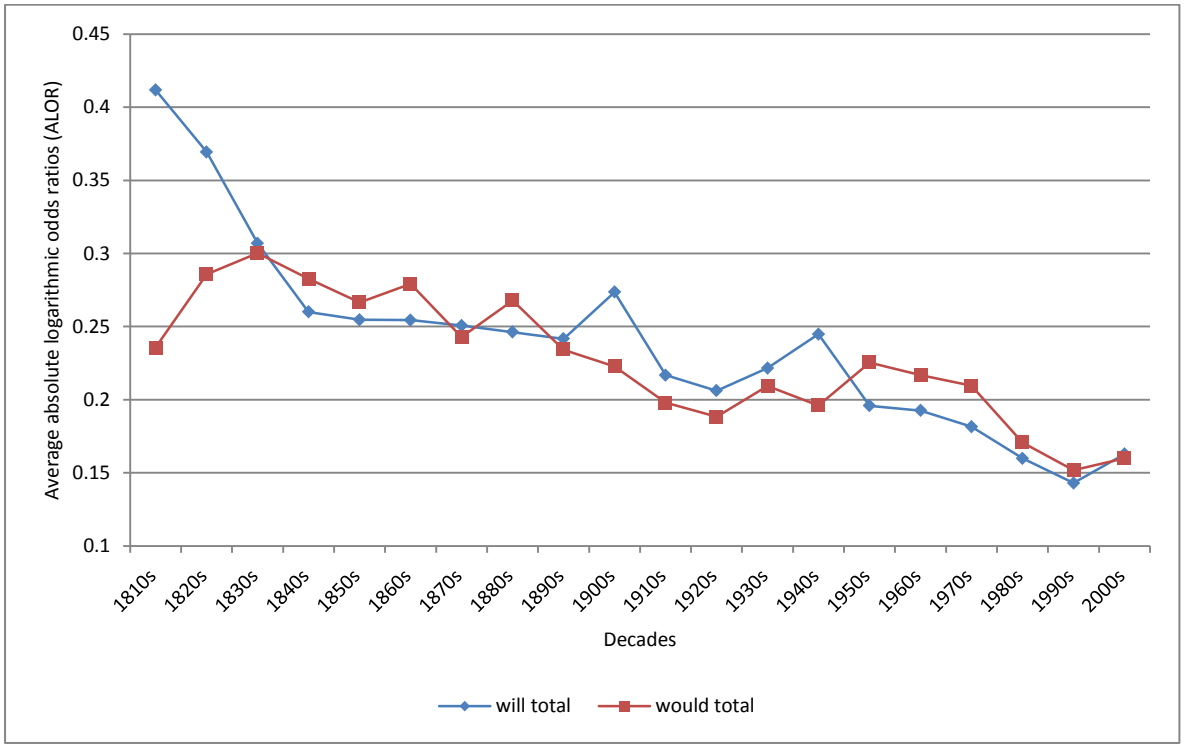

Figure 37. Average absolute log odds ratios of both positive and negative *will* and *would* constructions over time.

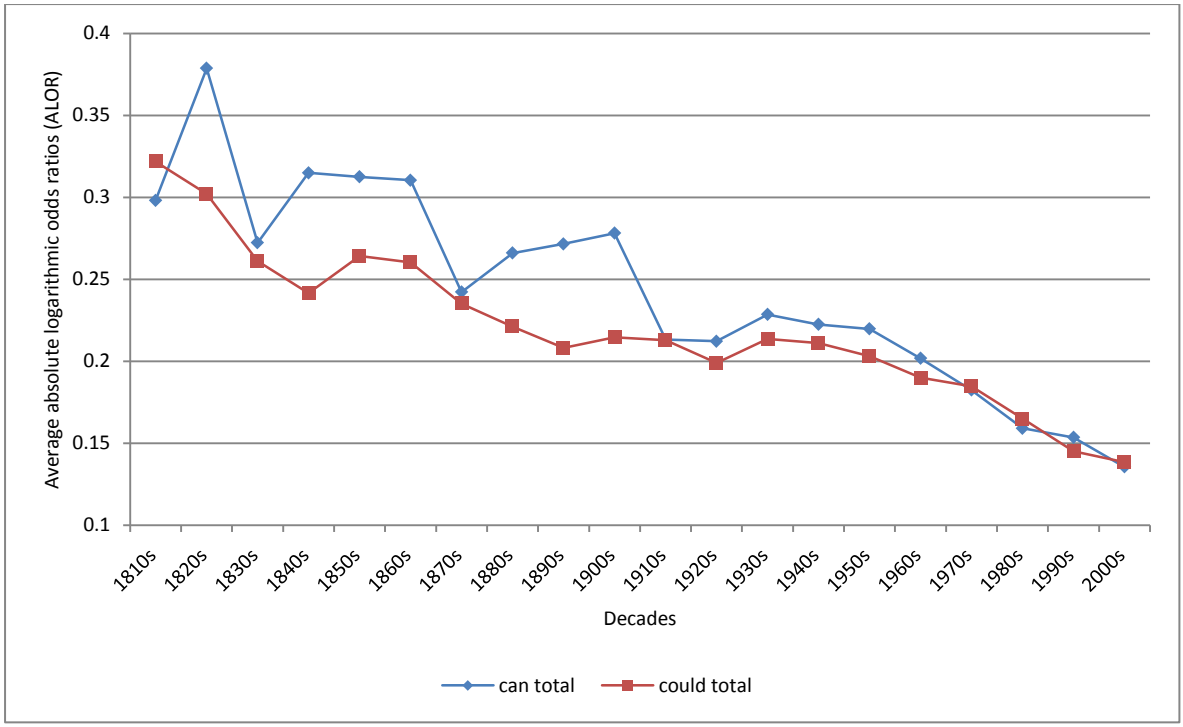

Figure 38. Average absolute log odds ratios of both positive and negative *can* and *could* constructions over time.

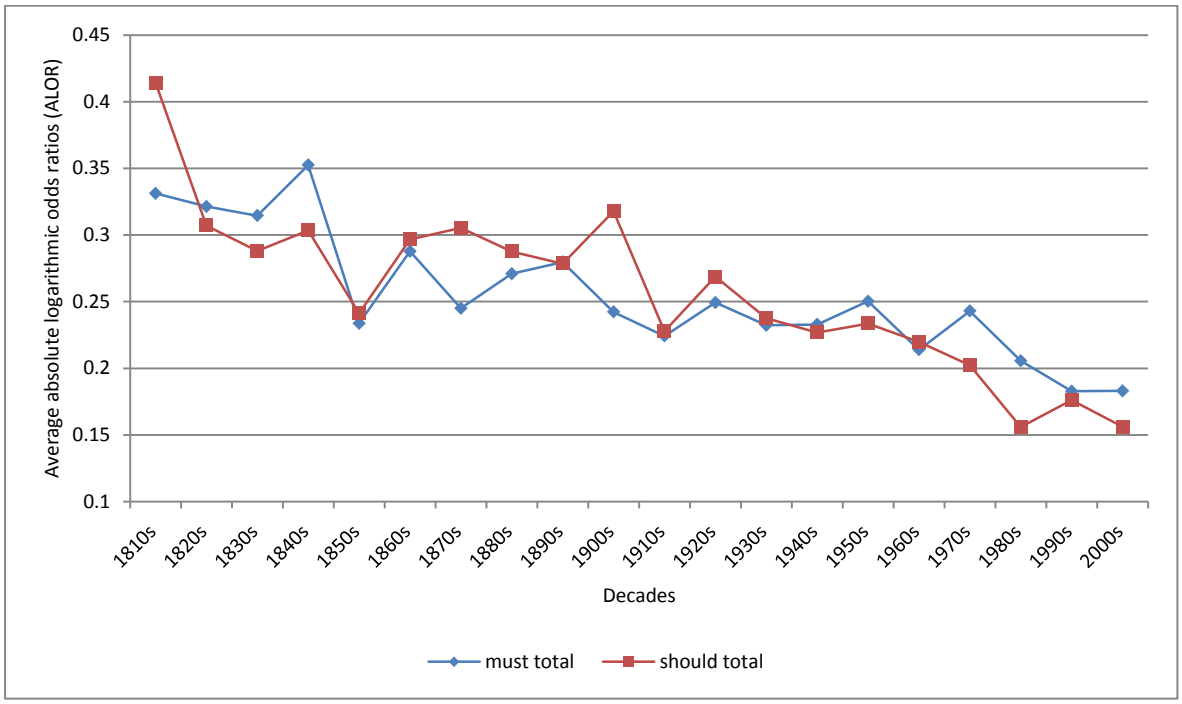

Figure 39. Average absolute log odds ratios of both positive and negative *must* and *should* constructions over time.

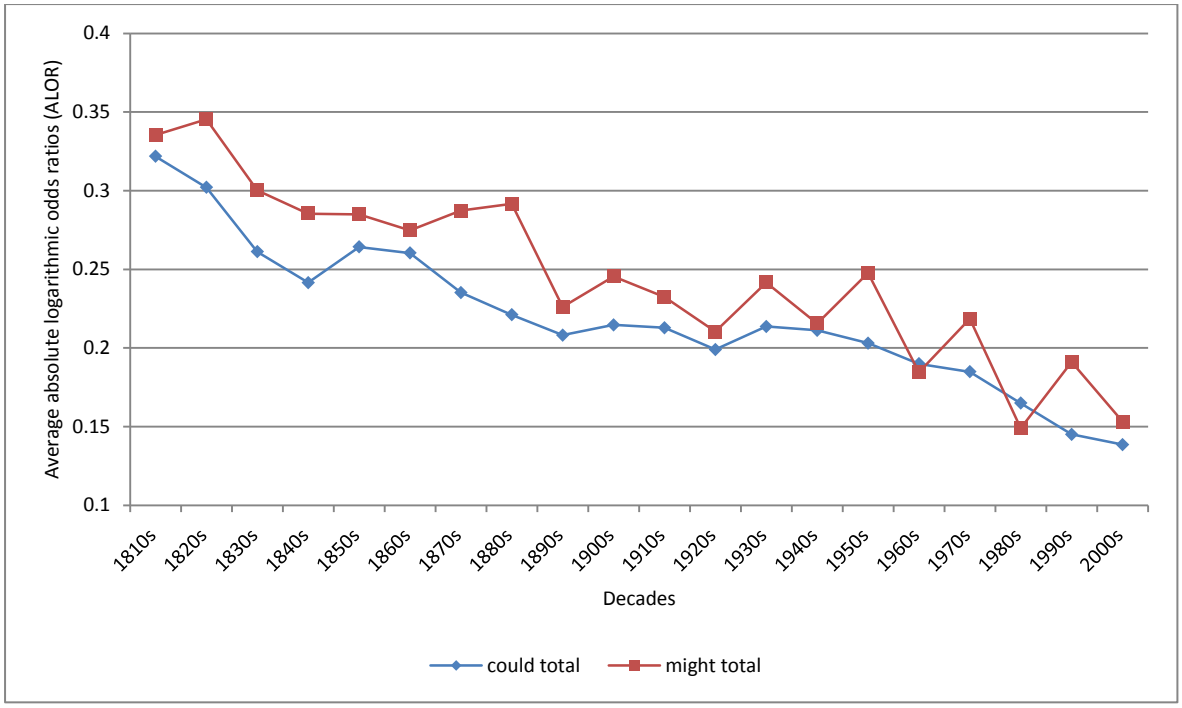

Figure 40. Average absolute log odds ratios of both positive and negative *could* and *might* constructions over time.

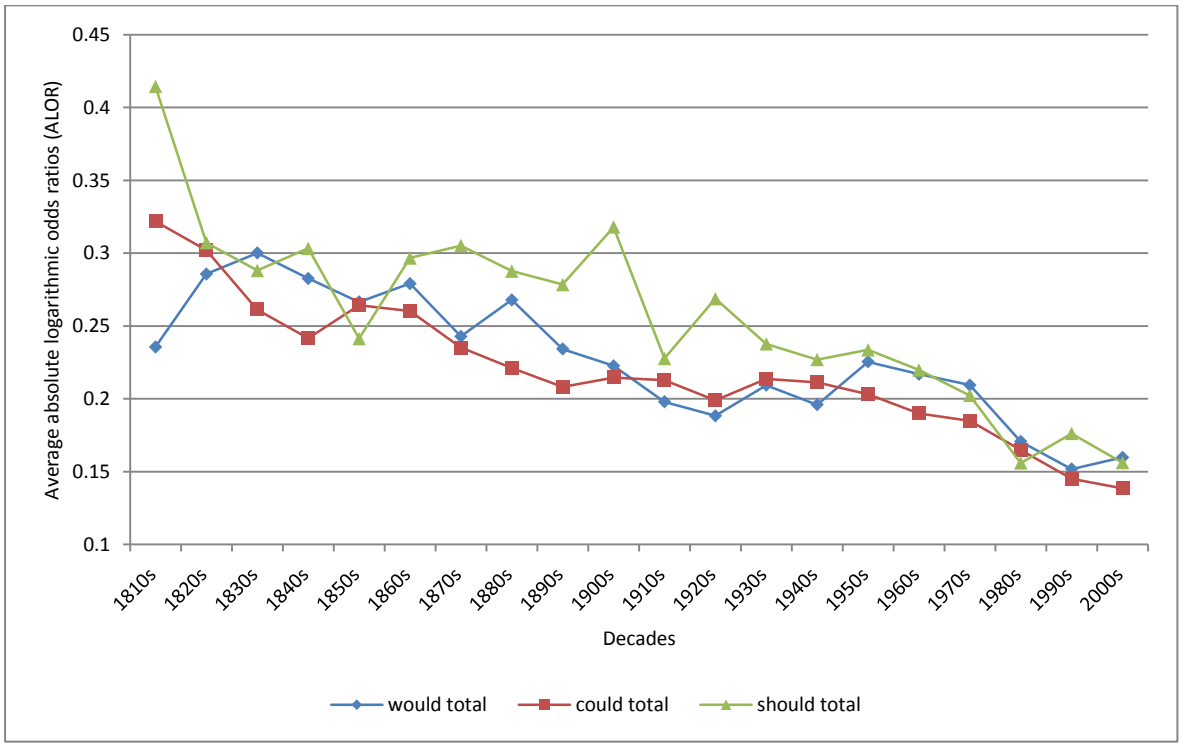

Figure 41. Average absolute log odds ratios of both positive and negative *would*, *could* and *should* constructions over time.

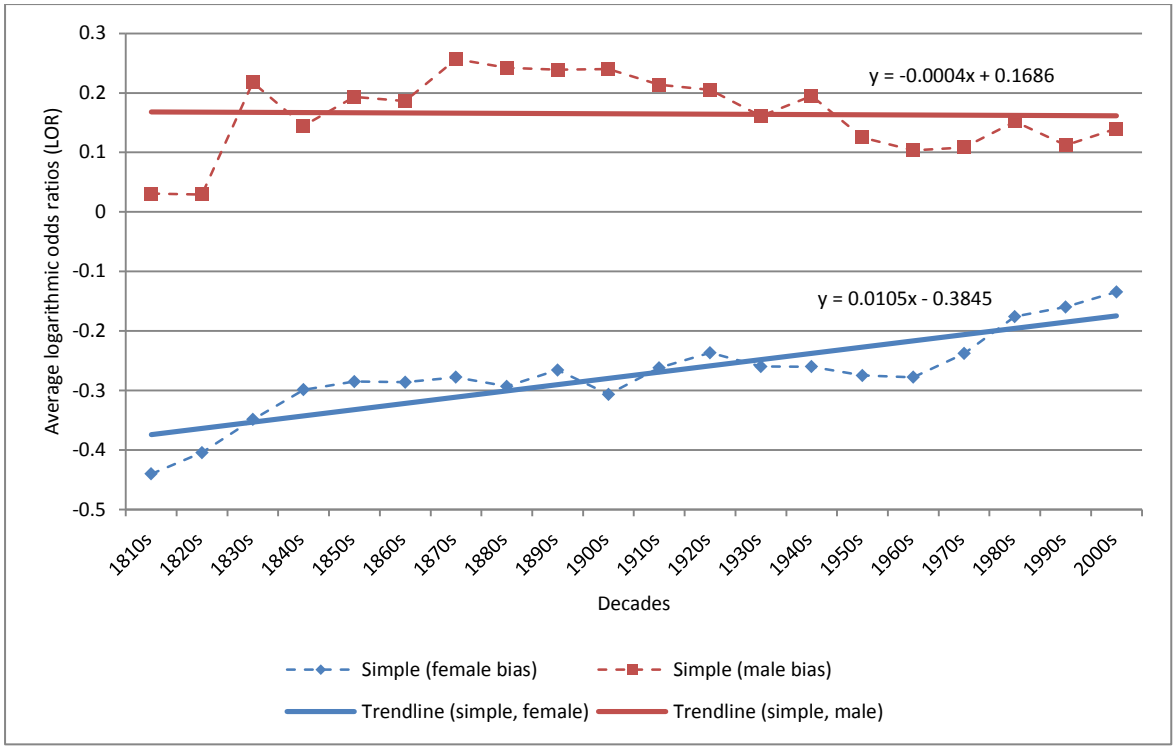

Figure 42. Average log odds ratio trajectories and accompanying trendlines for simple tense constructions.

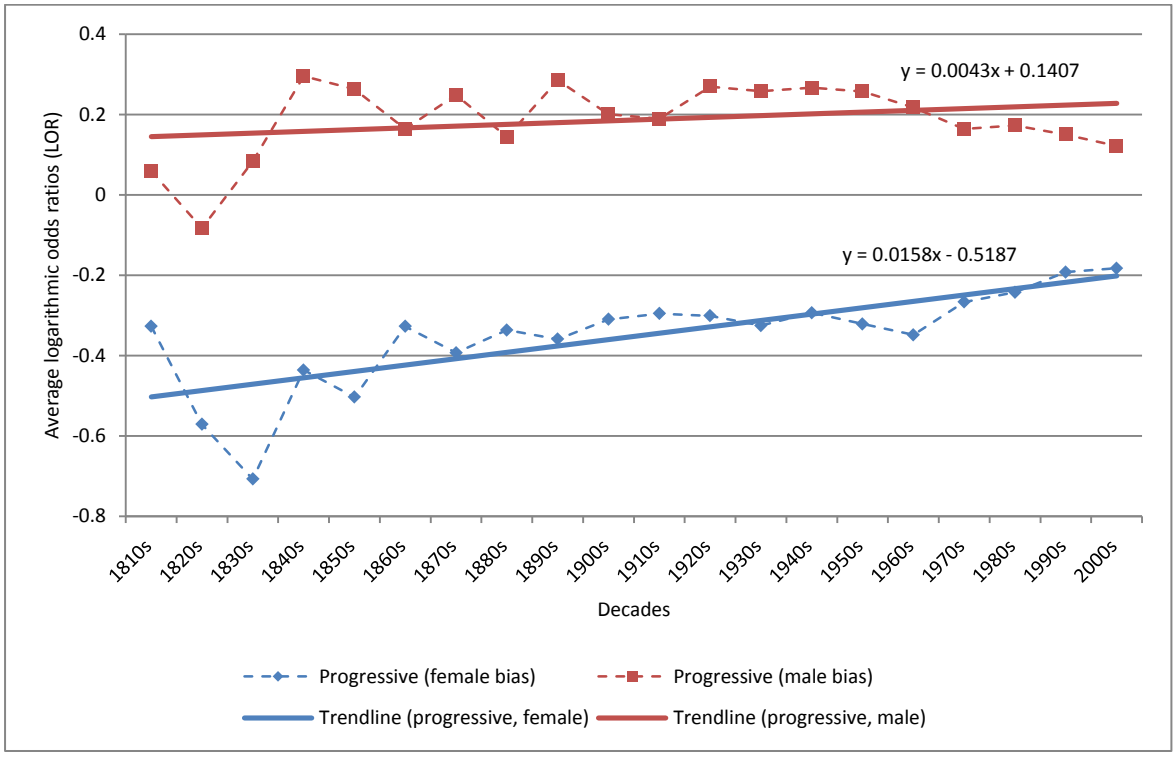

Figure 43. Average log odds ratio trajectories and accompanying trendlines for progressive constructions.

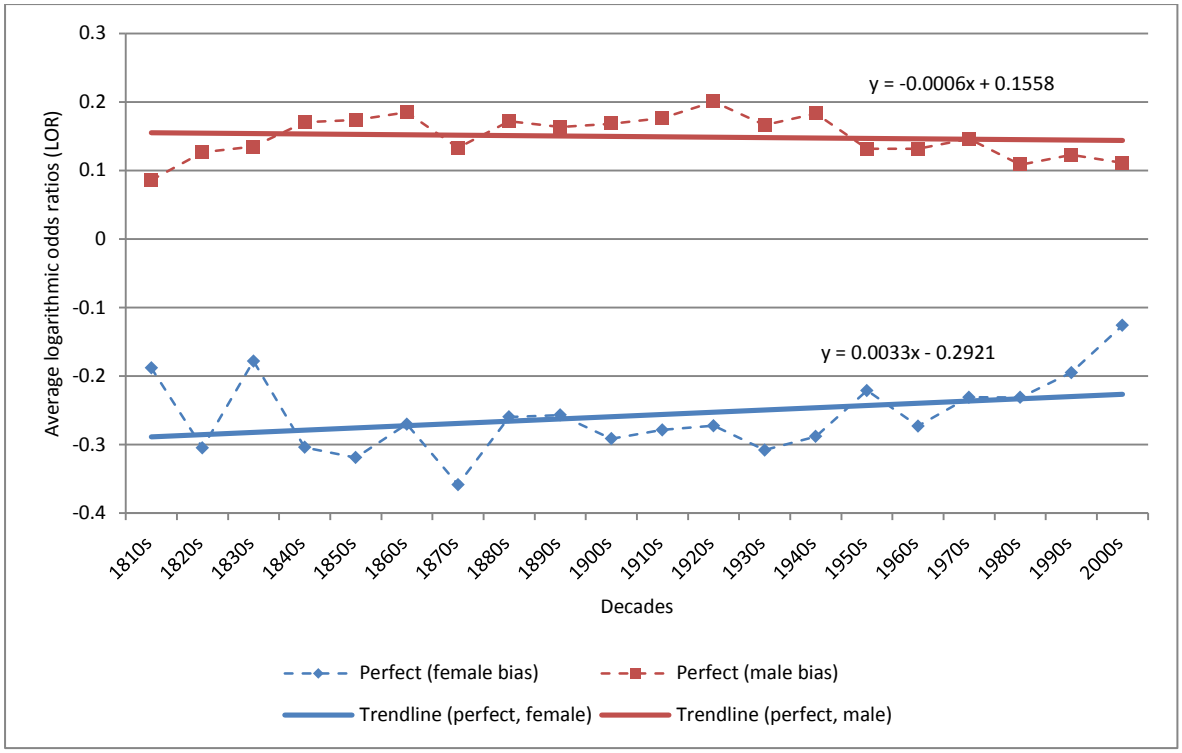

Figure 44. Average log odds ratio trajectories and accompanying trendlines for perfect constructions.

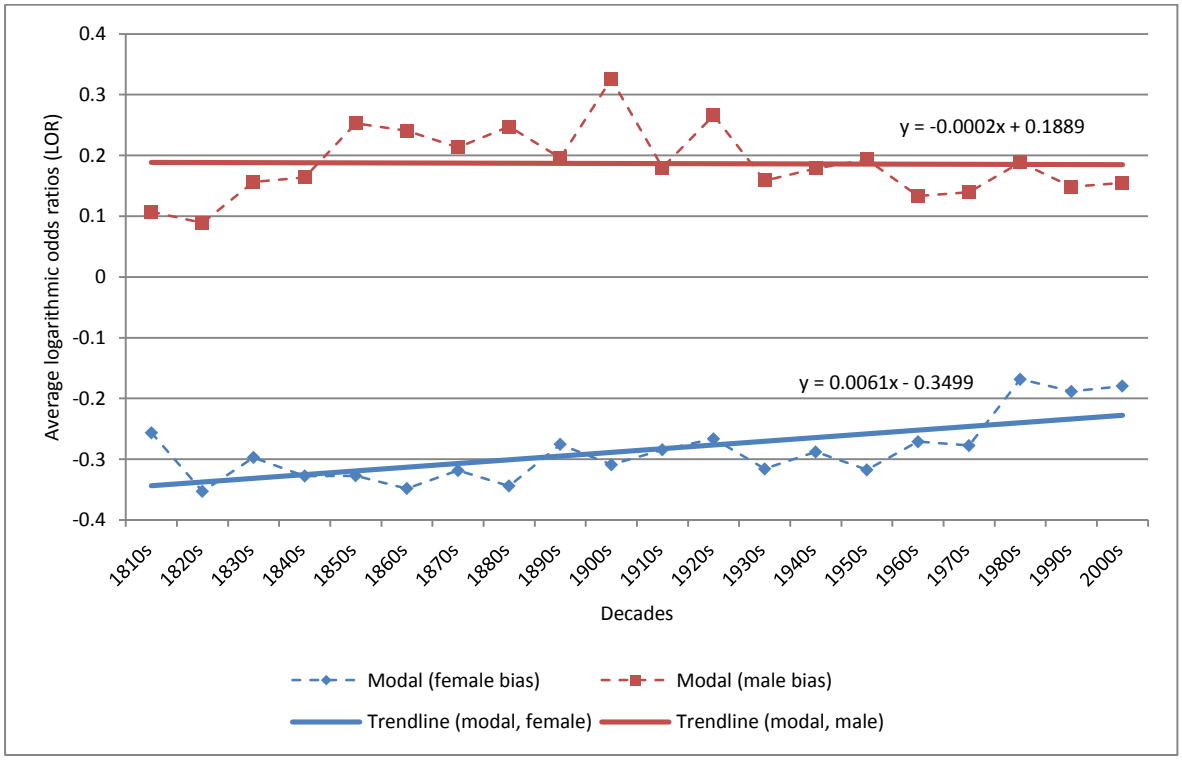

Figure 45. Average log odds ratio trajectories and accompanying trendlines for modal constructions.

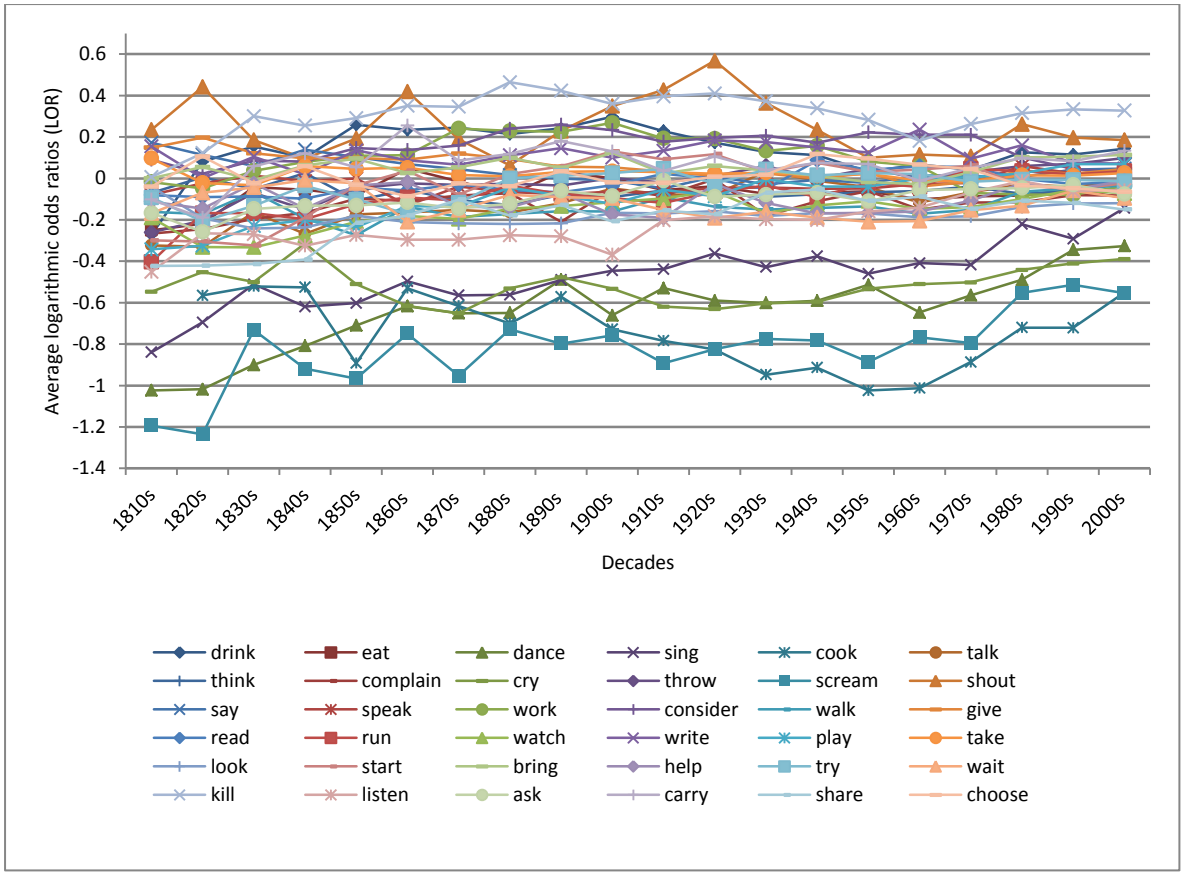

Figure 46. Average log odds ratio trajectories of the 36 verbs taking all 26 constructions into consideration.

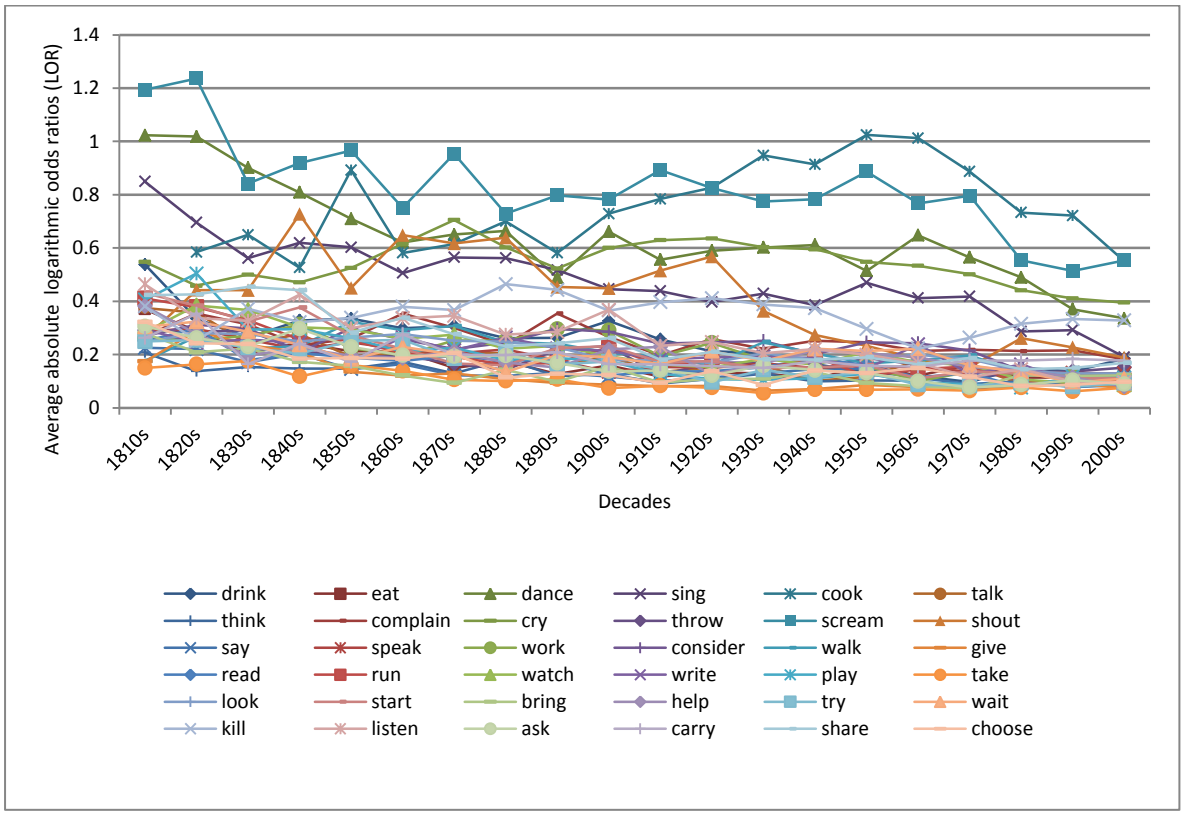

Figure 47. Average absolute log odds ratio trajectories of the 36 verbs taking all 26 constructions into consideration.

Table 7. Total number of male constructions per decade in GBA, 1810s-1900s.

Construction (query)	1810s (male)	1820s (male)	1830s (male)	1840s (male)	1850s (male)	1860s (male)	1870s (male)	1880s (male)	1890s (male)	1900s (male)
he|He|HE [vvz*]	222,663	339,809	725,474	991,615	1,429,392	1,060,514	1,269,126	2,006,859	2,461,932	3,209,335
he|He|HE does not [v*]	5,151	8,506	19,219	27,740	40,303	32,114	39,092	63,469	84,323	113,063
he|He|HE [vvd*]	289,166	486,250	1,162,489	1,570,868	2,527,091	2,116,494	2,584,065	4,039,583	6,036,337	7,909,973
he|He|HE did not [v*]	12,216	19,576	47,440	65,770	103,037	84,285	100,850	163,429	228,124	299,566
he|He|HE will [v*]	25,622	39,771	86,944	114,664	165,689	125,720	143,176	208,334	282,128	363,712
he|He|HE will not [v*]	3,362	5,399	11,395	15,154	21,125	15,629	17,617	26,006	33,683	42,355
he|He|HE is [v?g*]	2,863	4,787	12,282	18,387	29,727	26,050	32,648	52,156	79,725	115,812
he|He|HE is not [v?g*]	79	135	408	620	880	908	1,079	1,928	2,973	4,308
he|He|HE was [v?g*]	7,619	13,896	35,972	50,319	83,782	77,481	90,379	149,727	235,479	317,017
he|He|HE was not [v?g*]	103	230	532	745	1,315	1,366	1,806	3,335	5,488	7,941
he|He|HE has [v?n*]	19,298	32,024	71,432	93,961	139,093	107,093	127,974	196,596	262,281	343,171
he|He|HE has not [v?n*]	874	1,468	3,224	4,344	6,438	5,010	5,482	8,533	11,596	15,103
he|He|HE had [v?n*]	40,681	71,017	169,991	234,645	370,128	296,453	336,220	528,741	713,699	879,982
he|He|HE had not [v?n*]	2,165	3,533	8,325	11,544	17,688	14,697	17,186	26,699	37,990	46,616
he|He|HE would [v*]	35,045	56,498	136,707	189,349	295,473	234,110	272,220	417,321	572,611	734,232
he|He|HE would not [v*]	5,782	9,615	22,071	29,501	43,853	33,374	39,102	59,638	77,882	100,286
he|He|HE can [v*]	8,214	13,431	31,866	43,186	64,355	49,166	60,780	97,028	128,552	177,253
he|He|HE cannot [v*]	6,863	10,528	23,007	30,374	38,235	27,085	34,532	54,296	66,074	85,243
he|He|HE could [v*]	13,546	23,639	62,310	85,098	137,580	116,587	133,534	217,518	305,245	391,479
he|He|HE could not [v*]	9,998	16,893	40,837	56,905	84,153	67,496	80,192	122,136	163,825	206,316
he|He|HE must [v*]	12,213	20,374	49,154	68,333	101,374	78,234	95,492	151,928	212,162	283,878
he|He|HE must not [v*]	439	686	1,688	2,422	3,723	2,893	3,490	5,646	8,702	11,505
he|He|HE might [v*]	21,191	32,657	71,634	91,908	134,769	97,427	107,423	160,839	202,833	246,215
he|He|HE might not [v*]	971	1,559	3,336	4,084	5,906	4,550	4,508	7,152	9,452	10,738
he|He|HE should [v*]	27,126	42,840	90,996	118,862	170,240	118,126	133,325	195,884	236,631	304,775
he|He|HE should not [v*]	1,726	3,002	6,418	8,283	12,140	8,501	9,820	14,617	18,504	24,198

Table 8. Total number of male constructions per decade in GBA, 1910s-2000s.

Construction (query)	1910s (male)	1920s (male)	1930s (male)	1940s (male)	1950s (male)	1960s (male)	1970s (male)	1980s (male)	1990s (male)	2000s (male)
he|He|HE [vvz*]	4,022,786	2,977,995	2,352,622	2,365,892	3,196,482	5,398,857	5,015,922	4,484,767	5,601,453	7,420,381
he|He|HE does not [v*]	138,846	104,012	79,781	86,033	119,812	198,462	178,894	139,551	154,830	177,299
he|He|HE [vvd*]	9,685,620	7,005,877	6,095,117	6,429,543	7,736,310	10,783,339	9,621,961	10,683,388	14,809,733	22,843,133
he|He|HE did not [v*]	357,495	263,805	223,119	224,272	282,977	423,084	354,315	323,792	370,926	471,717
he|He|HE will [v*]	440,694	320,574	235,981	247,411	335,096	499,286	461,053	330,614	362,411	451,282
he|He|HE will not [v*]	47,745	33,310	24,260	24,457	32,495	49,520	44,812	33,919	35,842	45,064
he|He|HE is [v?g*]	171,252	136,807	109,712	118,988	168,028	276,657	278,737	212,778	239,410	303,359
he|He|HE is not [v?g*]	7,033	5,660	4,504	5,269	7,466	12,901	13,719	10,635	12,319	15,669
he|He|HE was [v?g*]	432,631	343,360	329,861	358,725	448,595	640,996	623,897	729,120	1,057,049	1,697,802
he|He|HE was not [v?g*]	11,799	10,233	9,437	9,913	13,453	20,260	18,090	18,401	22,633	32,528
he|He|HE has [v?n*]	459,203	292,902	201,413	188,992	241,244	399,355	366,341	299,691	325,908	407,783
he|He|HE has not [v?n*]	16,890	11,566	8,285	8,320	10,700	17,794	15,825	12,125	12,654	14,643
he|He|HE had [v?n*]	1,006,514	744,152	650,865	636,337	786,958	1,174,519	973,122	939,738	1,109,090	1,472,084
he|He|HE had not [v?n*]	54,637	40,294	32,758	31,694	39,892	57,513	46,659	43,238	48,520	61,945
he|He|HE would [v*]	860,321	665,645	572,222	562,435	701,607	1,061,076	949,514	925,758	1,168,206	1,642,874
he|He|HE would not [v*]	109,568	79,786	65,303	61,057	78,077	119,704	102,822	90,061	106,393	139,653
he|He|HE can [v*]	248,341	188,475	150,790	175,837	247,637	392,029	379,494	262,366	280,594	377,881
he|He|HE cannot [v*]	107,373	75,049	54,810	58,381	81,211	133,490	114,432	79,695	80,109	91,174
he|He|HE could [v*]	483,944	375,207	356,320	384,908	473,258	675,514	618,374	649,620	894,157	1,424,785
he|He|HE could not [v*]	235,549	171,351	144,034	132,324	164,362	245,123	197,384	172,104	194,967	247,980
he|He|HE must [v*]	379,501	284,637	226,797	226,086	287,287	424,041	351,708	245,359	250,897	319,053
he|He|HE must not [v*]	14,927	10,599	7,899	7,502	9,310	12,856	9,860	6,862	7,270	8,758
he|He|HE might [v*]	268,209	191,712	157,724	144,493	179,546	272,768	230,080	200,123	250,432	381,999
he|He|HE might not [v*]	11,310	8,150	6,397	5,973	7,633	11,369	9,998	9,637	13,152	21,961
he|He|HE should [v*]	331,723	234,830	177,613	171,314	224,742	331,118	277,271	196,269	215,981	302,431
he|He|HE should not [v*]	26,502	18,967	14,176	13,883	18,342	27,723	23,605	16,080	15,790	19,011

Table 9. Total number of female constructions per decade in GBA, 1810s-1900s.

Construction (query)	1810s (female)	1820s (female)	1830s (female)	1840s (female)	1850s (female)	1860s (female)	1870s (female)	1880s (female)	1890s (female)	1900s (female)
she|She|SHE [vvz*]	20,997	36,659	84,501	127,866	212,322	172,603	216,800	348,529	472,912	623,008
she|She|SHE does not [v*]	378	805	2,110	3,284	5,303	5,144	6,380	10,024	14,524	18,876
she|She|SHE [vvd*]	40,133	95,903	265,111	394,073	694,575	702,147	889,753	1,290,090	2,235,336	2,965,100
she|She|SHE did not [v*]	1,277	3,082	8,309	12,560	23,432	24,817	30,473	44,563	74,723	94,742
she|She|SHE will [v*]	2,355	4,697	11,892	15,932	27,735	26,180	29,930	44,619	67,103	82,631
she|She|SHE will not [v*]	337	649	1,576	2,085	3,621	3,327	4,072	5,980	8,545	10,022
she|She|SHE is [v?g*]	332	680	2,167	3,057	6,117	6,515	8,543	13,441	21,458	28,891
she|She|SHE is not [v?g*]	1	6	27	56	129	139	184	323	519	785
she|She|SHE was [v?g*]	1,210	3,033	8,005	12,362	22,550	25,227	31,439	48,128	88,258	118,854
she|She|SHE was not [v?g*]	18	33	100	176	345	528	758	1,146	2,152	2,992
she|She|SHE has [v?n*]	1,838	3,782	9,974	14,072	23,874	20,570	24,503	37,674	56,606	70,899
she|She|SHE has not [v?n*]	73	193	427	726	1,150	1,040	1,187	1,750	2,714	3,146
she|She|SHE had [v?n*]	5,360	12,615	34,058	50,129	83,220	80,006	95,214	138,729	218,197	276,220
she|She|SHE had not [v?n*]	251	669	1,601	2,418	4,048	4,272	5,241	8,073	13,270	16,469
she|She|SHE would [v*]	4,438	9,042	23,461	36,254	63,120	61,138	73,624	107,381	169,843	214,504
she|She|SHE would not [v*]	753	1,517	3,761	5,391	9,288	9,191	11,221	16,056	25,583	32,264
she|She|SHE can [v*]	866	1,506	3,997	5,405	9,634	8,986	10,548	16,822	24,054	31,222
she|She|SHE cannot [v*]	589	1,132	2,466	3,484	5,339	4,369	5,588	8,685	11,121	13,220
she|She|SHE could [v*]	1,998	4,600	12,849	18,576	32,753	33,176	40,663	57,393	92,762	117,377
she|She|SHE could not [v*]	1,532	3,594	9,638	13,704	22,995	23,481	28,476	38,145	59,239	71,393
she|She|SHE must [v*]	1,357	2,785	7,839	12,004	20,796	20,115	24,898	36,549	57,132	74,947
she|She|SHE must not [v*]	62	108	388	594	941	1,007	1,136	1,815	2,665	3,335
she|She|SHE might [v*]	2,128	4,611	11,376	15,869	24,981	22,560	26,859	36,982	56,846	69,502
she|She|SHE might not [v*]	122	251	584	810	1,334	1,309	1,409	1,835	3,330	3,670
she|She|SHE should [v*]	2,712	5,687	13,621	18,419	28,996	24,874	28,630	40,550	57,007	70,821
she|She|SHE should not [v*]	223	494	1,149	1,639	2,321	1,998	2,329	3,264	4,412	5,520

Table 10. Total number of female constructions per decade in GBA, 1910s-2000s.

Construction (query)	1910s (female)	1920s (female)	1930s (female)	1940s (female)	1950s (female)	1960s (female)	1970s (female)	1980s (female)	1990s (female)	2000s (female)
she|She|SHE [vvz*]	803,304	619,064	517,147	501,196	634,268	986,498	1,197,713	1,662,839	2,721,288	3,894,189
she|She|SHE does not [v*]	22,432	17,130	13,106	12,882	16,440	25,678	32,302	41,045	60,008	74,951
she|She|SHE [vvd*]	3,659,276	2,669,606	2,234,101	2,309,091	2,593,195	3,117,404	3,373,707	5,246,585	8,612,431	14,171,781
she|She|SHE did not [v*]	119,570	90,261	69,720	66,846	76,801	91,288	95,183	118,070	163,652	216,850
she|She|SHE will [v*]	99,188	69,329	51,000	50,938	59,926	83,850	95,976	122,265	169,974	229,134
she|She|SHE will not [v*]	10,925	7,516	5,556	5,380	6,150	8,912	9,617	11,482	15,590	19,511
she|She|SHE is [v?g*]	42,526	32,906	25,312	25,463	31,338	46,979	63,583	87,342	131,072	181,286
she|She|SHE is not [v?g*]	1,229	1,029	745	832	952	1,563	2,342	3,322	4,947	6,895
she|She|SHE was [v?g*]	168,483	133,595	120,929	131,290	159,957	192,594	221,212	343,706	579,094	967,679
she|She|SHE was not [v?g*]	4,762	3,753	3,424	3,368	4,554	5,089	5,275	7,589	11,617	16,668
she|She|SHE has [v?n*]	89,829	59,779	44,467	39,535	46,424	71,804	84,610	104,789	154,070	198,852
she|She|SHE has not [v?n*]	3,918	2,449	1,585	1,544	2,003	2,908	3,385	4,081	5,563	6,920
she|She|SHE had [v?n*]	333,494	251,742	205,729	197,937	226,289	285,701	291,983	397,103	556,374	752,785
she|She|SHE had not [v?n*]	20,830	14,888	12,074	10,902	12,612	15,501	14,623	18,906	25,547	33,085
she|She|SHE would [v*]	265,166	209,725	179,343	170,460	197,255	256,634	274,812	388,450	578,536	842,857
she|She|SHE would not [v*]	37,341	27,286	22,267	19,681	22,857	28,768	28,917	36,062	50,403	69,143
she|She|SHE can [v*]	45,636	33,432	27,538	29,864	37,615	54,580	72,927	100,131	152,420	225,505
she|She|SHE cannot [v*]	17,146	11,967	9,268	8,448	10,671	17,321	19,812	24,321	33,935	39,909
she|She|SHE could [v*]	153,548	120,233	112,285	116,200	129,131	156,326	183,000	296,714	506,841	850,220
she|She|SHE could not [v*]	86,738	64,044	49,908	44,349	49,346	59,957	60,177	74,016	99,906	127,642
she|She|SHE must [v*]	103,391	76,501	62,831	56,639	63,138	80,827	83,664	99,252	131,556	172,527
she|She|SHE must not [v*]	4,871	3,568	2,898	2,494	2,569	3,061	2,910	3,025	3,599	4,645
she|She|SHE might [v*]	77,951	57,260	46,720	40,738	46,468	62,040	64,296	84,887	131,183	214,069
she|She|SHE might not [v*]	4,052	2,877	2,210	1,998	2,329	2,946	3,224	4,411	7,650	13,293
she|She|SHE should [v*]	80,906	57,069	42,131	38,435	45,720	63,000	67,847	82,256	117,291	179,540
she|She|SHE should not [v*]	6,175	4,262	3,425	3,028	3,666	4,984	5,323	5,797	7,394	9,596

Table 11. Summary table for the positive simple present construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink	-0.48849	-0.35561	-0.20636	-0.13573	-0.09189	0.02621	-0.06160	-0.17291	-0.09544	-0.07740	0.01439	-0.03683	-0.07322	0.00648	-0.05865	-0.08562	-0.01989	0.03806	0.06090	0.04992
eat	0.18499	0.01986	0.14262	-0.14565	0.02501	-0.06109	-0.10073	-0.01885	-0.03478	0.03374	0.02313	-0.00592	0.00287	0.03732	-0.03429	-0.03086	-0.03434	-0.04799	-0.04413	-0.09661
dance	-1.30237	-1.04584	-1.00434	-1.02392	-1.01767	-0.92050	-0.85940	-0.75417	-0.85297	-0.86887	-0.81899	-0.83185	-0.68234	-0.66727	-0.74797	-0.65281	-0.53806	-0.44246	-0.36905	-0.37664
sing	-0.97386	-0.90272	-0.80294	-0.71545	-0.70143	-0.67000	-0.62653	-0.52565	-0.47418	-0.41974	-0.39117	-0.40549	-0.39137	-0.46514	-0.50872	-0.46990	-0.36911	-0.31506	-0.20250	-0.20581
cook		-1.26811		-0.55861	-0.71705	-0.40828	-0.44796	-0.53521	-0.55255	-0.39929	-0.73821	-0.74926	-0.64072	-0.82293	-0.78366	-0.86704	-0.70133	-0.50962	-0.42756	-0.30013
talk	-0.18942	-0.14603	-0.14546	-0.17404	-0.20089	-0.07786	-0.18077	-0.16418	-0.13833	-0.16534	-0.11584	-0.10544	-0.07531	-0.08690	-0.10754	-0.10260	-0.11508	-0.05846	-0.03334	-0.00712
think	0.07091	0.07982	-0.00226	0.04216	0.02090	-0.04258	-0.05073	-0.01427	-0.04107	-0.02883	-0.01971	-0.01156	-0.01820	-0.03058	-0.01455	0.00619	-0.01148	-0.04855	-0.03663	-0.02836
complain	0.01511	0.22711	0.11445	0.13545	0.00046	-0.07063	-0.03426	-0.00926	0.10904	-0.00190	0.03128	0.06840	0.04983	-0.00067	-0.06183	-0.10538	-0.14445	-0.10829	-0.08877	-0.12559
cry	-0.47918	-0.38578	-0.33467	-0.24133	-0.31020	-0.35738	-0.43754	-0.42223	-0.38686	-0.33757	-0.38608	-0.34645	-0.35456	-0.32371	-0.35693	-0.40798	-0.36360	-0.34553	-0.29817	-0.28464
throw	-0.41867	-0.30554	-0.24847	-0.18964	-0.22019	-0.23949	-0.20332	-0.19786	-0.17038	-0.19883	-0.26817	-0.25812	-0.23590	-0.24368	-0.21267	-0.23883	-0.15228	-0.10878	-0.00261	-0.00990
scream		-1.53145	-1.21013	-0.93015	-0.89523	-0.86327	-0.93037	-0.96276	-0.85212	-0.84910	-0.96899	-0.90292	-0.80684	-0.78058	-0.75733	-0.82588	-0.66867	-0.56863	-0.41993	-0.36392
shout	-0.02549	0.43091	0.32950	0.34408	0.19781	0.29590	0.58591	0.71878	0.77014	0.47752	0.53288	0.34942	0.36243	0.33057	0.09684	-0.02834	0.03893	0.12331	0.09891	0.11241
say	0.35174	0.34050	0.31026	0.24376	0.27368	0.24227	0.21043	0.16538	0.16858	0.14964	0.11826	0.13306	0.09138	0.04225	-0.00358	0.01454	-0.03573	-0.01608	-0.02114	-0.03306
speak	0.23285	0.07780	0.14361	0.14621	0.11589	0.13550	0.11785	0.12180	0.12385	0.10837	0.11248	0.07147	0.07328	0.12100	0.12534	0.12224	0.11945	0.14539	0.12520	0.14186
work	0.27591	-0.04605	-0.01286	0.13668	-0.01543	-0.05031	0.03798	0.01965	0.06457	0.02357	0.02110	0.03757	0.03307	0.02200	0.01736	-0.02280	-0.09889	-0.14992	-0.13828	-0.12898
consider	0.24106	0.36400	0.31520	0.33253	0.37885	0.30621	0.37450	0.44697	0.39107	0.42069	0.33973	0.37437	0.35274	0.32915	0.33465	0.32648	0.25668	0.20109	0.15597	0.14820
walk	-0.33237	-0.19396	-0.27414	-0.29837	-0.26484	-0.30720	-0.29864	-0.30148	-0.27637	-0.28513	-0.22458	-0.26160	-0.21933	-0.18618	-0.25274	-0.26603	-0.21165	-0.13117	-0.09970	-0.10137
give	0.14021	0.13009	0.13648	0.10233	0.11899	0.14389	0.15369	0.14375	0.09159	0.07165	0.06823	0.03032	0.01649	0.01489	-0.00394	0.00694	0.00643	0.06180	0.06508	0.06799
read	-0.21249	-0.07120	-0.09456	-0.05417	-0.11697	-0.14588	-0.09709	-0.09450	-0.11567	-0.06602	-0.04040	0.02829	0.04654	0.03978	0.04317	0.04895	0.02518	-0.05661	-0.08627	-0.10172
run	-0.22761	-0.29872	-0.23259	-0.16739	-0.18872	-0.15297	-0.16811	-0.18231	-0.22515	-0.24065	-0.29573	-0.31591	-0.30418	-0.30530	-0.30072	-0.31210	-0.21522	-0.19067	-0.11049	-0.09621
watch	-0.21131	-0.22593	-0.30345	-0.42911	-0.37541	-0.37808	-0.37041	-0.33600	-0.24541	-0.25406	-0.25538	-0.20265	-0.19589	-0.20560	-0.22844	-0.22578	-0.18186	-0.14816	-0.11809	-0.09221
write	0.07620	-0.00839	0.09977	-0.01361	0.03052	0.09860	0.17957	0.06726	0.13239	0.11093	0.16881	0.15000	0.15029	0.14024	0.17012	0.20259	0.14660	0.16042	0.11896	0.13712
play	-0.56114	-0.33507	-0.34140	-0.33254	-0.31366	-0.28604	-0.29831	-0.24813	-0.19727	-0.20378	-0.17962	-0.18926	-0.06857	-0.03888	-0.09847	-0.09486	-0.05795	-0.00186	0.03045	0.06845
take	-0.03815	-0.09435	-0.02419	-0.03935	-0.05762	-0.07141	-0.07370	-0.07226	-0.07752	-0.07623	-0.06580	-0.07624	-0.06323	-0.07754	-0.08194	-0.07854	-0.03490	-0.01059	0.01631	0.00714
look	-0.30994	-0.34232	-0.35617	-0.33632	-0.32837	-0.36142	-0.39269	-0.37933	-0.34332	-0.32647	-0.32633	-0.31779	-0.29997	-0.31134	-0.29530	-0.29313	-0.24475	-0.19965	-0.13485	-0.13708
start	-0.48559	-0.43651	-0.51262	-0.40708	-0.23678	-0.22600	-0.28142	-0.17171	-0.13145	-0.13545	-0.14211	-0.17868	-0.18583	-0.22171	-0.17421	-0.17278	-0.10499	-0.09755	-0.07728	-0.06233
bring	-0.13693	-0.24477	-0.11316	-0.12546	-0.09841	-0.11530	-0.06948	-0.06679	-0.07022	-0.05803	-0.08155	-0.07460	-0.07603	-0.05124	-0.04370	-0.06866	-0.03410	0.02157	0.01184	0.01064
help	0.26087	-0.09618	-0.03068	0.12779	-0.05556	-0.08834	-0.07951	-0.19395	-0.21946	-0.17755	-0.23739	-0.19601	-0.29794	-0.25231	-0.31209	-0.26235	-0.23070	-0.18191	-0.18875	-0.14307
try	-0.43539	-0.13356	-0.26646	-0.17932	-0.20128	-0.14111	-0.14369	-0.07430	-0.06473	-0.04376	-0.02146	-0.07422	-0.02539	0.00104	-0.01770	-0.01796	-0.02154	-0.01642	-0.01897	-0.02194
wait	-0.11179	0.03745	-0.13026	-0.28100	-0.23625	-0.27121	-0.31862	-0.30401	-0.37495	-0.34062	-0.38571	-0.33384	-0.31927	-0.28729	-0.31645	-0.31441	-0.26170	-0.19070	-0.15035	-0.12389
kill	0.09327	0.08678	0.22826	0.07677	0.16251	0.23208	0.30899	0.20168	0.19233	0.17676	0.18037	0.21424	0.13186	0.10216	0.12298	0.16092	0.17479	0.20965	0.29026	0.35563
listen	-0.63237	-0.00001	-0.42928	-0.29879	-0.31883	-0.34844	-0.27357	-0.35558	-0.30425	-0.31134	-0.31078	-0.27321	-0.24157	-0.14058	-0.16028	-0.13370	-0.08673	-0.12064	-0.08802	-0.10052
ask	0.02811	-0.07784	-0.00584	0.03121	-0.05645	-0.06721	-0.14173	-0.13898	-0.03472	-0.05002	-0.05903	-0.06580	-0.03664	-0.05642	-0.10776	-0.13288	-0.13962	-0.09930	-0.09244	-0.09963
carry	-0.33238	-0.23706	-0.24722	-0.17729	-0.13980	-0.12991	-0.13136	-0.11594	-0.14120	-0.17041	-0.16684	-0.11976	-0.12396	-0.17087	-0.16479	-0.15788	-0.12231	-0.09984	-0.07946	-0.04395
share	-0.26594	-0.28294	-0.32623	-0.07627	-0.12870	-0.13963	-0.18043	-0.01651	-0.02589	0.03844	0.07084	0.11693	0.14145	0.14936	0.06843	-0.00786	-0.02347	-0.06105	-0.07609	-0.07001
choose	0.20839	0.13890	0.10604	0.06757	0.11305	0.01310	0.00631	0.01647	0.04178	0.02855	0.11425	0.12529	0.15367	0.16226	0.20124	0.24218	0.13745	0.01141	-0.06010	-0.05075

Table 12. Summary table for the negative simple present construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink			-0.05638	0.21951		0.01753	0.46827		0.54122	0.10638	0.84752	-0.01743	0.25726	-0.10740	-0.08456	0.05147	-0.02652	-0.00711	-0.04609	-0.12310
eat	-0.65908	-1.15493	-0.00521	0.53602	0.21057	0.07004	-0.00913	-0.07591	0.14717	-0.27064	-0.03244	-0.01994	-0.34241	-0.19096	-0.17571	-0.02249	-0.14518	-0.17394	-0.16409	-0.07105
dance		-1.32599			-0.61795	-0.63628	-0.11831	-0.88909	-0.55166	-0.69059	-0.57933	-0.63389	-0.86413	-0.68370	-0.79632	-0.22811	-0.59465	-0.37023	-0.41177	-0.56493
sing				-0.32485	-0.34602	-1.06524	-0.73745	-0.53883	-0.46920	-0.60505	-0.39416	-0.17507	-0.63920	-0.26345	-0.48113	-0.32411	-0.54241	-0.10879	-0.12729	-0.15595
cook										-1.55570	-0.31458	-1.12601	-1.13689	-0.82477	-1.16373	-1.92969		-1.09585	-1.41185	-0.67508
talk		-0.28409	-0.14019	-0.00590	-0.29617	-0.09663	-0.18563	-0.15480	0.03209	0.10797	-0.14888	-0.04850	-0.11138	0.01831	-0.04481	-0.11923	-0.20916	-0.08393	-0.04180	-0.08983
think	-0.00549	0.20556	-0.05358	-0.03341	-0.10779	-0.00242	-0.00245	0.10676	0.02809	0.01189	0.05605	0.03727	0.01274	-0.01489	-0.02857	0.10991	0.02588	0.01962	0.05763	0.08216
complain	-0.05536	0.07318	-0.42862	-0.74512	-0.43381	-0.60386	-0.44751	-0.49227	-0.58206	-0.17822	-0.29674	-0.27182	-0.00178	-0.44971	-0.29309	-0.32522	-0.32882	-0.23068	-0.35722	-0.24471
cry	-0.83435			-0.05167	-0.84380	-0.61982	-0.88536	-0.48448	-0.47210	-0.69271	-0.68834	-0.14246	-0.51457	-0.28647	-0.57980	-0.41134	-0.23975	-0.36958	-0.26751	-0.40210
throw			0.67479	0.16456	0.15073	0.72348	0.14737	0.00011	0.13439	0.17353	0.15121	0.11986	0.04669	-0.03076	0.31846	-0.01722	0.30985	0.21161	0.11668	0.21945
scream																				
shout								-1.27877		-0.47655	-0.09270				0.11514	0.31604	0.03479	0.52928		0.59616
say	0.62441	0.40580	0.43432	0.55282	0.48822	0.48730	0.28345	0.38032	0.31003	0.36469	0.36596	0.22335	0.10931	0.26085	0.22491	0.27404	0.18277	0.28054	0.25774	0.28312
speak		0.14863	0.38483	-0.03149	0.05006	-0.10857	-0.05616	-0.12347	-0.10836	-0.11608	-0.11619	-0.20597	-0.17072	-0.05671	-0.05547	-0.07538	-0.06205	-0.07064	-0.01058	0.03071
work			0.16578	0.60558	0.39855	0.00625	0.00043	0.17192	0.15596	-0.09216	0.06401	0.06066	0.23931	-0.07671	-0.07740	0.09671	-0.26979	-0.10210	-0.21603	-0.17415
consider	0.70916	0.41082	0.26419	0.12758	0.21758	0.18619	0.03040	0.15826	0.22339	0.33942	0.14909	0.20651	0.35252	0.13440	0.19073	0.15125	0.17978	0.09036	0.10601	0.11475
walk			0.36304	0.11477		0.85834	0.87593	-0.14846	-0.41648	-0.15320	-0.17367	-0.17071	-0.43802	-0.38829	0.10360	-0.26692	-0.19745	0.01260	-0.06260	-0.08209
give	0.14035	0.38125	0.18568	-0.09377	-0.06071	0.21933	0.24314	0.15866	0.28477	0.20349	0.07211	0.21928	0.02448	0.05033	0.07002	0.05496	0.09304	0.08747	0.16562	0.18462
read				0.35841	0.33724	-0.06439	-0.08182	-0.23537	-0.32484	-0.04511	-0.02797	0.14542	0.17265	-0.19157	0.03243	0.12937	0.03484	0.01866	0.00040	0.02726
run	-1.26371	-1.02786	-0.92169	-0.78681	-0.63207	-0.63745	-0.44550	-0.23104	-0.41612	-0.28396	-0.18180	0.15667	-0.27298	-0.37916	-0.13670	-0.07287	-0.10090	0.00406	0.04849	-0.07763
watch					-0.14050		-0.34029		0.35012		-0.33943	0.69390	-0.42291	-0.12574	0.59988	-0.05350	-0.17915	0.07061	-0.21214	-0.00818
write		-0.02402	-0.25279	-0.15784	0.07083	0.01758	0.12881	-0.03421	0.08496	0.04894	-0.02904	0.08663	0.17307	0.08032	-0.12147	-0.03766	-0.03430	0.01264	0.02134	-0.06054
play		-1.02589	-0.78439	0.30388	0.13002	-0.43944	-0.17028	-0.16197	-0.21142	-0.31852	-0.30741	-0.18913	-0.10525	-0.08578	-0.13048	-0.21154	0.02795	-0.04203	-0.06571	-0.00059
take	0.54976	-0.10758	-0.17859	-0.16894	-0.08751	-0.08981	-0.20798	-0.06936	-0.02222	-0.00700	-0.10286	-0.03580	0.00190	-0.00510	0.01895	-0.00697	0.03253	0.03649	0.00603	0.11507
look	-0.29155	-0.09515	-0.45410	-0.45379	-0.39537	-0.33218	-0.43118	-0.37725	-0.41489	-0.24293	-0.37665	-0.24641	-0.35156	-0.27883	-0.29076	-0.22726	-0.28745	-0.20845	-0.16646	-0.15864
start				-0.32471	0.44155	0.20466	0.25418	0.02240	0.15264	0.75428	-0.01352	0.12815	0.04490	0.14544	0.25567	0.24382	0.17391	0.00001	0.05254	0.03066
bring	-0.62394	0.20685	-0.43814	-0.45147	-0.30297	0.12128	0.15393	0.12263	0.10784	0.19352	0.33684	0.12182	-0.00880	0.16167	-0.02110	0.01722	0.09905	0.27455	0.10970	0.09259
help					-0.03573	-0.28367	-0.03395	-0.05336	-0.25696	-0.29547	-0.29346	0.07705	-0.01543	-0.30216	-0.35543	-0.50092	-0.01300	-0.06777	-0.17480	-0.04578
try				-0.37440	-0.43141	-0.57447	-0.26504	-0.02125	-0.02260	-0.07067	-0.01188	-0.08060	0.00328	0.31036	0.04256	0.10610	-0.00094	0.14489	0.07662	0.06549
wait	-1.13976	-1.13760	-0.66165	-0.62719	-0.39825	-0.46232	-0.38374	-0.24850	0.04014	-0.03102	0.14341	-0.05918	0.00563	-0.01045	-0.16127	-0.12762	-0.06468	-0.00860	-0.09091	-0.17324
kill				0.07332	0.34973	-0.01725	-0.00912	0.68311	-0.06496	0.06774	0.26289	0.05485	0.20096	0.51790	0.11976	-0.19191	0.14355	0.21793	0.15126	0.28154
listen				-0.32492	0.11926	0.46004	-0.10610	-0.19959	-0.27855	-0.20263	-0.21881	-0.21101	-0.17424	-0.10272	-0.38131	-0.21498	-0.03644	0.11649	0.07412	0.07823
ask	-1.06050	-0.16746	0.02160	0.52556	0.00170	-0.04882	0.13808	0.02739	0.02023	0.00149	-0.11118	-0.01487	-0.26135	-0.10444	-0.26958	-0.02730	-0.00275	-0.03059	0.04720	0.02343
carry	-0.43689		0.00671	-0.09989	0.08418	0.20501	-0.15474	-0.11478	-0.05344	-0.00904	-0.01594	0.05296	0.13010	-0.09872	0.05957	-0.08165	0.02436	0.10746	0.02618	-0.08389
share	-0.83534	-1.32599		-0.62608	-0.40399	-0.37493	-0.12328	-0.01444	-0.18174	0.10681	0.06581	-0.08237	0.11891	-0.09137	-0.13048	0.03208	0.02704	0.04587	0.00607	0.00321
choose	0.35843	-0.11511	0.25381	0.06962	0.07390	0.21813	0.16762	-0.00807	-0.01549	-0.17594	-0.05731	0.06193	0.00564	0.23855	-0.01617	0.17738	0.02649	0.03473	-0.08749	-0.02303

Table 13. Summary table for the positive simple past construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink	-0.07141	0.14259	0.15391	0.19174	0.21057	0.21327	0.27287	0.22564	0.25084	0.24395	0.23193	0.22913	0.16678	0.14739	0.11951	0.07878	0.10931	0.14306	0.14651	0.14990
eat	-0.09661	0.04440	0.15994	0.07127	0.08374	0.10838	0.13683	0.11694	0.12838	0.11617	0.13759	0.08743	0.10342	0.07986	0.06057	0.00763	0.01550	0.02792	0.05001	0.04112
dance	-0.82124	-0.59781	-0.45860	-0.61578	-0.55925	-0.51427	-0.50546	-0.51742	-0.43283	-0.44877	-0.41872	-0.47147	-0.43869	-0.38867	-0.45773	-0.48207	-0.42956	-0.36110	-0.27540	-0.26168
sing	-0.66614	-0.51945	-0.48412	-0.55724	-0.47252	-0.55568	-0.40479	-0.39557	-0.33929	-0.28967	-0.24621	-0.25222	-0.25045	-0.28708	-0.31169	-0.32965	-0.32096	-0.23578	-0.19104	-0.15815
cook		-0.40400	-0.39893	-0.34301	-0.34063	-0.40828	-0.23962	-0.41356	-0.46201	-0.41311	-0.47061	-0.52277	-0.58233	-0.63563	-0.60216	-0.72350	-0.70803	-0.61928	-0.51487	-0.44217
talk	-0.14150	-0.09490	-0.02233	-0.02701	0.00508	0.03814	-0.01949	-0.00771	0.00483	0.01814	0.04371	0.07323	0.04226	0.05867	0.03579	0.01820	0.01001	0.05156	0.04351	0.07023
think	0.04437	0.01688	0.01635	0.02861	0.03343	0.02052	0.00793	0.02811	0.03500	0.05266	0.04236	0.02818	-0.02032	-0.01991	0.03355	0.06813	0.03808	-0.02480	-0.03907	-0.02653
complain	-0.22498	0.06078	-0.03102	-0.01184	0.08715	0.02408	0.00420	-0.01166	0.06943	0.00375	0.02643	0.01688	0.02545	0.00674	-0.00902	0.02122	-0.01583	0.02011	0.03987	0.06143
cry	-0.32799	-0.18212	-0.16970	-0.16845	-0.16817	-0.25545	-0.28295	-0.30465	-0.26839	-0.30677	-0.29725	-0.29400	-0.31726	-0.36389	-0.36649	-0.37786	-0.35053	-0.34400	-0.32390	-0.31277
throw	-0.24685	-0.20280	-0.11774	-0.10214	-0.08443	-0.07019	-0.03909	-0.03685	-0.02593	-0.03964	-0.02985	-0.00878	-0.00985	-0.03138	-0.04769	-0.06907	-0.02236	-0.00628	0.01530	0.01240
scream	-1.19370	-1.08058	-1.09164	-0.73006	-0.81811	-0.74310	-0.71931	-0.70149	-0.66358	-0.66717	-0.61206	-0.61570	-0.64721	-0.60823	-0.58481	-0.63418	-0.55623	-0.45288	-0.41124	-0.35496
shout	0.49458	0.45265	0.61095	0.76397	0.69096	0.70566	0.68572	0.65507	0.79086	0.83609	0.77707	0.66341	0.59230	0.53075	0.44898	0.33933	0.32200	0.31504	0.29072	0.28226
say	-0.04762	-0.00069	-0.02280	-0.00309	-0.02297	-0.05296	-0.04147	-0.10094	-0.08320	-0.08234	-0.05622	-0.03928	-0.04363	-0.05025	-0.08374	-0.08301	-0.05140	-0.02497	-0.00376	-0.00815
speak	-0.18290	-0.17575	-0.10606	-0.03453	-0.02248	-0.01025	0.00923	-0.01827	0.01001	0.00806	0.02306	0.04139	0.04842	0.06752	0.07388	0.08234	0.08632	0.11256	0.10686	0.11718
work	0.39598	0.16690	0.24527	0.06195	0.12194	0.22034	0.33752	0.38145	0.39652	0.42955	0.39823	0.33664	0.26129	0.25082	0.24406	0.22527	0.14139	0.16069	0.13570	0.10824
consider	0.25396	0.36794	0.30671	0.32627	0.40459	0.43071	0.40088	0.39645	0.33751	0.32913	0.32838	0.33484	0.33416	0.37110	0.39160	0.42819	0.39706	0.35011	0.29747	0.25191
walk	0.04782	0.02250	0.00866	0.02281	0.00603	0.05205	-0.01157	0.02135	0.05402	0.04870	0.03475	0.03686	0.01781	0.00449	0.01380	-0.03168	-0.01486	-0.01567	-0.00798	-0.00008
give	0.09750	0.16569	0.14960	0.17313	0.15492	0.14850	0.16502	0.15849	0.10143	0.10483	0.08873	0.07908	0.06862	0.07612	0.05726	0.05491	0.05584	0.07513	0.06056	0.04826
read	-0.02373	0.02051	0.05440	0.00711	0.03818	0.03557	0.06300	0.08402	0.08464	0.09569	0.07977	0.11019	0.10781	0.11867	0.12527	0.13446	0.10237	0.08143	0.05408	0.04096
run	-0.34034	-0.25947	-0.20835	-0.23113	-0.20447	-0.20265	-0.20043	-0.19894	-0.15725	-0.12444	-0.11747	-0.13860	-0.12934	-0.12442	-0.12588	-0.13414	-0.07092	-0.04527	-0.03133	-0.02633
watch	-0.25797	-0.32565	-0.26041	-0.29628	-0.29304	-0.21452	-0.22173	-0.19628	-0.14877	-0.16217	-0.11830	-0.11481	-0.09597	-0.07857	-0.05062	-0.09921	-0.08597	-0.09160	-0.07897	-0.05394
write	0.61733	0.54699	0.49464	0.41922	0.38952	0.48180	0.46619	0.46799	0.48320	0.48221	0.50759	0.47896	0.42771	0.45566	0.44734	0.43103	0.33058	0.32095	0.29252	0.31921
play	-0.21597	-0.25234	-0.22944	-0.15491	-0.12672	-0.12000	-0.04936	-0.07036	-0.03650	-0.01232	0.06093	0.05272	0.04978	0.07355	0.09394	0.07532	0.05768	0.20097	0.23400	0.26095
take	0.14795	0.18751	0.17175	0.16811	0.18995	0.17957	0.19070	0.19726	0.17903	0.20219	0.20051	0.17867	0.13876	0.12128	0.12334	0.10799	0.10742	0.10982	0.10031	0.08545
look	-0.26539	-0.28446	-0.25778	-0.23993	-0.25715	-0.23254	-0.25692	-0.25058	-0.19679	-0.20327	-0.19324	-0.17304	-0.16589	-0.15109	-0.15310	-0.18903	-0.15279	-0.11717	-0.08162	-0.06206
start	-0.50496	-0.40453	-0.29027	-0.13688	-0.13370	-0.09055	-0.03945	0.03329	0.06360	0.12308	0.16423	0.16589	0.10771	0.08449	0.08253	0.04074	0.01142	0.01961	0.00696	0.01592
bring	-0.06447	0.05072	0.10820	0.13596	0.12227	0.13664	0.15965	0.15714	0.15279	0.17858	0.15775	0.14555	0.13673	0.12452	0.08899	0.06680	0.08068	0.11798	0.12803	0.13158
help	-0.12875	0.29506	0.22651	0.05584	0.07931	0.11048	0.07248	0.13448	0.17705	0.23616	0.18299	0.16874	0.11169	0.12352	0.13399	0.10946	0.08486	0.13639	0.13366	0.14022
try	-0.09498	-0.08945	-0.15226	-0.13137	-0.14672	-0.12766	-0.12171	-0.05618	-0.02613	0.00844	0.00161	0.02548	0.03326	0.02594	0.05941	0.07529	0.03945	0.00253	-0.01911	-0.01649
wait	0.02404	-0.12654	0.00601	0.01086	-0.06601	-0.08688	-0.10438	-0.07225	-0.05794	-0.07031	-0.10299	-0.09692	-0.07894	-0.07553	-0.05480	-0.06513	-0.06235	-0.05113	-0.04096	-0.02220
kill	0.54242	0.33891	0.41064	0.64428	0.59115	0.56877	0.59822	0.62822	0.60194	0.57161	0.55588	0.49805	0.43200	0.38484	0.36386	0.34649	0.37891	0.40812	0.40592	0.38927
listen	-0.46588	-0.36878	-0.37396	-0.32629	-0.27159	-0.26387	-0.26880	-0.24117	-0.17485	-0.17872	-0.15149	-0.12665	-0.10121	-0.07844	-0.07846	-0.08476	-0.06186	-0.02967	-0.01647	-0.00058
ask	-0.07381	-0.05049	-0.03905	-0.04866	-0.06413	-0.08028	-0.10039	-0.12569	-0.11166	-0.13282	-0.10842	-0.08930	-0.06637	-0.06845	-0.07368	-0.07078	-0.05473	-0.03985	-0.03328	-0.03866
carry	0.10931	0.19265	0.24474	0.21793	0.20989	0.19101	0.22483	0.21143	0.14758	0.16686	0.14517	0.13872	0.11578	0.09563	0.09754	0.08641	0.08150	0.07969	0.08933	0.11327
share	-0.09312	-0.02186	-0.06535	0.00924	0.03545	0.09813	0.10685	0.10885	0.15294	0.10460	0.09174	0.07908	0.08032	0.08707	0.13308	0.13969	0.08330	0.08711	0.03820	0.03850
choose	0.23514	0.19293	0.24047	0.21964	0.22501	0.18386	0.18854	0.19712	0.18218	0.20464	0.19926	0.21601	0.21209	0.25056	0.25060	0.28046	0.20868	0.18896	0.14091	0.13871

Table 14. Summary table for the negative simple past construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink			0.28489		0.73719	0.33448	0.07393	0.28591	0.47557	0.58496	0.34365	0.51352	0.16267	0.20301	0.03685	0.08767	0.02371	0.05904	0.13594	0.19354
eat		-0.10404	-0.15477	-0.09227	0.05584	0.14085	-0.07359	0.10147	0.00228	0.08631	-0.06454	0.22806	-0.13102	-0.21595	-0.18017	-0.18261	-0.21910	-0.10055	-0.03907	-0.13591
dance		-0.80313	-0.66001	-0.45593	-0.31427	0.03329	-0.10762	-0.50152	-0.27277	-0.25269	-0.37317	-0.31037	-0.31867	-0.16401	-0.37544	-0.49011	-0.37909	-0.36938	-0.18196	-0.11926
sing	-0.98104	-0.92833	-0.58066	-0.76789	-0.81984	-0.63357	-0.47553	-0.58336	-0.27692	-0.25230	-0.12491	-0.37841	-0.16287	-0.22871	-0.48651	-0.29018	-0.40161	-0.18052	-0.26753	-0.16774
cook				-0.71907				-0.08724	-0.78578		-0.82786	-1.06791	-1.24561	-0.95171	-0.82170	-0.66605	-0.95114	-1.02650	-1.24756	-0.57833
talk	-0.68025	-0.32607	-0.42606	-0.64539	-0.30784	-0.09605	-0.08340	-0.15650	-0.18815	-0.15858	-0.18922	-0.18100	-0.03965	-0.05426	-0.16496	-0.23078	-0.13149	-0.11717	-0.05972	-0.00295
think	0.04064	0.05388	0.05478	0.02378	0.08079	0.08379	0.09556	0.07520	0.04336	0.12411	0.09073	0.08756	0.06992	0.11151	0.11829	0.14352	0.12153	0.06742	0.03389	0.04480
complain	-0.28193	0.37358	-0.11996	-0.35489	-0.27905	-0.37724	-0.29846	-0.45815	-0.25668	-0.35468	-0.24864	-0.35455	-0.26170	-0.46447	-0.34967	-0.28757	-0.36489	-0.29086	-0.29409	-0.27789
cry	-0.58380	-0.50254	-0.71507	-0.73373	-0.82326	-0.73267	-0.83110	-0.71759	-0.63736	-0.79815	-0.76603	-0.81229	-0.66342	-0.69832	-0.73574	-0.85860	-0.63195	-0.56847	-0.64385	-0.54365
throw	-0.50432		0.48261	0.16837	0.07703	0.43550	0.16671	0.10872	0.30785	0.04921	0.21658	0.29317	0.30509	0.06303	0.00836	-0.04236	0.02578	0.13596	0.13812	-0.08396
scream			-1.45607	-0.96229	-0.94448	-1.29491	-1.38950	-0.80757	-1.36807	-1.08005	-1.17492	-0.85041	-1.26350	-1.25309	-1.18657	-1.02731	-1.28325	-0.87268	-0.76430	-0.69809
shout					-0.64321					0.30171	0.74625	0.58350	0.69902	0.52231	0.15855	0.01203	0.01038	0.00774	0.07278	0.15683
say	-0.01080	-0.06960	0.03491	-0.12238	-0.03958	-0.16762	-0.14991	-0.08375	-0.06993	-0.06824	-0.08942	-0.03934	-0.08098	-0.03820	-0.07146	-0.00218	-0.00325	0.02603	0.07206	0.05606
speak	-0.44731	-0.23282	-0.31747	-0.34251	-0.37138	-0.34834	-0.30351	-0.39478	-0.34614	-0.26611	-0.27317	-0.26926	-0.27012	-0.27841	-0.26558	-0.21845	-0.21105	-0.10918	-0.10921	-0.06568
work		-0.36388	0.81802	0.27639	0.07933	0.41954	0.60367	0.36065	0.48343	0.33227	0.48447	0.38176	0.49072	0.24136	0.17553	0.07171	0.01386	-0.09800	-0.04046	-0.03656
consider	0.01432	0.01165	0.15024	0.06485	0.29455	0.46410	0.38671	0.34325	0.23038	0.25279	0.31914	0.27110	0.30965	0.29023	0.32659	0.31462	0.30675	0.21739	0.17713	0.24441
walk		-0.10398	-0.21273	-0.36717	0.03026	0.12229	-0.43689	-0.41553	-0.36127	-0.16687	-0.03213	-0.06143	-0.29786	0.00825	-0.12538	-0.22961	-0.16694	-0.15085	-0.16179	-0.06078
give	0.05563	-0.07416	0.25263	0.11767	0.11040	0.08592	0.06085	0.02280	-0.00003	0.05414	0.07521	0.05111	0.11442	0.06665	-0.00462	-0.00809	0.05867	0.03498	0.05099	0.05916
read	-0.80695	-0.24689	-0.18876	-0.25319	-0.03617	0.04938	-0.07693	-0.07510	0.08116	0.06526	0.06787	-0.03899	-0.09990	0.03595	-0.03176	0.05217	0.08572	0.08425	0.04170	-0.01055
run	-0.28193	-0.57348	0.17293	-0.29915	-0.01996	-0.19882	0.10169	-0.07217	-0.10460	-0.12400	-0.04611	-0.14066	-0.04764	-0.00655	-0.12786	-0.07021	-0.02820	-0.00042	0.05046	-0.06433
watch		-0.32584	-0.61861	-0.33892	-0.47704	-0.37617	0.12125	0.22807	-0.43515	0.03270	-0.18129	-0.15785	-0.17078	-0.26865	-0.27136	-0.37465	-0.41970	-0.08597	0.06219	-0.07475
write	0.59529	0.48636	0.62546	0.30211	0.34580	0.57815	0.28641	0.27276	0.23295	0.32347	0.25241	0.32908	0.25808	0.34832	0.14606	0.28299	0.18549	0.19271	0.16011	0.17650
play	-0.67998	-1.45733	-0.41010	-0.62955	-0.45272	-0.26805	0.10087	-0.09353	-0.07885	-0.10815	-0.03869	0.05231	0.04029	0.09110	-0.04737	-0.10117	-0.07305	0.12344	0.14336	0.15480
take	-0.00049	0.02649	-0.08360	0.10747	0.12894	0.12789	0.13280	0.12499	0.16419	0.06951	0.09639	0.10373	0.07907	0.10050	0.06934	0.10253	0.07982	0.09871	0.09715	0.05473
look	0.01934	-0.28204	-0.24778	-0.33314	-0.38458	-0.30302	-0.35511	-0.37246	-0.32891	-0.31948	-0.28292	-0.27878	-0.29227	-0.23853	-0.21409	-0.29684	-0.23405	-0.19249	-0.14627	-0.13126
start		-1.10418	-0.88208	-0.41822	-0.87765	-0.44199	-0.38749	-0.44125	-0.51333	-0.25055	-0.26815	-0.25173	-0.05114	-0.07214	-0.11841	0.05241	-0.10665	0.05119	0.04577	0.03554
bring	0.44372	0.05264	0.21815	-0.14820	0.15491	0.28521	0.17408	0.16047	0.15522	0.18989	0.11350	0.06971	0.14649	0.23784	0.05696	0.07661	0.08023	0.13309	0.08118	0.04221
help	-1.45885		0.17287	-0.33901		0.23677	-0.30596	-0.16646	0.14203	-0.09100	-0.14042	0.00514	-0.19061	-0.18424	-0.08319	-0.27621	-0.08309	-0.00725	-0.06989	0.07373
try	-0.16804	-0.54863	-0.70432	-0.56623	-0.23137	-0.29543	-0.24180	-0.15120	-0.08790	-0.06809	-0.12367	-0.06887	-0.04155	-0.08699	-0.03322	0.00182	0.01080	0.00361	0.00028	0.00480
wait	-0.55052	0.31844	0.01921	0.22371	0.06883	0.19123	-0.13065	0.07487	0.02069	-0.09707	-0.03385	-0.01082	-0.05301	0.02608	-0.11209	-0.06283	-0.03067	-0.08461	0.00367	0.05684
kill			0.48666	0.87247	-0.02925	-0.24543	1.00255	0.76113	0.52279	0.62404	0.35609	0.52967	0.41976	0.34850	0.41412	0.16263	0.35477	0.34695	0.17681	0.27698
listen			-0.42100	0.42726	-0.32306	-0.23800	-0.35586	-0.26791	-0.36200	-0.41436	-0.15636	-0.26011	-0.29443	-0.37569	-0.25650	-0.20229	-0.05523	0.01198	0.09923	-0.08161
ask	0.16571	-0.37383	-0.13303	-0.09485	-0.22745	-0.12402	-0.12217	-0.30358	-0.18044	-0.20783	-0.15587	-0.12991	-0.14394	-0.18898	-0.18277	-0.18722	-0.18505	-0.11402	-0.07678	-0.07657
carry	-0.13573		0.16781	0.50233	0.22683	0.50854	0.42569	0.13094	0.24575	0.36127	0.30118	0.26359	0.34038	0.18558	0.28918	0.22315	0.31318	0.36026	0.26269	0.19742
share	-0.50385	-0.32592	-0.30957	-0.52991	-0.03731	0.21692	-0.12769	0.04475	0.00201	0.04888	-0.00303	0.05259	0.10050	0.13122	0.21039	0.19800	0.10288	0.07776	0.07967	0.08469
choose	0.00459	-0.01042	-0.00289	0.04745	0.06860	0.08254	0.11115	0.05653	0.11584	0.08234	0.09598	0.15056	0.09515	0.24812	0.19298	0.10989	0.18182	0.11977	0.09372	0.04253

Table 15. Summary table for the positive simple future construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink		0.24842	0.21529	-0.06806	0.03668	0.23154	0.01923	0.22753	0.11566	0.14652	0.21510	0.15396	0.23045	0.11440	-0.12259	0.15254	0.03003	0.32287	0.00307	0.17463
eat	-0.29644	-0.29119	0.11840	0.13566	-0.11239	0.10216	0.12982	0.20389	0.17482	-0.01245	-0.00282	-0.12446	-0.10635	-0.05580	-0.13294	0.01120	-0.02266	-0.01158	0.02611	-0.03377
dance		-1.70644	-1.86435	-1.20957	-1.01069	-0.77855	-0.71228	-0.84563	-0.25388	-0.67271	-0.80111	-0.85845	-0.83762	-1.46507	-0.69955	-0.86883	-0.68721	-0.52243	-0.39590	-0.44902
sing	-1.06912	-1.15305	-1.13060	-1.08700	-0.93740	-0.74426	-0.82884	-0.88982	-0.58513	-0.53810	-0.50797	-0.43314	-0.54984	-0.59157	-0.58432	-0.52247	-0.45248	-0.29064	-0.24756	-0.20005
cook		-0.62679		-0.07901	-0.60024	-0.38044	-0.11550	-0.06718	-0.84564	-0.53452	-1.15972	-1.04536	-0.82035	-1.01103	-0.70622	-1.11748	-0.86891	-0.39870	-0.61416	-0.69242
talk	-0.21293	-0.58157	-0.49000	-0.03904	0.08531	0.07197	-0.07783	0.07774	-0.07546	-0.16000	-0.03392	0.16347	-0.09367	0.07545	0.02944	-0.03783	-0.09391	-0.00846	-0.07342	-0.04243
think	-0.14084	-0.02859	-0.33783	-0.14911	-0.24052	-0.16044	-0.29291	-0.23185	-0.15210	-0.16373	-0.10850	-0.09929	-0.06539	-0.03632	-0.06436	-0.09160	-0.01873	-0.02845	0.01015	-0.06470
complain	-0.49280	-0.57584	-0.71815	-0.32900	-0.34063	-0.20436	0.16536	0.00999	-0.22582	0.20835	0.29151	0.31729	-0.21305	0.20892	-0.34974	-0.16746	-0.06682	-0.11064	-0.17396	0.14615
cry	-0.63925	-0.36886	-0.61778	-0.46375	-0.74281	-0.66178	-0.35266	-0.73210	-0.32297	-0.50424	-0.54491	-0.23124	-0.35904	-0.13728	-0.38439	-0.23011	-0.28389	-0.23943	-0.12066	-0.28949
throw	-0.03665	-0.11496	-0.06244	-0.35246	-0.10600	-0.06174	-0.10174	-0.16170	-0.04943	-0.20230	-0.13009	-0.13707	-0.16491	-0.03551	-0.13768	-0.14854	-0.20141	0.09815	0.11347	0.03667
scream						-0.68145		-0.30129	-0.53660	-0.78201	-0.73491	-0.93810	-0.56845	-0.35906	-0.67301	-0.49616	-0.38748	-0.47222	-0.43279	-0.30890
shout													0.65697	-0.02362	0.26670	-0.01006	0.11830	0.44886	0.23233	0.26285
say	-0.16760	-0.14898	0.02627	0.04427	0.07049	0.03087	0.07815	0.02572	-0.01402	0.01049	0.07629	0.09450	0.02616	-0.00218	-0.04791	0.02074	-0.01875	0.07306	0.12169	0.07088
speak	-0.24200	-0.34881	-0.14073	-0.01211	-0.13958	-0.14374	-0.09294	-0.14814	-0.15939	-0.13778	-0.08158	-0.02892	-0.07768	-0.06331	0.05081	0.11728	0.13965	0.24180	0.26264	0.22506
work		0.09358	0.40363	0.76468	0.23972	0.31092	-0.00132	0.30847	0.24530	0.28346	0.13344	0.26198	0.07530	0.09082	0.13411	0.12169	0.01357	-0.06848	-0.06124	-0.06462
consider	0.45577	0.12603	-0.13161	0.05596	-0.13654	0.02404	0.06527	0.25983	-0.11135	0.05545	0.12144	0.14920	0.25453	0.09017	0.22124	0.12638	0.16773	0.10029	-0.04049	-0.00996
walk	0.02410	-0.09322	-0.46137	-0.19320	0.17813	-0.15015	0.02480	-0.02009	-0.19306	0.05205	-0.15082	-0.14319	-0.10658	-0.23256	-0.22691	-0.17210	-0.24971	-0.10767	-0.11996	-0.05847
give	0.35243	0.40678	0.26124	0.26596	0.17409	0.21384	0.21154	0.16765	0.09843	0.01404	0.01341	-0.01148	-0.03571	0.03256	-0.02266	-0.07241	-0.03370	0.11143	0.15938	0.18420
read	0.37041	0.03846	0.02066	-0.13317	-0.00528	0.04311	-0.08743	0.01902	-0.05427	0.16732	0.10498	-0.02207	0.00526	0.06601	0.05514	0.05568	0.05157	-0.13306	-0.15055	-0.15109
run	-0.52090	0.05002	-0.21591	-0.00329	-0.10549	-0.03659	0.01107	-0.20615	-0.14510	-0.11776	-0.00923	0.00354	-0.11957	-0.03121	-0.00579	-0.14265	-0.04173	-0.10972	0.03660	-0.03810
watch	0.19393		-0.16516	-0.43344	-0.11877	-0.19970	-0.12000	0.24336	-0.16550	-0.12948	-0.00650	0.08253	0.18763	0.08394	0.09593	-0.15935	-0.10957	-0.12289	-0.09481	-0.05055
write	-0.40427	-0.14482	-0.21139	-0.02574	0.01399	-0.06782	-0.03962	0.04074	-0.01981	-0.02603	0.00436	-0.07371	0.00977	-0.00748	-0.10324	0.02519	-0.06103	0.00451	-0.00387	0.01336
play	-1.38099	-0.76021	-0.55537	-0.67119	-0.54404	-0.31162	-0.43608	-0.39011	-0.35082	-0.23819	-0.17602	-0.18827	-0.14211	-0.16923	-0.15801	-0.12148	-0.11729	-0.06958	-0.06889	-0.02900
take	-0.01780	0.05059	0.00486	0.05058	0.06724	0.00457	-0.00596	0.01873	0.02998	0.08458	0.02398	0.04362	0.05719	-0.01997	-0.00495	-0.03168	0.02537	0.09198	0.09877	0.12662
look	0.59848	0.16105	-0.16103	-0.11834	-0.08009	-0.08148	-0.22344	-0.15049	-0.14735	-0.11953	-0.11167	-0.16892	-0.12977	-0.05881	-0.07497	-0.06579	-0.09640	-0.10685	-0.14017	-0.15528
start	-0.66908	-0.02466	-0.16510	0.40621	0.01144	-0.20444	-0.29415	0.05187	-0.35135	-0.11415	-0.13843	-0.07125	-0.22287	-0.12190	-0.06984	-0.11907	0.04276	-0.02567	-0.16568	-0.12277
bring	0.46097	0.12141	0.21430	0.15611	0.08677	0.06521	0.03523	-0.04721	-0.02248	-0.05311	-0.01405	-0.10377	-0.18633	-0.13930	-0.16152	-0.05317	0.03378	0.15661	0.21091	0.25205
help			0.04551	0.11501	0.18403	0.21237	0.19455	-0.05797	-0.06325	-0.12421	-0.16235	-0.19464	-0.27335	-0.19424	-0.28738	-0.18882	-0.18541	-0.01738	-0.03400	-0.00928
try	0.74236	-0.30520	-0.04861	0.14513	0.01101	-0.22727	-0.13188	-0.07866	-0.00847	-0.05448	-0.05799	0.02282	0.07396	0.07711	0.08742	0.00025	0.07590	0.10543	0.02169	0.04629
wait	-0.31019	0.66365	0.13128	-0.07910	-0.17930	-0.11174	-0.38907	-0.24051	-0.17376	-0.07562	-0.16126	-0.18184	-0.03225	-0.42211	-0.32751	-0.32991	-0.23643	-0.12269	-0.07605	-0.04885
kill	0.10961		0.84826	0.15221	0.08053	0.13869	0.05936	0.14220	-0.00908	0.24569	0.16691	0.13916	0.01701	-0.01446	-0.13508	-0.02915	0.00768	0.24206	0.36029	0.38947
listen	-0.56030	-0.58347	-0.24702	-0.26746	-0.10264	-0.04644	-0.17040	-0.15726	-0.24401	-0.17510	-0.05620	-0.15309	-0.16648	-0.05930	-0.18078	-0.17853	-0.03655	0.04794	-0.01142	0.01508
ask	-0.73665	0.03136	-0.37269	-0.30712	-0.14160	-0.06478	-0.06226	-0.10739	0.11537	0.01376	0.09888	-0.05656	-0.09744	0.01630	-0.06094	-0.07707	-0.02314	0.01319	-0.08859	-0.10830
carry	-0.06682	-0.19148	-0.15162	-0.08913	0.00540	0.17166	0.05593	0.06703	0.12922	0.01809	0.07344	0.07078	0.08960	-0.04970	0.11988	-0.09065	-0.01638	0.00895	0.08231	0.08462
share	-0.73589	-0.62693	-0.80224	-0.59659	-0.32412	-0.45323	-0.44018	-0.30140	-0.25702	-0.32874	-0.18444	-0.11417	-0.10973	0.07504	-0.00720	-0.16171	-0.20231	-0.11706	-0.16329	-0.18856
choose		0.34571	-0.10173	-0.07911	0.06514	0.08985	0.22356	0.04903	0.08321	0.05388	-0.14443	0.04085	0.06002	0.05204	0.13099	0.18241	0.24085	0.01033	-0.08409	-0.11646

Table 16. Summary table for the negative simple future construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink					0.41050	-0.32980	0.52575	0.21920	-0.05169	0.90614	0.24075	0.59280	0.00614	0.37073	0.04302	0.20269	0.19710	0.34756	0.20764	0.56303
eat		-0.18016	-0.27636	-0.29808	-0.45830	-0.03778	0.08422	-0.01287	-0.11487	-0.24022	-0.05869	-0.01898	0.13516	-0.22643	-0.29372	-0.15266	-0.19211	0.06231	-0.05417	-0.15664
dance						-0.95284		-1.16261	-0.89753	-1.73023	-1.17262	-0.98305	-0.64043	-1.31153	-0.24588	-0.56886	-0.70658	-0.85103	-0.40312	-0.53988
sing				-0.86160	-0.86340	-0.67219	-0.63637	-0.24062	-0.47716	-0.85703	-0.37214	-0.03120	-0.39784	0.12061	-0.82852	-0.26784	-0.73968	-0.41276	-0.36177	-0.43628
cook																				
talk				-0.56054	-0.36817	-0.54750	-0.40632	-0.22369	-0.18299	-0.17120	0.00638	0.11602	0.54269	-0.21477	-0.14279	-0.26389	-0.31099	-0.16557	-0.16831	-0.16065
think	-0.02918	-0.14324	0.02390	-0.12198	-0.13591	-0.31101	-0.11537	-0.32727	-0.03021	-0.19745	-0.07130	-0.10177	-0.28407	-0.03914	-0.14513	-0.24132	-0.06906	0.03661	0.05475	-0.06004
complain			-0.33673	0.25269	-0.26570	-0.51014	-0.47966	-0.58808	-0.52163	-0.43023	0.01642	-0.47080	-0.57353	-0.60695	-0.24597	-0.51119	-0.38505	-0.39518	-0.27741	-0.13170
cry							0.61949	0.85343	0.37455	0.19811	0.12137	0.05243		-0.11363	0.14643	0.28336	-0.23637	-0.22082	-0.16144	-0.03342
throw						-0.19495	0.72606	0.36191	-0.41265	-0.11901	-0.07631	-0.27886	0.77523		0.33792	0.14739	0.26121	0.19244	0.46797	0.65092
scream																				
shout																				
say		0.42457	0.02908	0.03009	0.22895	-0.16129	-0.26770	-0.07142	-0.00563	-0.10181	-0.07779	0.02890	-0.03832	-0.03778	-0.01157	-0.12513	-0.07265	-0.03861	0.05211	0.08640
speak	-0.66930	-0.18703	-0.36231	-0.58612	-0.46610	-0.43906	-0.41683	-0.25211	-0.30584	-0.33133	-0.41192	-0.29607	-0.37037	-0.35359	-0.23561	-0.38462	-0.25778	-0.03655	0.04572	0.16812
work					0.82084		1.18494	0.73398	1.61586	1.38892	0.44381	0.39344	0.00145	1.13328	0.16787	-0.14933	0.08527	0.08158	-0.08243	0.00308
consider		-0.04514	0.11891	0.07655	-0.19228	-0.09795		0.13989	0.11126	0.31544	-0.29730	0.18062	-0.12028	0.05216	-0.13225	-0.01074	0.27826	-0.01263	-0.01069	0.14186
walk					0.13717	-0.19490	0.42493	0.11509	-0.07913	0.08732	0.43888	-0.04457	0.17286	0.51866	0.07873	0.06814	-0.34292	0.33147	0.10656	0.16961
give	0.20352	0.18097	0.05886	0.21736	0.11050	-0.02294	0.07172	0.04940	-0.00841	-0.00363	-0.19216	-0.14025	-0.03523	-0.17871	-0.13551	-0.23526	-0.03386	-0.05578	-0.00339	0.06644
read		-0.92066	-0.01408		0.16367		-0.21052	0.15186	-0.35311	0.17800	-0.01730	-0.07463	-0.16325	0.79037	0.24111	0.09475	0.01161	-0.06592	-0.30385	-0.28035
run				-0.25983	0.01221	0.04419	-0.14841	-0.07797	-0.31574	0.15257	0.20495	-0.19840	0.05905	0.28762	0.13478	0.18822	-0.07991	0.04159	0.07631	0.10882
watch																				
write	-0.09613		0.22021	0.07655	0.82578	0.52401		0.40335	0.74268	-0.45098	-0.01729	0.10550	0.25870	-0.14296	0.10511	0.00487	-0.04519	0.12579	0.01590	0.02898
play		-0.07508	-0.80279		-0.11281	0.28248	0.26706	-0.29995	0.04781	-0.25666	-0.13766	-0.04461	0.00869	-0.10971	0.17359	-0.00445	-0.31257	0.06562	0.03365	-0.05750
take	0.00836	0.20615	-0.18321	-0.02203	0.20974	-0.14976	-0.00925	-0.09763	0.03666	0.03362	0.14021	-0.02867	0.03612	-0.05066	0.04326	0.01628	-0.09636	0.05029	-0.04254	0.06334
look	-0.39768	-0.25195	-0.49287	-0.38590	-0.01268	-0.08983	-0.29236	-0.22996	-0.15867	-0.22307	-0.31298	-0.03496	-0.25751	-0.29083	-0.15932	-0.17513	-0.24533	-0.28769	-0.19424	-0.19222
start									0.00636	-0.44991	0.59006	0.71538	0.40138	0.50410	-0.18800	0.17609	0.26121	-0.08230	0.33755	0.35975
bring		-0.37678	-0.29570	-0.12930	-0.11298	-0.18096	0.14223	-0.15319	0.23718	0.30101	0.00143	-0.05305	0.37745	0.07717	0.00197	-0.30177	0.04292	0.04818	0.13937	-0.06263
help		-0.74508		-0.25974	0.59647	0.10638	-0.03417	-0.26733	0.05439	-0.14907	0.22234	0.14757	-0.18622	-0.48913	-0.24640	-0.24458	-0.32789	-0.09494	-0.16435	-0.08756
try		-0.44338	0.28723	0.41767	0.06116	-0.38447	0.34680	0.69154	0.00162	0.23915	0.11510	0.06883	0.17348	-0.14943	0.28050	-0.01542	0.00183	0.09204	0.03844	0.04834
wait		-0.50741	0.34530	0.15308	0.50172		-0.42145	0.01486	0.03642	0.48880	-0.13117	-0.24581	-0.22833	-0.06567	-0.26746	-0.30091	-0.17860	-0.01966	-0.00253	0.00139
kill				-0.20853				0.17461	0.97920	0.44471	0.43930	0.46803	0.44588	0.15550	0.27744	0.63605	0.41548	0.53910	0.41276	0.27049
listen		-0.49547	-0.15137	-0.34024	-0.30956	-0.31179	-0.22538	-0.30637	-0.22544	-0.41274	-0.24686	-0.31814	-0.35744	-0.23357	-0.21500	-0.18905	-0.03429	-0.06744	-0.01852	0.10484
ask		-0.48254	-0.17075	0.07667	0.25580	0.15973	-0.03416	0.28947	-0.18302	-0.03133	-0.10530	-0.29108	-0.22821	-0.24204	-0.28926	-0.18217	-0.11060	0.03871	0.00422	-0.09858
carry	-1.00129	-1.04745	0.04399	0.12406	0.14636	0.62310	0.12749	0.30237	0.47802	0.35687	0.16301	-0.06128	-0.31635	0.68542	0.19107	-0.09983	0.02335	0.67627	-0.00616	0.48182
share				-0.12116	-0.06703	0.69036	-0.15904		0.02987		-0.66341	-0.07259	-0.02036	0.32028	-0.36684	-0.15997	-0.24520	-0.14332	-0.27043	0.03253
choose				-0.16258	0.41031	-0.01868	0.66529	0.13986	0.58068	-0.19039		-0.23173	0.01309		-0.02399		0.27922	-0.03112	-0.10636	-0.16532

Table 17. Summary table for the positive present progressive construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink			0.21324		0.45273	0.32813	0.28746	0.18232	0.15374	0.12000	-0.09244	0.18071	0.03212	0.22061	-0.08893	-0.07336	-0.05200	0.03558	0.09135	0.06275
eat	-0.46287	0.01286	0.20955	0.30788	0.07629	0.36748	0.40946	-0.01076	0.35940	0.01862	-0.04216	0.08037	0.19728	-0.00478	-0.13051	-0.09196	-0.07788	-0.13482	-0.05436	-0.14385
dance		-1.14977	-1.26160	-0.81845	-0.70900	-0.58914	-0.59453	-0.41362	-0.66622	-0.63941	-0.69233	-0.57876	-0.77064	-0.62195	-0.76591	-0.74119	-0.63457	-0.49444	-0.37715	-0.35032
sing		-0.84920	-0.85903	-0.60462	-0.64881	-0.48474	-0.59041	-0.63729	-0.51579	-0.48526	-0.43578	-0.35520	-0.43129	-0.44641	-0.54671	-0.42638	-0.45568	-0.27937	-0.19131	-0.14431
cook			-1.35624		-1.72877	-0.82431	-0.62839	-1.01508	-0.81818	-0.74938	-0.71182	-0.65253	-0.61951	-1.00105	-1.18725	-1.20124	-0.97150	-0.76276	-0.53368	-0.50516
talk	-0.23032	0.01438	0.07147	0.09179	0.06540	0.02050	0.06430	0.12706	0.09700	0.12160	0.18767	0.15497	0.24074	0.16188	0.17819	0.14125	0.09810	0.09827	0.10500	0.07718
think		0.11939	-0.26814	-0.02970	-0.23024	-0.08514	-0.09282	-0.21440	-0.12855	-0.10220	0.04680	-0.03551	0.07581	0.06345	0.01449	0.02908	-0.00292	-0.02030	-0.02893	-0.04846
complain			-0.41160		0.21657	0.43962	-0.40634	-0.22101	-0.32710	-0.33281	-0.53523	0.49522	0.09009	0.20557	-0.56129	-0.39997	-0.30499	-0.14219	-0.15681	-0.19187
cry		-0.99221	-0.41547	-0.79654	-0.78862	-0.60385	-0.79622	-0.76730	-0.56848	-0.74373	-0.87979	-0.82234	-0.79159	-0.79113	-0.73030	-0.69657	-0.64434	-0.55662	-0.52668	-0.47708
throw		-0.14877	-0.31449		0.20759	0.01355	0.03175	0.05102	-0.18652	-0.02402	0.09166	0.22954	0.24343	0.36001	-0.17317	0.05879	0.02478	0.03846	0.11558	0.02685
scream						-1.00013			-1.04718	0.17518	-1.20712	-0.46399	-0.00011	-0.45584	-0.48919	-0.46922	-0.39674	-0.34725	-0.45152	-0.31266
shout						-0.42589	-0.88336		0.23167	0.95343		1.16668	0.06871	-0.04489	0.32514	-0.26362	-0.25570	0.02829	0.15130	0.20372
say	0.29733	0.27693	0.55119	0.30880	0.18560	0.12832	0.06211	-0.00560	0.10876	-0.12201	0.01725	-0.06000	-0.04575	0.12625	0.08003	0.08042	0.06751	0.11234	0.09295	0.08781
speak	0.55486	1.00822	0.75455	1.08383	0.86254	0.58345	0.58496	0.70513	0.56287	0.43666	0.37709	0.30440	0.31482	0.41084	0.28139	0.25401	0.27140	0.29059	0.23172	0.27849
work	-0.36463	-0.39196	-0.13006	-0.03912	0.24330	0.04610	0.07665	0.28039	0.15056	0.20479	0.17063	0.11897	-0.01933	0.02067	0.11465	0.07869	-0.02249	-0.07040	-0.10504	-0.09919
consider			0.42334		0.90495	0.34043		0.27005	1.21595	0.33652	0.55347	0.32324	0.32960	0.39992	0.33421	0.35618	0.22428	-0.03101	-0.06574	-0.14592
walk		-0.26033	0.06160	-0.24740	-0.09591	-0.15707	-0.15820	-0.07313	-0.20184	-0.22465	-0.08070	-0.14480	-0.09278	-0.12847	-0.08913	-0.17345	-0.09900	-0.04412	-0.05006	0.00617
give	-0.10458	0.54940	0.32709	-0.14357	-0.03306	-0.11317	-0.07215	-0.06413	-0.02776	-0.02542	0.00570	-0.06022	-0.03309	-0.08176	-0.17643	-0.16167	-0.19279	-0.04458	0.04076	0.01766
read	0.76070	0.69136	0.39659	0.36335	0.15391	0.20315	0.28390	0.17771	0.20928	0.12573	0.07480	0.15899	0.09085	0.11738	0.23123	0.18188	0.13547	0.00778	-0.04732	-0.11268
run	-0.73694	0.69906	-0.11280	-0.29780	-0.20548	-0.06298	-0.12219	-0.21135	-0.05973	-0.02285	-0.05971	-0.00809	-0.04431	-0.02266	-0.01463	0.02001	0.06791	0.02089	0.04843	-0.01708
watch	-0.09083	-0.00246	-0.64535	-0.15133	-0.23763	-0.13313	-0.18226	-0.20483	-0.09955	0.09841	-0.09986	-0.22227	-0.25529	-0.17836	-0.30306	-0.25785	-0.12338	-0.02092	-0.05304	-0.05482
write		0.44134	0.30286	0.58901	0.10986	0.26308	0.32302	0.31521	0.36194	0.33594	0.33108	0.33554	0.32951	0.27882	0.22767	0.35819	0.20164	0.17307	0.10413	0.11955
play	-0.57350	-0.29441	-0.27141	-0.00807	-0.28246	-0.11581	-0.27385	-0.10064	-0.16778	-0.19648	-0.07872	-0.12787	-0.08676	-0.11373	-0.12577	-0.04939	0.00851	0.01829	0.09593	0.11435
take	-0.18982	0.13589	-0.28080	0.06649	0.09917	-0.14302	-0.10273	0.11766	-0.02850	-0.06975	-0.15183	-0.13575	-0.09551	-0.05381	-0.07273	-0.09189	-0.07283	-0.10444	-0.10947	-0.12283
look	-0.03307	-0.57990	-0.40608	-0.17731	-0.02442	-0.21037	-0.23380	-0.26218	-0.23370	-0.14311	-0.13292	-0.19697	-0.20054	-0.15437	-0.09340	-0.12505	-0.08481	-0.05320	-0.07190	-0.04239
start				-1.25674	0.01236	-0.12484	-0.22072	-0.17401	-0.17220	0.24234	0.01483	0.14282	-0.27052	0.07873	-0.15550	-0.13010	-0.17133	-0.20443	-0.18849	-0.14743
bring	0.14401	0.37165	0.20178	-0.01493	-0.02919	-0.06022	-0.08939	0.08644	0.08176	0.03030	-0.02976	-0.07992	0.07778	-0.16744	-0.23301	-0.11257	-0.01046	0.07182	0.11937	0.10653
help				-0.42739	-0.00993	-0.16640	0.49017	-0.11827	-0.09584	-0.01947	-0.10492	-0.12346	-0.10347	-0.19658	-0.23934	-0.18795	-0.22108	-0.16704	-0.15356	-0.08599
try	0.01862	-0.71339	-0.02528	-0.07288	0.08032	0.17359	0.03510	-0.00107	0.06300	0.09039	0.09893	0.00994	0.10106	0.06092	0.13086	0.09532	0.05070	0.05650	0.00776	-0.00375
wait	-0.31583	0.24644	-0.02298	-0.16113	-0.20113	-0.01091	-0.13154	-0.04509	-0.20754	-0.26616	-0.23148	-0.43852	-0.24507	-0.26471	-0.41955	-0.45300	-0.28034	-0.15521	-0.14403	-0.03556
kill				-0.00106	-0.18158	0.01951	-0.19128	0.17098	-0.12490	-0.05903	0.27886	-0.01394	-0.16453	0.02940	0.22778	-0.18164	0.08462	0.13481	0.38045	0.37110
listen		-0.03474	-0.25671	-0.27684	-0.32519	-0.14814	-0.31701	-0.09132	-0.11266	-0.33538	-0.19894	-0.24122	-0.00431	-0.14765	-0.10071	-0.12720	-0.00417	-0.12519	-0.13159	-0.09112
ask			-0.45303	0.10551	-0.17031	0.05513	-0.27589	-0.12364	0.00885	0.02055	0.05576	0.01499	0.02524	0.14329	-0.02187	0.01097	0.02154	0.04394	-0.00083	0.01102
carry	-0.28757	-0.49796	-0.43423	-0.01756	-0.14715	0.15666	-0.07908	0.04629	0.00737	-0.14349	-0.24875	-0.16861	-0.32394	-0.31983	-0.45415	-0.42868	-0.39595	-0.33975	-0.28928	-0.22103
share				-1.08051			0.37207		0.23167	-0.44338	0.02341	0.12716	0.06207	-0.00539	0.24198	-0.06571	-0.19647	-0.11872	-0.12804	-0.23074
choose					-0.08458	-0.47709	-0.40622	0.21282	-0.00180	-0.22046	-0.20719	-0.20252	-0.07503	0.36530	0.02722	-0.06498	0.00387	-0.04090	-0.21439	-0.26089

Table 18. Summary table for the negative present progressive construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink													-0.30461			-0.20620	-0.16615	0.17827	0.21621	-0.17777
eat									0.22899	-0.15430	-0.26613	0.32122	-0.43073	-0.20012	0.08347	0.02958	-0.34214	-0.32176	-0.46911	-0.35982
dance								-1.08190	-0.15622	-1.52079	-0.96424		-1.26188	0.10162	-1.15356	-1.35893	-0.68191	-1.03884	-0.39639	-0.61710
sing						-0.64384		-0.41040	-0.65346	-0.58340	-0.26613	0.40638	-0.70461	-0.38818	-0.59444	-0.51091	-0.19397	-0.47428	-0.62782	-0.14763
cook																				
talk				-0.34954	-0.31627	-0.34310	-0.54611	-0.39622	-0.28454	-0.16544	0.15203	0.26274	0.47849	0.17409	0.04143	0.32633	0.14828	0.12500	0.17594	0.24637
think			-0.69679	0.47614	-0.14252	-0.09634	-0.00632	-0.51123	-0.33053	0.01005	-0.08264	0.08581	-0.01038	-0.17872	-0.07309	-0.01050	0.00816	-0.00738	0.10439	0.11247
complain											-0.69233	0.21411					-0.55796	-0.38984	-0.27179	-0.35736
cry				-1.36033			-0.06958	-0.60387		-1.56654	-1.26087	-1.50958	-1.11511	-1.57256	-1.07006	-0.85921	-0.57831	-0.45527	-0.63273	-0.70353
throw																				
scream																				
shout																				
say									0.18112		-0.37944	0.75292	0.65195	0.55786	1.14891	0.50313	0.43462	0.34336	0.35634	0.40076
speak				0.48965	1.09513	0.32237	0.95370	0.95104	0.21040	0.04847	0.19489	0.21935	0.29667	0.28291	0.32622	0.27159	0.14469	0.46857	0.34231	0.27817
work					0.28313	0.33467	0.18860	0.34124	-0.30496	0.19828	0.12264	-0.15288	0.16164	-0.10663	0.12132	-0.13631	-0.15506	-0.12461	-0.23944	-0.04565
consider																				
walk					-0.71719	-0.60030	-0.29341		-0.28211	-0.58340	-0.30637	0.10481	-0.23787	-0.32531	-0.04945	-0.24044	0.11704	0.46545	0.06769	0.30085
give									0.69571	0.47583	0.27871	-0.22305	-0.02384	0.11135	0.05495	0.09458	-0.08768	-0.05671	-0.18757	-0.02659
read				0.03564	-0.21071	-0.16383	-0.16843	0.06990	-0.76928	-0.11007	-0.15649	-0.49088	-0.61156	-0.43386	-0.28386	-0.18785	0.12606	-0.35808	-0.03848	-0.27656
run			-1.70650	-1.16909	-1.14761		-0.77614		-0.03120	-0.91843	-0.37944	-0.26521	-0.00330	-0.06166	0.23145	-0.22951	-0.00600	0.31251	-0.35094	0.07407
watch						0.08873	-1.25206	-0.55698	-0.15647	-0.05556		-0.18483	-0.28609	-0.36346	0.03512	0.17483	-0.05669	-0.02826	-0.28498	-0.13203
write					0.08922	-0.14879	0.49706	-0.19655	0.01534	0.26965	0.59094	0.15228	0.20036	0.42097	0.21551	0.51172	0.20065	0.22757	0.18443	-0.03064
play						0.39363	-0.47189	0.00227	-0.31089	-0.17449	0.08863	-0.25278	0.01266	0.08150	-0.12781	0.10427	0.07835	-0.05000	0.05358	-0.01141
take				-0.27336	-0.97103	-0.60030	-0.42193	-0.30217	-0.70733	-0.11007	-0.07116	-0.19600	0.00394	-0.24807	-0.19833	-0.15588	-0.33432	-0.30936	-0.26156	-0.31107
look				-0.66065	-0.19728	-0.72930	-0.42229	-0.63417	-0.42878	-0.02819	-0.26865	-0.18921	-0.15194	-0.06621	-0.42749	-0.33156	-0.24988	-0.10705	-0.17951	-0.18784
start																				
bring																				
help							-0.46873				-0.18405	0.10493	-0.00331	-0.46089	-0.53770	-0.44096	-0.41406	-0.30823	-0.56100	0.01460
try					-0.05615		0.13575	-0.72821	0.72266	0.27626	0.24436	0.19020	0.12201	0.40523	0.20200	0.14337	0.09048	0.13765	0.15535	0.01362
wait						-0.51622			-0.10518	0.51713		-0.56554	-0.78339	-1.08957	-0.64094	-0.87308	-0.50951	-0.50642	0.06631	-0.09527
kill																				
listen					-0.61979	-1.12192	-0.98961	-0.67268	-0.72140	-0.93684	0.25426	-0.37142	-0.39641	-0.40621	-0.52833	-0.42022	-0.06443	-0.22380	-0.16820	-0.11106
ask								-0.10779			0.60736	-0.48434	0.68318	0.32183	-0.42450	0.15850	0.12419	0.01334	-0.09659	-0.02870
carry					-0.53525		-0.77017	-0.55698	-0.05938		-0.42726	0.49182		-0.06152	0.04355	-0.40121	-0.09457	0.03407	0.07064	0.13986
share																				
choose																				

Table 19. Summary table for the positive past progressive construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink			0.42334	0.88568	0.43081	0.25813	0.49974	0.31969	0.33641	0.24199	0.23776	0.32684	0.23095	0.16044	0.18778	0.16096	0.14379	0.16787	0.17751	0.15869
eat	0.11779	-0.15014	0.01187	-0.02995	0.21019	0.27387	0.06263	0.23207	0.12769	0.13475	0.17621	0.05862	0.01715	0.05449	-0.01433	-0.00268	-0.04076	-0.03530	0.01297	-0.01850
dance		-0.96307	-0.69698	-0.58710	-0.46843	-0.64865	-0.55120	-0.74373	-0.51065	-0.69527	-0.48886	-0.54919	-0.58100	-0.49200	-0.67364	-0.57415	-0.50684	-0.40753	-0.39219	-0.37141
sing	-0.62467	-0.47681	-0.75198	-0.64274	-0.55971	-0.61092	-0.50867	-0.65975	-0.57986	-0.47338	-0.38395	-0.42267	-0.30885	-0.26417	-0.35922	-0.36935	-0.30338	-0.22113	-0.20251	-0.16499
cook		0.03797	0.19253	-0.73493	-0.81816	-0.41183	-0.63494	-0.71241	-0.36737	-0.39496	-0.48903	-0.37354	-0.71699	-0.65383	-0.59910	-0.73748	-0.73592	-0.67869	-0.54924	-0.44569
talk	-0.23376	-0.19406	-0.03069	-0.08622	-0.06561	-0.02437	-0.04214	-0.03178	-0.03287	0.00727	0.05953	0.03686	0.03547	0.07027	0.04948	0.01862	0.04400	0.06080	0.04425	0.04008
think	-0.46961	-0.30375	-0.22736	-0.27347	-0.29157	-0.26481	-0.24425	-0.27340	-0.14987	-0.14397	-0.13386	-0.08089	-0.08152	-0.03880	-0.00572	-0.02071	-0.01381	-0.02121	-0.02534	-0.04210
complain	-0.70336	-0.36007	0.19256	-0.13260	-0.48135	-0.27450	0.00301	-0.11404	-0.27015	-0.06745	0.03716	0.03722	0.06809	0.19746	0.10525	-0.11723	-0.19881	-0.05256	-0.10652	-0.07904
cry		-0.37510	-0.61063	-0.44016	-0.68133	-0.77477	-0.85520	-0.81291	-0.81109	-0.76701	-0.81759	-0.79552	-0.84120	-0.75328	-0.77201	-0.79708	-0.70155	-0.62041	-0.54606	-0.55791
throw	-0.36059	-0.38869	0.20291	0.15531	0.42638	0.07385	-0.12588	-0.01854	0.13973	0.00230	0.03200	0.03958	0.09734	-0.00252	0.01574	-0.04136	0.07539	0.06276	-0.04299	0.02658
scream			-0.17550	-0.60970	-1.17214	-0.55436	-0.69281	-0.37646	-0.28012	-0.60987	-0.38207	-0.75661	-0.62667	-0.49858	-0.49837	-0.58718	-0.47650	-0.39836	-0.38700	-0.30716
shout					0.33313	0.91073	0.84260	0.29559	0.36533	0.55275	0.52989	0.51394	0.46309	0.47463	0.30106	0.13039	0.06946	0.14090	0.14705	0.14661
say	-0.02108	-0.22041	-0.08038	-0.04183	-0.08445	-0.14462	-0.11759	-0.17475	-0.15121	-0.15492	-0.12033	-0.09672	-0.10303	-0.11436	-0.11300	-0.11099	-0.07292	-0.02735	0.00682	0.01262
speak	-0.05008	0.13282	0.04160	0.06067	0.11570	0.03889	0.13966	0.13296	0.10232	0.06707	0.08405	0.09876	0.10087	0.12287	0.11734	0.16888	0.17234	0.18548	0.14399	0.12220
work	-0.19784	0.15226	-0.01583	0.19487	0.12031	-0.03240	0.16855	0.23380	0.14400	0.25004	0.27384	0.24548	0.26573	0.22649	0.25555	0.21697	0.14324	0.14237	0.11872	0.09408
consider	0.31521	0.20480	-0.03991	0.32492	0.16065	0.30967	0.13962	0.14962	0.12394	0.11015	0.10729	0.25369	0.31708	0.22624	0.38538	0.44400	0.22647	0.17803	0.11064	0.07316
walk	0.00256	0.07147	0.02908	-0.01109	0.06511	0.03738	0.07262	-0.02845	-0.05861	-0.04182	0.05863	-0.00555	0.00183	-0.02151	0.00615	-0.06706	-0.05372	-0.01955	0.00013	0.00302
give	-0.03258	-0.06801	0.07985	-0.02694	0.03464	-0.01414	0.16154	-0.02597	0.07870	-0.03093	-0.03488	0.02616	-0.01501	0.02435	0.03298	-0.02507	-0.00152	0.00556	0.00367	-0.00020
read	-0.03886	-0.19166	-0.14594	-0.06868	-0.13261	-0.06281	-0.11304	-0.09859	-0.08166	-0.00684	-0.03846	0.01969	-0.01399	-0.02911	0.01546	0.02048	-0.01035	-0.01947	-0.03014	-0.02861
run	-0.32388	-0.05202	-0.09733	-0.34617	-0.17508	-0.01862	-0.04441	0.01242	-0.04310	-0.00868	0.07979	0.04593	0.02498	0.05884	0.03581	0.02374	0.11806	0.07051	0.04775	0.05371
watch	-0.40161	0.07952	-0.05070	-0.29967	-0.30141	-0.21208	-0.22326	-0.08976	-0.06000	-0.16733	-0.10465	-0.10477	-0.08763	-0.13571	-0.11698	-0.15007	-0.13036	-0.03411	0.02784	0.01902
write	0.55538	0.29082	0.30466	0.58678	0.44029	0.43223	0.50700	0.39209	0.34366	0.40601	0.38389	0.44592	0.33537	0.39448	0.39608	0.41708	0.32146	0.28615	0.22219	0.22448
play	-0.36549	-0.30831	-0.13198	-0.04678	-0.07105	-0.05234	-0.10633	-0.03499	-0.08103	-0.09741	-0.05590	-0.04452	-0.04275	-0.00824	0.01512	0.00092	0.02533	0.09937	0.15334	0.18186
take	0.30490	0.08830	0.20344	0.07256	0.02229	0.03981	0.03477	0.03512	0.09758	0.03674	0.03464	0.05684	0.04060	0.03104	0.05332	0.02653	-0.00701	0.00521	-0.00809	0.01023
look	-0.33270	-0.31375	-0.20860	-0.24118	-0.22355	-0.25354	-0.30864	-0.24749	-0.25827	-0.21829	-0.15693	-0.16119	-0.13279	-0.11137	-0.10130	-0.12114	-0.06779	-0.01238	0.00486	0.01280
start			-0.25471		0.42138	0.31843	0.25705	0.24772	0.22157	0.14838	0.18530	0.24697	0.09245	0.13578	0.17169	0.06841	0.06806	-0.00947	-0.07737	-0.08487
bring	0.78248	0.14330	-0.05415	0.12863	0.04277	0.07263	0.10103	0.06704	0.16694	0.17624	0.09895	0.03462	0.11362	0.09386	0.11191	0.08764	0.01955	0.15276	0.12978	0.06323
help		0.18416	-0.17563	-0.58786	0.03882	-0.08409	0.06671	-0.02245	0.02409	0.16765	-0.01649	0.03231	-0.05778	-0.02010	-0.03760	-0.02286	-0.01905	-0.01451	0.01953	-0.00721
try	0.10417	-0.21554	0.00192	-0.13603	-0.13758	-0.11168	-0.11985	-0.07971	-0.01713	0.02584	0.03499	0.05159	0.04558	0.06833	0.08181	0.05705	0.03259	0.05431	0.02491	0.04010
wait	-0.27098	-0.18077	-0.17269	0.06501	-0.09902	-0.05389	-0.05850	-0.14329	-0.10080	-0.06960	-0.11687	-0.15521	-0.10399	-0.16328	-0.14008	-0.18193	-0.13070	-0.11667	-0.05297	-0.04540
kill			0.19256	-0.09306	-0.23431	0.10254	0.05331	0.59099	0.41498	0.09139	0.24775	0.39768	0.55663	0.29327	0.22618	0.11831	0.31121	0.16590	0.27598	0.21759
listen	-0.75033	-0.44109	-0.42041	-0.29987	-0.27453	-0.35462	-0.33679	-0.40805	-0.16664	-0.28031	-0.18035	-0.20017	-0.14735	-0.14502	-0.23224	-0.18947	-0.10382	-0.10853	-0.02099	-0.01929
ask	-0.05884	-0.48550	-0.14574	0.27534	-0.02926	-0.23459	-0.32617	-0.21161	-0.06092	-0.07733	-0.07660	-0.00195	0.02372	0.01945	-0.05619	0.02037	-0.02872	0.02979	0.01423	0.00129
carry	-0.00770	0.11834	0.10764	0.06173	0.12573	0.04112	0.03326	0.02863	0.01211	0.04706	-0.02240	-0.04045	-0.08365	-0.10669	-0.12506	-0.15581	-0.14604	-0.13108	-0.09312	-0.06396
share					-0.26900	-0.16001		0.14395	-0.27784	0.12050	-0.14279	-0.21442	-0.06225	-0.19286	-0.05862	-0.18837	-0.15529	-0.06872	-0.06477	-0.08736
choose				0.66924		-0.48745	-0.24199	0.06341	-0.17909	-0.17748	0.02315	-0.39607	-0.09738	-0.02156	-0.00133	-0.03411	0.09716	-0.18660	-0.16417	-0.13486

Table 20. Summary table for the negative past progressive construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink							-0.55438	0.13843			0.40800		0.18585	0.13355	0.44365	0.03360	-0.00162	0.15799	0.19907	0.28796
eat					-0.88876	0.54347	0.10052	-0.62171	-0.44301	0.23687	-0.25673	-0.40437	-0.16828	0.00832	-0.10525	-0.31310	-0.32994	-0.46754	-0.23201	-0.47320
dance					-1.36655	-0.71786	-1.28451	-0.90691	-0.35353	-0.39517	0.20027	-0.26823	-0.38231	-0.67458	-0.16980	-1.21964	-0.57373	-0.51023	-0.33148	-0.36628
sing					-0.88433	0.09992	-0.60322	-0.21808	-0.79473	-0.42646	-0.30858	-0.69149	-0.52824	-0.22642	-0.21274	-0.20838	-0.68619	-0.35559	-0.14094	-0.04904
cook																				
talk			0.17939		-0.17385	0.36752	0.08255	-0.33371	0.20797	-0.07919	0.18050	0.03590	0.28827	0.21235	-0.23659	-0.02279	0.26980	0.08793	0.24933	0.11828
think			-0.30048	0.34521	-0.30779	-0.21919	-0.37750	-0.35344	-0.25074	-0.21185	-0.07125	-0.18099	-0.06259	-0.05645	0.08909	0.03010	0.02400	-0.00969	-0.01835	-0.01927
complain									-0.70793	0.57649	0.25960	-0.46673	-0.13955	-0.19621	-0.12825	-0.24842	-0.05824	-0.30382	-0.19898	-0.13352
cry				-0.93205	-1.28609	-1.12728	-0.86930	-0.84082	-1.06788	-0.85308	-0.92181	-1.09349	-0.83921	-0.84536	-0.82393	-0.98342	-0.90238	-0.57724	-0.55031	-0.60564
throw																				
scream																				
shout																				
say			-0.25070			-0.14165	0.16756	-0.01683	0.29298	0.09137	-0.31869	0.13514	-0.00899	0.26372	-0.40758	-0.05642	-0.01764	0.08467	0.13283	0.16661
speak	-0.38917	-0.64196	0.72496	-0.01579	0.21246	-0.28745	-0.01660	0.07866	0.09952	0.19583	0.20047	0.08270	0.08941	0.22574	0.08917	0.24751	0.14837	0.34324	0.21822	0.24664
work			-1.21544		0.07256	0.28770	0.60412	-0.17533	0.17672	0.16720	0.11392	0.10208	0.12264	0.02592	0.27572	0.01113	0.04425	0.00582	0.05070	0.08984
consider						-0.41332	-0.07620					0.79604	1.02331	0.00831	0.62710	0.45482	0.46542	0.25669	0.08615	0.14909
walk						0.15249	-0.45848	0.65267	-0.56316	-0.25864	0.10007	-0.19375	-0.15156	-0.32406	-0.37102	-0.14683	-0.23515	-0.06259	-0.10425	-0.22505
give			-1.29986	-0.99928	-0.25829	-0.08282	0.51747	0.27309	0.21087	0.15531	0.21315	0.16402	0.12794	-0.04843	-0.07655	-0.15693	-0.00916	0.01459	-0.05071	-0.06507
read		-1.49950	-0.56928	-0.45702	0.02127	-0.33049	-0.32528	-0.53011	-0.34651	-0.25734	-0.12891	-0.03576	-0.08043	-0.17838	-0.15126	-0.16711	-0.13331	-0.00110	-0.19791	0.00783
run							-0.44658		0.48761	0.24931	0.14743	0.50947	0.27327	0.16048	0.03580	0.26782	0.11648	0.21219	0.02430	0.10823
watch		-0.85469				0.06457	0.46916	-0.06629	0.00844	-0.37704	-0.34883	-0.05569	-0.17654	0.03649	-0.18659	-0.07742	-0.05029	-0.15825	-0.01906	-0.12286
write						0.74066	-0.02512	0.64060	0.58592	0.33660	0.48359	0.39014	0.23715	0.44678	0.29034	0.35959	0.18920	0.22464	0.18801	0.26727
play				-0.55647	-0.90243	-0.45252	-0.30074	-0.13433	-0.39258	-0.00585	-0.24808	-0.31376	-0.06303	-0.22961	-0.15830	-0.01052	0.03094	-0.13631	0.10401	0.04605
take		-1.16950	-1.44634	-0.02472	0.56841	0.09332	0.36634	-0.25576	0.23726	-0.01541	-0.08923	-0.05544	0.13559	0.03178	0.25230	-0.03617	-0.03247	-0.06777	-0.11418	-0.02002
look			-0.43318	-0.07947	-0.47942	-0.06251	-0.11628	-0.26383	-0.36573	-0.24822	-0.23158	-0.21313	-0.26196	-0.29245	-0.21035	-0.31846	-0.26497	-0.11269	-0.06406	-0.05400
start																				
bring										0.80732	-0.12537	0.51915	0.16187	0.43470	0.22881	-0.08348	0.19192	0.11033	0.13261	0.32270
help							-1.07865	-0.76557	-0.40706	-1.60213	-1.55800	0.13876	-0.39523	-0.52161	-0.24248	0.06522	0.16402	-0.08383	0.06951	0.05061
try				0.07323	0.50080	0.13183	0.24534	0.45534	0.47828	0.24574	0.17248	0.21246	0.21289	0.17691	0.21632	0.20509	0.32406	0.19810	0.14218	0.05120
wait				-0.15024			-0.20138	0.01322	-0.58401	-0.12328	-0.35278	-0.32777	-0.25774	0.13383	-0.17788	-0.07481	-0.41831	-0.30433	-0.11019	-0.27636
kill																				
listen				-1.33755	-0.28249	-0.11197	-0.38177	-0.27229	-0.27746	-0.47342	-0.33413	-0.61879	-0.51555	-0.44902	-0.50332	-0.50753	-0.35993	-0.31703	-0.12519	-0.13844
ask						-0.15881	0.02095	0.01323	0.33473	-0.02613	0.16206	0.09540	0.21188	0.12976	0.03880	0.17471	0.03633	0.06828	0.15504	0.11727
carry					0.11826		-0.07611	1.01676	0.57225	0.17874	0.28542	0.41056	0.06708	-0.00031	-0.10674	0.03697	-0.24069	0.19132	0.00398	-0.03848
share											-0.17239	-0.61224	-0.13937	0.37644	-0.09491	0.12026	-0.73225	0.17178	0.20068	0.05648
choose							-0.55438	0.13843			0.40800		0.18585	0.13355	0.44365	0.03360	-0.00162	0.15799	0.19907	0.28796

Table 21. Summary table for the positive present perfect construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink	-0.72033		0.32126	-0.06411	0.13097	0.19034	0.22965	-0.02808	-0.02955	-0.07773	-0.11069	-0.11079	-0.06454	0.15319	0.12818	0.02464	0.12158	0.18801	0.21521	0.05909
eat	0.05805	-0.32618	-0.12826	-0.26436	-0.10089	-0.15188	-0.05981	-0.00020	0.01379	-0.13923	-0.09056	-0.07175	-0.06273	-0.02297	-0.13979	-0.15848	-0.02728	-0.07738	-0.10052	-0.12412
dance	-1.02159		-0.45714	-0.72778	-0.28829	0.01047	-1.10644	-0.77576	-0.51996	-0.80736	-0.75458	-1.01798	-0.94582	-0.79628	-0.78132	-0.83580	-0.81097	-0.54982	-0.54735	-0.33646
sing	0.05805	-0.02468	-0.11472	-0.25491	-0.42517	-0.05492	-0.56460	-0.24066	-0.25161	-0.42500	-0.65078	-0.36986	-0.35169	-0.45701	-0.56154	-0.55679	-0.47846	-0.31429	-0.32056	-0.22841
cook				-0.82462	-1.61060	0.28349	-1.11601	-1.29173	-0.85507	-0.72923	-1.09868	-1.13278	-0.83232	-0.60391	-0.77143	-0.96730	-0.62871	-0.44882	-0.54565	-0.43166
talk		0.15152	-0.11476	-0.14671	-0.33220	-0.25527	-0.14620	-0.25783	-0.25848	-0.26000	-0.15762	-0.11971	-0.12928	-0.10320	-0.03223	-0.08633	-0.04733	-0.13871	-0.06797	-0.01349
think	-0.10885	0.03957	0.19357	0.20309	0.14746	0.22443	0.07991	0.13014	0.16477	0.04230	0.02537	0.12410	0.09573	0.12399	0.07918	0.04892	-0.01787	-0.02592	-0.07642	-0.03943
complain	-0.02118	-0.62716	-0.42156	-0.24109	-0.23398	0.12862	-0.45494	-0.24819	-0.48747	-0.61520	-0.60310	-0.10474	-0.54901	-0.32742	-0.10179	-0.31709	-0.26299	-0.31344	-0.25069	-0.31207
cry			-0.21825	-0.52378	-0.57662	-0.33133	-0.95031	-0.85800	-0.42087	-0.62249	-0.63744	-0.58391	-0.19538	-0.48652	-0.53518	-0.35284	-0.50826	-0.28373	-0.44162	-0.46291
throw	-0.26030	-0.11239	-0.03592	-0.03739	-0.08799	0.01101	-0.15373	-0.17789	-0.13432	-0.12676	-0.30902	-0.27310	-0.07527	-0.13490	-0.03124	-0.10460	-0.02115	0.10745	0.08415	0.08873
scream																				
shout																				
say	0.51287	0.23962	0.26440	0.25249	0.23043	0.19540	0.09271	0.06115	0.11239	0.05242	0.01951	0.06369	0.15068	0.10765	0.10408	0.13982	0.14149	0.08016	0.12582	0.06514
speak		0.41252	0.20536	0.31834	0.24704	0.04723	0.12381	0.01195	-0.00535	-0.02247	0.04882	-0.02392	0.08824	0.04019	0.10715	0.14452	0.12664	0.23102	0.19851	0.20085
work		0.15152	-0.40445	0.02053	-0.03598	-0.09873	0.10148	0.04786	0.09724	0.13291	0.05507	0.15594	0.01933	0.01676	0.03348	-0.03059	-0.17325	-0.19397	-0.17346	-0.13471
consider	-0.06706	-0.27076	0.27157	0.01686	0.12589	0.55791	0.29780	0.23458	0.10341	0.42805	0.18365	0.42779	0.13417	0.07829	0.14152	0.24121	0.10906	-0.08891	-0.15007	-0.13596
walk	-0.17631	-0.80398	-0.10180	-0.16904	-0.33625	-0.44374	-0.33298	-0.22599	-0.09703	-0.20104	-0.15633	-0.11407	0.02609	-0.02822	-0.12394	-0.15209	-0.21115	-0.10492	-0.01000	-0.04221
give	0.35714	0.29436	0.28194	0.27126	0.21995	0.15291	0.14588	0.14228	0.07487	0.03460	0.01819	0.01190	-0.00824	-0.02691	0.00566	-0.01029	0.03298	0.06280	0.14130	0.18201
read	-0.05374	-0.02310	0.00074	0.13786	0.11618	0.14737	0.09990	0.12632	0.13167	0.05653	0.13451	0.20297	0.08704	0.14483	0.11399	0.15245	0.12893	-0.04325	-0.05681	-0.13110
run	-0.28148	-0.24996	-0.04782	0.13986	-0.00706	0.15896	-0.15069	-0.15049	-0.11762	-0.19948	-0.32607	-0.24902	-0.24039	-0.17746	-0.16885	-0.23487	-0.22058	-0.12245	-0.06712	-0.18997
watch	-0.11813	-0.47236	-0.37828	-0.30129	-0.54228	-0.45274	-0.23237	-0.18168	-0.16336	-0.21979	-0.14527	-0.14228	-0.06120	-0.13958	-0.11989	-0.14017	-0.10653	-0.19318	-0.12731	-0.21665
write	0.13869	0.21837	0.12253	0.12136	-0.00786	0.10102	0.04645	0.11136	0.01972	0.12553	0.11771	0.14154	0.12835	0.16740	0.13239	0.12803	0.11093	0.08479	0.03307	0.01342
play	-0.26817	-0.39125	-0.14346	-0.05931	-0.22079	-0.19147	-0.07497	-0.05213	-0.16856	-0.18941	-0.22677	-0.13679	-0.08437	-0.07205	-0.14422	0.00376	-0.08188	0.05746	0.07556	0.04588
take	-0.00664	0.00045	0.05573	-0.07241	-0.02101	-0.04000	-0.04129	-0.02821	-0.02198	-0.04676	-0.07891	-0.04448	-0.03415	-0.05708	-0.03241	-0.03158	0.00172	0.00027	0.00682	-0.00125
look	0.05808	-0.11882	0.00532	-0.18465	-0.36129	-0.35833	-0.37906	-0.18146	-0.24186	-0.26045	-0.13919	-0.10857	-0.17440	-0.12575	-0.04218	-0.07326	-0.05545	-0.05788	0.02703	0.05230
start	0.02024	0.24843	-0.33228	0.39307	0.13051	0.20656	-0.15661	-0.13523	-0.09949	-0.08405	-0.16416	-0.11592	-0.05021	-0.05018	-0.05458	-0.04267	-0.08209	-0.16758	-0.17810	-0.15213
bring	0.00363	-0.02508	-0.04552	0.13459	0.03213	0.09022	-0.01264	-0.01739	-0.01207	-0.03857	-0.07199	-0.04732	-0.04062	-0.09378	-0.12157	-0.10730	-0.03106	0.05664	0.09156	0.06624
help			0.44624	0.16458	-0.06649	-0.00775	-0.08311	-0.09423	-0.03722	-0.12288	-0.06719	-0.01241	-0.06290	-0.09662	-0.10390	-0.07286	-0.13171	-0.10297	-0.10388	-0.06273
try		-0.07598	0.16163	0.19021	-0.00711	-0.06899	-0.08082	0.04787	0.02498	0.03023	-0.10731	0.03389	0.00225	-0.07393	-0.03066	0.03923	-0.03201	-0.04268	-0.08515	-0.14484
wait		-0.14967	-0.80122	-0.34766	-0.13342	-0.42148	-0.43392	-0.42115	-0.30867	-0.49165	-0.28061	-0.33834	-0.44115	-0.55119	-0.55606	-0.44120	-0.34301	-0.28285	-0.23977	-0.22257
kill	0.59230	0.02655	0.57437	0.50004	0.43209	0.33277	0.19088	0.33719	0.25338	0.26572	0.20660	0.23245	0.40704	0.26045	0.23642	0.15696	0.30284	0.32174	0.41968	0.36665
listen		-0.48098	0.18646	0.20689	-0.29309	-0.35515	0.02550	-0.25612	-0.27222	-0.03811	-0.26233	-0.08329	-0.06360	-0.32559	-0.06143	0.08449	-0.00959	-0.18459	0.03416	-0.11612
ask	-0.11813	-0.39312	-0.44961	-0.00726	-0.17467	-0.14277	-0.13183	-0.27605	-0.13335	-0.23258	-0.27539	-0.19099	-0.19868	-0.16079	-0.11697	-0.12319	-0.05967	-0.10230	-0.04945	-0.07769
carry	0.55275	-0.25141	0.09069	-0.17986	-0.01801	-0.08487	0.03441	0.10097	0.16904	0.16197	-0.16429	0.02632	0.06369	-0.11750	-0.07080	-0.04276	-0.09528	0.02834	0.00704	-0.00020
share	-0.47741	0.02650	-0.12041	-0.73812	-0.35904	-0.58682	-0.75234	-0.42239	-0.28519	-0.35319	-0.23334	-0.31864	-0.27434	-0.13099	-0.17003	-0.21186	-0.34963	-0.27332	-0.20114	-0.19070
choose	0.23837	0.12560	0.04203	0.04458	0.08246	0.05014	0.00313	-0.04448	0.04250	-0.00469	0.02191	0.06278	0.10897	0.05352	0.11335	0.15745	0.10439	0.01540	0.00658	-0.00841

Table 22. Summary table for the negative present perfect construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink																				
eat			-0.40146	-0.87646	0.15520	-0.23154	0.06268	-0.52354	0.05599	-0.21728	-0.19934	-0.03449	-0.26161	-0.44506	-0.06184	-0.19075	-0.11801	-0.27350	-0.22369	-0.31269
dance																				
sing									0.32372	-0.22935	-0.77388	-0.67503	-0.24138	-0.43080	0.11746	-0.78743	-0.67029	0.56877	-0.75518	-0.14947
cook																				
talk										0.52314	-0.25460	-0.59634				-0.29241	-0.14180	0.08361	-0.07842	0.05865
think	-0.54009	-0.13138	-0.06710	-0.04545	-0.04367	0.81416	0.11601	-0.01655	0.18114	-0.09199	0.03840	0.08793	-0.40201	-0.01378	-0.17169	-0.18542	0.09072	-0.13147	0.08980	0.08006
complain					-0.27111			0.09014	-0.39729	-0.33926	-0.10320	-0.19727	-0.89507	-0.43080	-0.12575		0.20543	-0.29693	-0.27009	0.41524
cry							-0.48889			-0.10739	-0.36616				0.05048	-0.30971	-0.66989	-1.28714	-0.13526	-0.22872
throw		-1.18643		-0.30013	-0.80860	-0.50787	-0.48939	-0.18798	-0.73088	-0.47422	-0.15765	-0.27659	-0.01932	-0.33414	-0.22763	-0.40196	0.04629	-0.17202	-0.02115	0.41504
scream																				
shout																				
say		0.38036	0.06338	0.57865	0.43416	0.23441	0.03933	0.24702	0.08237	-0.07336	0.06842	-0.02686	0.06394	0.24796	-0.05256	0.00740	-0.13925	0.10844	0.09991	0.22157
speak		-0.22981	-0.36205	-0.15465	0.08001	-0.30066	-0.33224	-0.29034	-0.20564	-0.11489	0.04399	-0.10566	-0.30331	-0.29739	-0.28273	0.06184	-0.13892	-0.14172	-0.12541	-0.18681
work						-0.38275	0.61541		-0.28324	0.18660	-0.00872	0.45455	0.34202	0.37714	0.14877	-0.00056	-0.24567	-0.35828	-0.09899	-0.06507
consider					0.55395	0.05768		0.04446	0.24460	0.30166	0.41922	0.17133	-0.15448	0.70102		0.55231	0.40644	0.24389	0.28763	0.07520
walk					-1.22624		-0.96617	-0.74105	-0.63242	-0.49026	-0.45920	-0.77258	-1.02178	-0.55579	-0.87515	-0.30994	-0.56129	-0.29714	-0.41979	0.09565
give	-0.03288	1.09910	0.27070	0.28081	0.25771	0.30510	0.18434	0.19305	0.02391	0.07801	0.08933	0.14101	0.09305	-0.06002	0.11696	0.07692	0.10290	0.10710	0.03567	0.19278
read		0.11954	-0.22725	0.11811	0.09292	0.32037	0.37887	0.30697	0.22586	0.34000	0.19738	0.40308	0.02136	0.79949	0.24739	0.19789	0.28180	0.04698	0.13362	0.23503
run			0.07647		0.25224	0.09555	-0.36367	-0.12398	0.10985	-0.64783	-0.59267	-0.57008	-0.37664	-1.57850	-0.17175	-0.47180	-0.01661	0.22649	0.02334	0.57809
watch																				
write		-0.30713	0.72737	-0.44953	-0.33716	-0.15629	-0.18207	-0.12721	-0.02797	-0.02878	-0.07152	0.05168	-0.00476	-0.19456	0.00926	0.02600	0.16935	0.08865	-0.03449	0.19248
play						-0.28527		0.39148	-0.17872	-0.55296	0.31050		0.33706	-0.29880	0.70454	-0.11350	-0.05597	0.09158	-0.09189	0.04621
take	0.34391	-0.26764	0.10245	-0.08792	0.04909	0.25510	-0.07074	0.06700	0.08261	0.07809	0.07817	0.05299	0.09544	0.07640	0.10728	-0.01254	0.02651	-0.08130	0.04092	0.00853
look			-0.24245	-0.65623	0.61485	0.35913	0.11410	-0.36869	-0.21500	-0.07968	0.12754	0.13078	-0.03044	0.11413	0.17615	0.13632	-0.13328	-0.19278	0.07444	-0.07058
start							0.03451		0.50033	0.59786		0.22914			0.05050	0.05846	-0.06787	-0.53151	-0.44375	-0.18260
bring			0.48796	0.57855	0.63465	-0.24422	0.12271	0.19783	0.17669	0.07093	0.28430	0.05507	0.38405	-0.05254	-0.06306	-0.10161	0.28225	0.35380	0.18166	0.27021
help					-0.35053			-0.38733	-0.06650		-0.44188	0.29636	0.37941	0.22314		-0.06372	-0.03784	0.90805	0.04110	0.09373
try					-0.04916		-0.18827	0.19740	0.20834	0.33806	-0.13335	-0.04328	-0.26594	0.02203	0.41335	0.23102	0.02060	0.16897	-0.29227	0.12754
wait																				
kill							0.11377		0.42755	0.05922	0.53257	0.98053	0.53768	-0.36398			0.67359	0.20429	0.18288	0.23010
listen																				
ask			-0.57769	0.26494	0.55493	0.24790	0.14932	0.03011	0.39874	-0.02397	0.13668	-0.08280	-0.56149	-0.16409	-0.05268	0.09222	-0.00329	-0.08048	-0.03541	-0.14104
carry		-1.37814	-0.03302	-0.56034	-0.56202	0.29573	0.73507		0.76802	-0.00006	0.07679			0.12925	0.20746	0.16108	0.29463	0.61655	1.21091	0.77808
share								0.01091	-0.00746	0.05910	-0.03253	-0.33898	0.60478	0.17175		0.32760	0.16302	0.05547	-0.13986	-0.28084
choose		0.16226	-0.50480	0.50442	0.14672	-0.42227	-0.40980	-0.24250	-0.12893	0.05671	-0.09370	0.19936	0.03888	0.44663	-0.00994	0.18651	-0.03814	-0.23061	0.09175	-0.09028

Table 23. Summary table for the positive past perfect construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink	0.75362	0.31051	0.01949	0.26630	0.25299	0.27657	0.31267	0.17322	0.33229	0.29540	0.22244	0.24400	0.17443	0.19764	0.06660	0.07038	0.06627	0.04849	0.09861	0.14308
eat	-0.12425	-0.05651	0.06726	0.11659	0.13295	0.02419	-0.03331	0.05544	0.08604	0.10032	0.07071	0.08075	0.02759	-0.01947	-0.01967	-0.05365	-0.08755	-0.04789	-0.06784	-0.06434
dance		-1.17696	-0.86129	-0.81681	-0.89765	-0.45016	-0.36248	-0.51022	-0.45277	-0.43894	-0.58075	-0.55832	-0.53138	-0.65628	-0.68711	-0.71316	-0.61809	-0.48254	-0.48807	-0.50141
sing	-1.00575	-0.83530	-0.32546	-0.67092	-0.41341	-0.49068	-0.52663	-0.38527	-0.47661	-0.37792	-0.29809	-0.48930	-0.38390	-0.37628	-0.43960	-0.43466	-0.37488	-0.27780	-0.26822	-0.24395
cook				-0.23100	-0.73537	-0.38324	-0.52456	-0.35266	-0.36265	-0.40927	-0.41398	-0.52644	-0.79900	-0.77869	-0.69287	-0.81184	-0.82839	-0.79990	-0.65940	-0.57523
talk	-0.43082	-0.20050	0.00076	-0.13426	-0.12578	-0.11985	-0.11035	-0.04462	-0.01913	-0.04956	-0.04693	0.00628	-0.00024	0.03042	-0.00132	0.03588	0.07236	0.05043	-0.04968	-0.02449
think	0.15081	0.00000	-0.01213	-0.13695	-0.12734	-0.20372	-0.17175	-0.18302	-0.16440	-0.15985	-0.17254	-0.15929	-0.16927	-0.18959	-0.13021	-0.13314	-0.15014	-0.19143	-0.19152	-0.17686
complain	-0.29424	-0.02349	-0.20542	0.02487	-0.13691	-0.22286	-0.15357	-0.09264	-0.13173	-0.19535	-0.15205	-0.15656	-0.14280	-0.20507	-0.19742	-0.26685	-0.16000	-0.14895	-0.14744	-0.16609
cry	-0.27836	0.02769	-0.40969	0.04628	-0.16630	-0.40883	-0.50635	-0.38023	-0.57505	-0.50962	-0.53707	-0.63596	-0.64396	-0.62581	-0.63414	-0.62777	-0.61672	-0.61278	-0.54422	-0.55038
throw	-0.38162	-0.22319	-0.23340	-0.21375	-0.21374	-0.11724	-0.19218	-0.17544	-0.13469	-0.19094	-0.14420	-0.13906	-0.09377	-0.06936	-0.09597	-0.13687	-0.13119	-0.10981	-0.07752	-0.07875
scream					-1.68957	-1.26783	-0.84901	-0.62249	-0.64339	-0.83908	-0.97867	-0.78300	-0.80144	-0.80302	-0.82393	-0.85076	-0.74492	-0.52325	-0.62352	-0.50545
shout						0.77362	0.53128	0.29975	-0.15479	0.45905	0.33909	0.40835	0.16291	0.22372	0.21095	-0.02843	-0.06820	0.02769	0.05572	0.08141
say	0.22885	0.19482	0.18305	0.24150	0.16042	0.01188	-0.01516	-0.04256	-0.08949	-0.14243	-0.13872	-0.12551	-0.16384	-0.12638	-0.12262	-0.14011	-0.09116	-0.08108	-0.06015	-0.06258
speak	0.39168	0.16851	0.05476	0.08301	-0.04667	-0.09200	-0.05938	-0.08522	-0.08076	-0.08989	-0.06078	-0.05447	-0.07283	-0.07033	-0.02934	-0.03441	0.02534	0.02592	0.01621	0.01183
work	-0.50862	-0.56973	-0.24518	-0.35781	-0.17216	0.00438	-0.01725	0.05258	0.13551	0.14060	0.10593	0.10879	0.11423	0.13650	0.10357	0.08361	0.04624	0.05141	0.02697	0.02378
consider	-0.01894	0.34832	0.03783	0.20235	0.12574	-0.02220	-0.01021	0.14428	-0.02837	0.06582	0.06764	0.07287	0.17968	0.10452	0.07856	0.17440	0.11199	0.01870	0.03353	-0.00045
walk	-0.13027	-0.28768	-0.16619	-0.13762	-0.13484	-0.09748	-0.11284	-0.08223	-0.09116	-0.07261	-0.07021	-0.09080	-0.06635	-0.04174	-0.08967	-0.11625	-0.09614	-0.08745	-0.08195	-0.06570
give	0.06154	0.05155	0.04786	0.04334	-0.00412	-0.01907	-0.01719	-0.03270	-0.05383	-0.04896	-0.06444	-0.07326	-0.05526	-0.06220	-0.04956	-0.06833	-0.04237	-0.03189	-0.03877	-0.03494
read	-0.03559	-0.00160	-0.00069	-0.06271	-0.04223	-0.06747	-0.06877	-0.05391	-0.03790	-0.06502	-0.04647	0.01332	0.01367	0.04465	0.06648	0.05971	0.02535	0.00871	-0.01427	-0.02199
run	-0.02404	0.28800	0.16366	0.02648	-0.05889	-0.06450	-0.19388	-0.06517	-0.13824	-0.07265	-0.17138	-0.12840	-0.15980	-0.18087	-0.16051	-0.14481	-0.15073	-0.10478	-0.06931	-0.08254
watch	0.23383	-0.22132	-0.26705	-0.40385	-0.29052	-0.31759	-0.29231	-0.27457	-0.20966	-0.18302	-0.14470	-0.17662	-0.11821	-0.06153	-0.08741	-0.10364	-0.09701	-0.09492	-0.08400	-0.06360
write	0.38764	0.26438	0.26500	0.22912	0.17695	0.15631	0.17228	0.16535	0.16897	0.19060	0.17591	0.22987	0.21621	0.22146	0.19341	0.23140	0.17351	0.16820	0.13396	0.13715
play	0.05433	-0.61074	-0.21594	-0.09720	-0.15346	-0.04995	-0.19643	-0.04076	-0.09845	-0.10689	-0.06346	-0.05511	-0.02145	-0.00489	0.04742	0.02127	0.01338	0.05366	0.14105	0.11542
take	0.02826	0.08702	0.05099	0.04276	0.07870	0.08690	0.05460	0.05557	0.04717	0.04245	0.04275	0.01881	0.02906	0.02298	0.01353	0.00876	0.01932	0.00842	0.02915	0.02060
look	-0.07526	-0.20942	-0.24123	-0.23538	-0.22263	-0.25453	-0.26013	-0.23123	-0.22236	-0.23230	-0.22173	-0.22182	-0.24427	-0.23713	-0.23389	-0.26295	-0.23483	-0.19264	-0.18717	-0.16252
start	0.37516	-0.43723	-0.22131	0.04386	-0.10804	-0.07643	0.01570	-0.02230	0.12156	0.07601	0.07957	0.09026	0.09892	0.10448	0.06157	0.04453	-0.00515	-0.03328	-0.03970	-0.03230
bring	-0.22470	-0.10243	-0.06051	0.01364	-0.02189	-0.03286	-0.00550	-0.02324	-0.01256	-0.00874	-0.00506	-0.01299	-0.02303	-0.01358	-0.03793	-0.03590	-0.01577	-0.01740	-0.01696	-0.00558
help	-0.77155	-0.23113	0.03017	-0.09901	-0.10240	-0.05681	-0.02061	0.03124	0.04511	0.07329	0.03789	0.11843	0.06707	0.16694	0.12119	0.15374	0.13491	0.18664	0.17401	0.13520
try	-0.15374	-0.04916	-0.17110	-0.16640	-0.04233	-0.15971	-0.14794	-0.08744	-0.11092	-0.07194	-0.07671	-0.06884	-0.08718	-0.03682	-0.04061	-0.01721	-0.04631	-0.07104	-0.08987	-0.07614
wait	-0.18142	0.04662	-0.06941	-0.15577	-0.11663	-0.20157	-0.25282	-0.15026	-0.15634	-0.16622	-0.19452	-0.21849	-0.21339	-0.19749	-0.19899	-0.18013	-0.19259	-0.16327	-0.17769	-0.15349
kill	0.29119	0.42822	0.64461	0.66147	0.54246	0.50808	0.55948	0.59599	0.50514	0.54365	0.52982	0.46288	0.49754	0.44162	0.44687	0.33060	0.37184	0.39303	0.37700	0.32252
listen	-0.41305	-0.27379	-0.18896	-0.35283	-0.33168	-0.25518	-0.17026	-0.18421	-0.24591	-0.26566	-0.17677	-0.16832	-0.19807	-0.16579	-0.13923	-0.12343	-0.09178	-0.08171	-0.08431	-0.07073
ask	-0.29557	-0.27406	-0.36158	-0.09972	-0.13883	-0.19121	-0.14811	-0.19915	-0.10317	-0.20475	-0.12811	-0.14856	-0.14972	-0.12050	-0.09609	-0.09668	-0.09776	-0.08116	-0.07468	-0.09413
carry	-0.09794	-0.02916	0.14536	0.23441	0.08872	0.13129	0.11791	0.13396	0.15133	0.11339	0.15781	0.09434	0.06655	0.09779	0.08623	0.04823	0.07646	0.05269	0.07111	0.05326
share	-0.00517	-0.07385	-0.27247	-0.16497	-0.12456	-0.11009	-0.02318	-0.17194	-0.14870	-0.06134	-0.17349	-0.07779	-0.20435	-0.05866	-0.15623	-0.16698	-0.22638	-0.23934	-0.21928	-0.22085
choose	0.27467	0.11143	0.04444	0.05854	0.03686	-0.01472	0.02507	0.00470	-0.02292	-0.03756	-0.01550	-0.05157	-0.01890	0.01451	0.01106	0.02427	0.00924	-0.05920	-0.06623	-0.07042

Table 24. Summary table for the negative past perfect construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink			0.06220	-0.20185	0.17258		0.18338		0.05256	0.50738	0.25363	0.22431	0.20604	0.26113	0.10928	0.30189	0.14951	0.12948	0.05472	0.10175
eat		-0.50221	-0.01708	-0.47393	-0.01918	-0.02010	-0.13274	-0.05945	0.15596	0.15879	0.09162	0.10802	0.14136	-0.01505	0.14703	0.01924	-0.00601	0.02718	-0.01887	-0.05472
dance				-1.52522	-1.03887	-1.01455	-0.51622	-0.39462	-0.79983	-0.94971	-0.63190	-0.47089	-0.34942	-0.31245	-0.29328	-0.89438	-0.32793	-0.86505	-0.35805	-0.19326
sing	-0.93732	-0.72377	0.18724	-0.20198	-0.49504	-0.12175	-0.56049	-0.31154	-0.72062	-0.21284	-0.61927	-0.20421	-0.39615	-0.50516	-0.53693	-0.35171	-0.18188	-0.53585	-0.32723	-0.29610
cook																				
talk					-0.12866	-0.90520	0.43877	0.28247	-0.06953	-0.32287	0.10200	0.09739	0.18921	-0.02535	0.07421	-0.05451	0.04508	-0.02005	-0.04975	-0.03642
think	-0.25844	0.06433	-0.02263	-0.30724	-0.31089	-0.25226	-0.17600	-0.14059	-0.13757	-0.15476	-0.06124	-0.15285	-0.03045	-0.10449	-0.02564	-0.10556	-0.12378	-0.17630	-0.16840	-0.12657
complain			0.12920	-0.37871	-0.71672	-0.60408	-0.71279	0.14961	-0.31569	-0.08786	-0.06414	-0.24063	-0.26738	-0.43067	-0.22282	-0.21435	-0.08378	-0.22703	-0.36162	-0.28969
cry					0.08663	-0.53696	-1.14546	-0.40338	-0.43028	-0.49853	-0.58705	-0.89846	-0.83681	-0.73443	-0.56165	-0.70920	-0.53803	-0.27006	-0.41376	-0.36924
throw	0.06449	-0.12082	-0.50664	-0.13502	-0.25542	-0.36131	-0.17725	0.13387	-0.00382	-0.23848	-0.02338	-0.11525	-0.22819	0.05323	0.26500	0.25791	-0.00658	0.16574	-0.05425	-0.14321
scream										-0.69505	-0.77114		-0.43353	-0.33860	-0.80125	-1.04683	-0.42481	-0.92376	-0.34559	-0.64049
shout																				
say	0.13302	0.29976	0.11584	-0.23708	0.04169	-0.08936	-0.14072	-0.12355	-0.04803	-0.05396	-0.14410	-0.19666	-0.10837	-0.13495	-0.15751	-0.08320	-0.02951	-0.06482	-0.02575	-0.10315
speak	0.01861	0.01210	0.05185	-0.19074	-0.10119	-0.25099	-0.22464	-0.11798	-0.07780	-0.22757	-0.13774	-0.16582	-0.21549	-0.18777	-0.21873	-0.18376	-0.17096	-0.09902	-0.13063	-0.10539
work				-0.73925	-0.80444	-0.31554	0.02836	0.18704	0.49526	0.26469	0.24642	0.49958	0.23369	0.07898	0.19403	0.06770	-0.06311	-0.12891	0.02917	0.01620
consider	-0.57162	0.62177	0.06241	0.47517	0.04312	0.16921	0.28680	0.09969	-0.06632	-0.15160	-0.11844	0.07351	0.06927	0.16447	0.01409	0.09129	0.01280	-0.04180	-0.12363	-0.13300
walk		-0.52087	-0.14226	-0.03273	-0.58667	-0.02737	-0.02849	-0.03285	0.00810	0.30961	-0.10462	-0.13165	0.16882	-0.10051	0.02282	-0.30280	-0.02683	-0.02442	0.13244	0.13424
give	0.19623	0.55035	0.28556	0.20014	0.22388	-0.03724	-0.01455	-0.03289	0.01223	-0.00195	0.05729	-0.00342	0.04146	0.06352	0.00488	0.03156	0.04518	0.03579	-0.02092	0.03140
read		0.13278	0.11891	0.22015	0.26157	0.12363	0.30836	0.10955	0.09625	0.04646	0.15593	0.22182	0.28843	0.32108	0.21448	0.32723	0.21879	0.20383	0.18593	0.19777
run	-0.09101		-0.07931	0.06157	-0.05873	0.00749	0.45814	0.34139	0.04158	-0.12938	-0.13866	-0.13639	-0.12391	0.06392	0.08796	0.13420	-0.05115	0.03176	-0.07975	-0.11676
watch					-0.21460	-0.23578	-0.11790	-0.17713	-0.26071	-0.54916	0.03861	-0.03449	-0.24458	-0.25953	0.26599	0.06265	-0.16518	-0.12526	0.21759	-0.01411
write	0.14476	0.08640	0.00188	-0.06427	0.08282	0.10386	0.24079	0.03416	0.08217	0.07633	0.03373	0.21707	0.25283	0.12223	0.16169	0.12007	0.06907	0.11326	0.12642	0.13945
play			-0.31833	-0.34370	0.03304	0.19057	-0.05346	0.34347	-0.08466	0.11033	-0.04478	-0.10083	-0.20829	-0.05787	0.02127	0.13601	0.12028	0.13756	0.07011	0.25241
take	-0.19082	0.03705	-0.01577	0.15692	0.16105	-0.01212	0.03084	0.08557	0.14949	0.02227	0.05459	0.03653	0.08031	0.05293	0.00839	0.05588	0.09729	0.08413	0.04838	0.02889
look			-0.25356	-0.19386	-0.27488	-0.30739	-0.39440	-0.44690	-0.41394	-0.30152	-0.02340	-0.06216	-0.15518	-0.07177	-0.15678	-0.16137	-0.27240	-0.15430	-0.10625	-0.02981
start						0.30861	-0.18014	0.26923	0.32157	0.23626	0.17437	0.03140	-0.11141	0.02561	0.31808	0.09633	-0.07500	0.15667	0.11101	-0.06342
bring	-0.32491	-0.15346	-0.29048	-0.14993	-0.15142	-0.21211	-0.03205	-0.19215	-0.07907	0.06894	0.09972	0.07334	0.03313	0.03649	0.00741	-0.04688	-0.05460	0.01381	-0.01874	0.10872
help					0.05855	0.57763	-0.32427	0.02463	0.44653	-0.20911	-0.24634	-0.11106	-0.07886	0.03041	0.16431	-0.04657	0.14091	0.11069	0.18174	0.21925
try			-0.19341	0.06152	0.08670	-0.20614	-0.15846	0.14554	0.17549	0.24663	0.12590	0.07351	0.00380	0.04270	-0.05459	0.12916	0.04046	-0.07141	-0.07274	0.10007
wait		0.05552	0.92949	0.40137	0.01723	-0.17177	0.08015	0.00468	-0.18927	-0.13629	-0.16004	-0.16620	-0.05214	-0.03549	-0.15396	-0.20314	0.02844	-0.08477	0.01892	-0.22927
kill			-0.27717	0.18669	0.15825	0.36690	0.04501	0.75455	0.51745	0.59067	0.56668	0.40631	0.68212	0.85813	0.43293	0.27017	0.34029	0.34464	0.47532	0.40766
listen	-0.23828	0.18071	0.15931	-0.23987	0.17279	-0.14868	-0.29975	-0.41860	-0.29540	-0.52857	-0.29469	-0.21103	-0.42538	-0.31142	-0.15245	-0.08889	-0.08727	-0.14394	0.15870	0.10246
ask		-0.42290	0.30240	-0.54554	-0.27443	-0.02494	-0.16919	-0.02737	-0.13501	-0.06548	-0.06313	-0.04755	-0.00149	-0.00089	-0.10088	-0.00324	-0.10754	-0.06466	-0.01528	-0.06741
carry	-0.23754	-0.72483	0.26218	0.62270	0.27217	0.73147	0.21690	0.11279	0.31104	0.16586	0.16748	0.11456	0.22013	0.44032	0.28621	0.17134	0.41414	0.29182	0.67649	0.28677
share			0.28426		0.53591	0.24178	0.12783	0.20784	0.13071	-0.12234	-0.03993	-0.16542	-0.23605	0.15520	-0.27874	-0.02538	-0.19601	0.00291	-0.18845	-0.04166
choose		0.39218	-0.01705	0.01216	0.21100	-0.09776	0.08681	-0.14151	-0.18732	-0.07630	-0.04542	-0.02677	-0.00754	0.24147	0.11508	0.23892	-0.16655	-0.06752	-0.02864	-0.03946

Table 25. Summary table for the *would* construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink	0.30688	0.23578	-0.12282	0.08748	0.15319	0.09071	0.28208	0.23095	0.40377	0.19181	0.09384	0.27181	0.26943	0.29965	0.15105	0.02486	0.10982	0.13932	0.14044	0.09965
eat	0.11130	0.07355	0.18906	0.01249	0.15753	0.18498	0.17537	-0.06730	0.06074	0.07690	0.14590	0.06018	0.00601	-0.03817	-0.01627	-0.02163	-0.11353	-0.03307	-0.00656	-0.02160
dance		-1.22210	-0.58949	-0.27080	-0.60688	-0.52529	-0.47112	-0.75011	-0.80238	-0.49651	-0.36280	-0.55950	-0.37914	-0.55709	-0.47646	-0.53456	-0.37955	-0.31540	-0.35424	-0.33156
sing		-0.19383	-0.69505	-0.62071	-0.52495	-0.46118	-0.45728	-0.65868	-0.47957	-0.45645	-0.43914	-0.32711	-0.26369	-0.37355	-0.42193	-0.36657	-0.31396	-0.12074	-0.19893	-0.19017
cook				-1.32006	-0.92573	-1.58318	-0.48500	-0.85286	-0.34281	-0.48745	-0.49162	-0.38108	-0.80509	-0.74676	-0.92051	-0.78652	-0.87598	-0.80535	-0.67214	-0.52664
talk	0.31595	-0.09077	-0.28218	-0.43182	-0.22742	-0.11843	-0.14624	-0.15268	-0.14652	-0.09172	-0.02111	-0.02204	-0.02066	-0.04558	-0.09514	-0.05095	-0.04357	-0.02428	-0.03394	-0.00510
think	-0.01884	-0.21787	-0.14352	-0.09372	-0.15135	-0.25583	-0.23964	-0.25053	-0.18923	-0.21501	-0.19331	-0.22228	-0.21802	-0.19506	-0.17562	-0.19297	-0.16095	-0.14833	-0.12424	-0.10264
complain	0.00565	-0.05542	-0.19731	-0.29201	-0.01181	-0.08118	0.01837	-0.13531	-0.07393	-0.18225	-0.29734	-0.33002	-0.34130	-0.32926	-0.35911	-0.34275	-0.29812	-0.24376	-0.17521	-0.15680
cry	-0.63510	-0.51100	-0.62002	-0.22399	-0.28847	-0.22964	-0.38329	-0.33411	-0.29024	-0.31692	-0.23473	-0.20629	-0.38461	-0.45573	-0.50193	-0.49647	-0.51875	-0.40374	-0.35925	-0.41007
throw	-0.28772	-0.10158	0.01707	-0.06045	-0.00467	-0.00561	-0.02868	0.04997	-0.04439	-0.07843	-0.02243	0.02690	0.01087	-0.09871	-0.01767	-0.10083	-0.05578	-0.00823	-0.00641	0.00601
scream		-1.09706	0.27596	-0.77598	-0.51211	-0.06024	-0.65520	-0.86494	-0.63113	-0.55768	-0.65556	-0.74490	-0.66781	-0.71812	-0.82424	-0.75110	-0.64924	-0.55364	-0.46245	-0.45968
shout				0.53739	0.38404	1.05043	0.39436	0.83788	0.67640	0.48359	0.62618	0.36962	0.32407	0.26394	0.12998	0.07300	0.12511	0.23462	0.13298	0.16260
say	-0.17301	-0.07817	-0.09495	-0.15168	-0.08008	-0.06515	-0.08299	-0.06468	-0.09927	-0.09584	-0.06455	-0.06720	-0.11079	-0.12496	-0.16081	-0.17211	-0.14465	-0.08633	-0.05661	-0.04361
speak	-0.28185	-0.04580	-0.03695	-0.06371	0.02989	-0.09457	-0.03598	-0.09940	-0.05130	-0.05529	-0.10387	-0.00248	0.01835	-0.01683	0.03172	-0.00982	0.04047	0.03712	0.03276	0.07424
work	0.19091	0.60246	-0.04631	0.24182	0.11063	0.23217	0.07270	0.12548	0.16461	0.18561	0.16736	0.24678	0.10304	0.14422	0.05252	0.09636	0.07171	0.00917	0.04639	0.06461
consider	0.23359	-0.03860	0.31151	-0.02845	0.05314	0.08924	0.31300	0.15104	0.20722	0.21832	0.19604	0.25014	0.27765	0.21505	0.31538	0.18928	0.20777	0.16882	0.09328	0.05497
walk	-0.01016	0.04940	-0.10286	-0.14693	-0.09492	-0.13899	-0.10164	-0.05611	-0.08219	-0.00930	-0.07077	-0.01252	-0.03179	-0.01319	-0.06812	-0.06933	-0.03696	-0.06559	0.00095	-0.00263
give	0.16727	0.11897	0.18047	0.18826	0.14973	0.16548	0.12840	0.12054	0.13632	0.10349	0.09905	0.08995	0.08753	0.07520	0.04114	0.01461	0.01824	0.06038	0.06581	0.04322
read	-0.36836	-0.21732	-0.20109	-0.27185	-0.20230	-0.11711	-0.06695	-0.18026	-0.14654	0.01122	-0.08185	0.08642	-0.02238	-0.01823	-0.01599	-0.08282	-0.07724	-0.04033	-0.04849	-0.07250
run	-0.28276	-0.24949	-0.26977	-0.11612	-0.06977	-0.19713	-0.06073	-0.02152	-0.09677	-0.09280	-0.06258	-0.11639	-0.11550	-0.07586	-0.05298	-0.03765	-0.00844	0.01016	-0.00858	-0.00378
watch	-0.11936	-0.07564	-0.04120	-0.25438	-0.27369	-0.33021	-0.20716	-0.26950	-0.11419	-0.07771	-0.07089	-0.09351	-0.15280	-0.08733	-0.20435	-0.16677	-0.09873	-0.10903	-0.08456	-0.06853
write	-0.06296	-0.12541	0.05498	-0.05626	-0.00828	-0.00203	0.06002	-0.03276	-0.02907	-0.03859	-0.00946	-0.00699	0.04586	0.00797	0.03765	0.09321	0.04942	0.06502	0.09634	0.09719
play	0.12382	0.41539	-0.05991	-0.22011	-0.22195	-0.23902	-0.21397	-0.01620	-0.06601	-0.07792	-0.05018	-0.08294	-0.06293	-0.06642	-0.08755	-0.06156	-0.05335	0.03740	0.13860	0.16131
take	-0.06919	0.02829	0.03269	0.05069	0.05782	0.07502	0.05241	0.00428	0.04024	0.05634	0.07034	0.08389	0.05154	0.06203	0.03715	0.02419	0.04345	0.06632	0.06660	0.09119
look	-0.22714	-0.28967	-0.30628	-0.25174	-0.29046	-0.29940	-0.23664	-0.29575	-0.23840	-0.16013	-0.18553	-0.21320	-0.20270	-0.18235	-0.21104	-0.22430	-0.18287	-0.13947	-0.13301	-0.12652
start	-0.52988	0.02544	-0.28865	-0.20219	-0.27652	-0.23328	-0.18771	-0.13946	-0.04889	-0.02747	0.02861	-0.01950	0.04744	0.06092	-0.07140	-0.06948	-0.06617	-0.03050	-0.05896	-0.05219
bring	0.11174	0.20088	-0.00826	0.07378	0.09319	0.10011	0.04031	0.01300	0.03826	0.09684	0.00410	-0.01957	0.04604	0.04260	-0.04896	-0.02845	0.03949	0.06059	0.08569	0.07056
help	0.18179	0.14881	-0.14121	-0.34198	-0.29203	0.02920	-0.13393	-0.05800	-0.04637	-0.02146	-0.08506	-0.06593	-0.10720	-0.03261	-0.10541	-0.13884	-0.02834	0.00920	0.00001	0.01449
try	-0.19894	-0.20948	-0.27979	-0.17073	-0.22689	-0.18797	-0.19805	-0.14568	-0.10566	-0.05573	-0.05277	-0.01266	-0.08842	0.00875	-0.00019	0.01251	-0.01568	0.01155	-0.00874	0.02244
wait	0.18205	-0.16662	0.16334	0.04708	-0.02609	-0.04757	-0.03513	-0.04207	-0.07750	-0.17129	-0.09458	-0.16195	-0.15106	-0.16027	-0.17198	-0.13143	-0.10225	-0.08217	-0.06785	-0.03798
kill	0.11500	0.26050	0.19861	0.14545	0.51781	0.33173	0.30540	0.40063	0.38396	0.37142	0.41619	0.28366	0.23755	0.22028	0.21981	0.16726	0.18098	0.23549	0.29329	0.26951
listen	-0.00814	-0.00238	-0.17198	-0.32145	-0.14505	-0.27136	-0.27978	-0.23518	-0.20878	-0.18365	-0.24187	-0.11319	-0.13249	-0.10900	-0.17256	-0.19426	-0.09732	-0.07880	-0.00612	0.03705
ask	0.30739	-0.19024	-0.03362	-0.24750	-0.13068	-0.13025	-0.09909	-0.16424	-0.09393	-0.12779	-0.10653	-0.09053	-0.11080	-0.12422	-0.12192	-0.13440	-0.04799	-0.09252	-0.03236	-0.05198
carry	0.31259	0.13220	0.19167	-0.09344	0.02198	0.11292	0.12780	0.24221	0.11388	0.09682	0.10532	0.10669	0.08670	0.05798	0.13525	0.14824	0.09295	0.10014	0.06706	0.07985
share	-0.59663	-0.92121	-0.43979	-0.26970	-0.29488	-0.33003	-0.24012	-0.28405	-0.13312	-0.16200	-0.21022	-0.10184	-0.12343	-0.18305	-0.11352	-0.13940	-0.17275	-0.17533	-0.17124	-0.12002
choose	-0.31867	-0.27604	-0.21547	-0.08122	-0.12190	0.03581	-0.07853	-0.05106	0.10425	-0.00208	-0.10065	0.10379	0.11997	0.12473	0.06957	0.12327	0.06654	-0.00507	-0.02793	-0.03706

Table 26. Summary table for the *would not* construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink			1.23334	0.95985	0.67645	0.64876	1.22057	0.73233	0.62821	1.12948	0.56145	0.16126	0.20890	0.24912	0.31443	0.26903	0.21417	0.39045	0.25720	0.25874
eat	-0.88780	-0.50231	-0.08158	-0.12322	-0.31294	0.53341	0.34880	-0.01197	-0.09770	-0.00872	-0.08512	0.05803	-0.20120	-0.10503	-0.21100	-0.15233	-0.31950	-0.16954	-0.08877	-0.04049
dance		-0.80220	-1.81242	-0.79670	-0.77157	-0.88810	-0.87010	-0.47319	-0.41128	-0.60005	-0.31443	-0.14386	-0.88269	-0.69623	-0.42890	-0.51877	-0.49121	-0.24781	-0.30226	-0.30539
sing		-0.80245		-0.83542	-0.44393	-0.31720	-0.91086	-0.90159	-0.47199	-0.51072	-0.35792	-0.55070	-0.60131	-0.41285	-0.49237	-0.43667	-0.56703	-0.16824	-0.07850	-0.07125
cook																				
talk		-0.80293	-0.06963	-0.01803	-0.15136	-0.05952	-0.24158	-0.10488	-0.05478	-0.02660	-0.04566	-0.08018	-0.09050	-0.02186	-0.23161	-0.18054	-0.16924	-0.17004	-0.12223	-0.04177
think	-0.20513	0.23247	0.27430	0.05744	-0.25033	-0.18858	-0.19447	-0.20565	-0.08707	-0.07425	-0.09745	-0.12104	-0.17167	-0.16782	-0.15554	-0.15784	-0.10696	-0.26780	-0.28363	-0.31546
complain	-0.58511	-0.50152	-0.76957	-0.11501	-0.24211	-0.58015	-0.36647	-0.55573	-0.49684	-0.25750	-0.07211	-0.07607	-0.30476	-0.04457	0.01057	-0.18866	-0.40822	-0.22150	-0.10266	-0.43982
cry			-0.29145		-0.59507	-0.99985	-0.48036	-0.39402	-0.67713	-0.35308	-0.53014	-0.73688	-0.82933	-0.79365	-0.61355	-0.63787	-0.87217	-0.61833	-0.72297	-0.89788
throw		-0.23805	-0.42646	-0.70911	-0.01133	0.32261	-0.06511	-0.02584	0.26268	0.27605	0.14137	-0.02435	-0.04133	-0.15294	0.06859	0.42730	0.12009	-0.05074	-0.17527	-0.04901
scream												-0.46600	-0.94444	-0.36683	-0.35743	-0.65713	-0.89350	-0.62809	-1.04066	-1.19763
shout																				
say	0.23027	0.14092	0.15566	0.03448	0.04005	0.03160	0.01281	-0.07213	-0.05338	-0.05556	-0.08221	-0.03813	0.02353	-0.01471	-0.19499	-0.08447	-0.01451	-0.07868	-0.00402	-0.01307
speak	0.01786	0.08300	-0.06991	-0.14987	-0.14141	-0.30627	-0.32301	-0.47527	-0.22166	-0.25340	-0.27785	-0.15975	-0.22406	-0.18136	-0.32631	-0.20370	-0.20869	-0.11734	-0.13143	-0.14582
work					0.07072	0.11676	-0.09991	0.63022	0.59294	0.72199	0.47223	0.49479	0.16308	0.03844	0.23657	-0.02718	-0.11364	-0.13382	0.01101	-0.11065
consider	0.01783	-0.75834	-0.31921	-0.01118	0.26809	-0.04519	0.24806	0.23403	0.11869	0.10124	0.14139	0.15106	0.16867	0.13762	0.06743	0.22754	0.12647	0.17805	0.00758	0.02919
walk				0.25338	0.11139	0.50493	0.17665	0.10089	0.38599	0.10415	0.03598	0.16252	0.37184	0.12179	-0.17938	0.03404	-0.10940	-0.10845	0.03774	-0.20568
give	-0.01047	-0.03928	0.06809	-0.01399	-0.06633	0.00087	-0.01094	-0.01444	0.04966	0.02302	0.03521	0.08786	-0.07377	-0.01082	-0.04794	-0.07590	-0.05913	-0.11773	-0.08733	-0.11203
read			0.23164	-0.33288	0.29405	0.10009	-0.17135	-0.13439	-0.08919	-0.15655	-0.00124	0.16624	0.09215	0.10051	0.12515	0.34495	-0.08761	0.16146	0.12473	-0.08353
run		-0.20001	0.40793	0.29355	0.37423	0.10916	-0.00955	0.18527	-0.16616	0.39274	0.00048	0.09850	0.25708	0.01247	0.44551	0.58469	0.56260	0.39244	0.42340	0.36124
watch				-0.73850		-0.56039	-0.54239	-0.57004	-0.48368	-0.49264	-0.46757	-0.46610	-0.46734	-0.49175	-0.53364	-0.61939	-0.55107	-0.39762	-0.32455	-0.30540
write	-0.23255	-0.10311	-0.43360	0.29810	0.32624	-0.35245	-0.19576	-0.14327	-0.45816	-0.29189	-0.09356	0.03542	0.11365	0.01569	0.07097	0.21634	0.11829	0.10037	0.14827	0.19204
play			0.27298	-0.37947	-0.34344	-0.39149	-0.10560	-0.41709	0.02508	0.00930	0.11440	-0.04422	0.24685	0.00893	-0.18660	0.00921	0.20471	0.04537	0.15293	0.06702
take	0.43488	0.06368	0.02670	0.00460	0.01495	0.13376	0.02882	0.17192	0.05920	0.04869	0.11935	0.06014	0.10166	0.14801	0.11519	0.08029	0.08328	0.06200	0.08391	0.12763
look	0.01783	0.19820	-0.42429	-0.62256	-0.32659	-0.56847	-0.48833	-0.45113	-0.28569	-0.54605	-0.35887	-0.49942	-0.32025	-0.29040	-0.48947	-0.36726	-0.41391	-0.35233	-0.23909	-0.28616
start					-0.54924			0.35967	0.36174	0.06566	-0.02818	0.16614	-0.06153	-0.21302	0.36059	0.16698	0.30048	0.28387	-0.00570	-0.14016
bring	-0.04029		-0.00083	0.05027	0.27621	-0.02017	-0.02367	0.18357	0.01386	0.03242	0.04108	0.05259	-0.09749	0.10577	-0.00940	0.19634	-0.02810	0.04286	0.11922	0.04881
help					0.64830	-0.16216	-0.00857	0.12040	-0.34027	-0.22458	-0.05374	0.02935	-0.04943	-0.09383	-0.33102	-0.10084	-0.16965	-0.07264	0.08490	0.03761
try			-0.76909		-0.67434	-0.56046	-0.54278	-0.57028	-0.48410	-0.49305	-0.46820	-0.46661	-0.46786	-0.49253	-0.53442	-0.61999	-0.55179	-0.39817	-0.32505	-0.30586
wait	0.39432		0.28893	0.63060	0.17829	-0.11616	0.00191	-0.07382	0.05304	0.03676	-0.01409	0.12437	-0.01380	-0.11067	-0.12111	0.11039	-0.07397	0.02770	0.02132	0.02341
kill		0.25920			0.80380	0.48183	0.88457	1.06444	0.80072	0.74780	0.87799	0.59386	0.36167	0.86585	1.15274	0.44330	0.24231	0.40037	0.38146	0.31589
listen	-0.14558	-0.30814	-0.42477	-0.23957	-0.22707	-0.25916	-0.30517	-0.22724	-0.13220	-0.17867	-0.20627	-0.21140	-0.09256	-0.11321	-0.12996	-0.19341	-0.15252	0.00374	0.07777	0.09785
ask	-0.47781	-0.20065	-0.43844	-0.29864	-0.34553	-0.22784	-0.09285	-0.27729	-0.15601	-0.29527	-0.11239	-0.03148	-0.20961	-0.17223	-0.19371	-0.11161	0.00415	-0.09346	0.00357	-0.03516
carry	-0.28378	0.19837		0.93454	0.35757	0.50086	0.07331	0.12918	0.33802	0.54319	0.62587	-0.00110	0.08803	0.35933	0.05075	0.30797	0.43465	0.37569	-0.03786	0.18013
share			-0.99089	0.01517		-0.16213	0.34258	0.03219	0.08386	0.11786	-0.24049	-0.46636	0.00986	-0.26418	-0.05643	-0.02874	-0.13450	-0.26280	-0.13350	-0.05552
choose		0.32393	0.42219	0.12449	-0.22212	-0.22638	0.03189	-0.04165	0.24784	0.06273	0.32376	0.20313	-0.07452	0.00892	0.29775	0.12978	0.16750	0.00045	-0.12603	-0.11912

Table 27. Summary table for the *can* construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink					0.50458	0.60462	0.13344	0.08781	0.45429	0.14507	0.04244	0.11340	0.27490	0.00550	-0.10255	-0.09745	0.01462	0.07088	0.01828	0.07690
eat	-0.21459	-0.36085	-0.03459	-0.14119	-0.21102	-0.08664	-0.24094	-0.32630	-0.08281	-0.11799	-0.06136	-0.17639	-0.09324	-0.13126	-0.27254	-0.05514	0.01652	-0.03254	-0.02087	-0.02347
dance	-0.67759	-0.40692	-0.78495	-0.95563	-0.60354	-0.32622	-0.59109	-0.70674	-0.46147	-0.62292	-0.65361	-0.97126	-0.78710	-0.70589	-0.74653	-0.81206	-0.57774	-0.55787	-0.23581	-0.12530
sing	-0.32435	-0.54155	-0.61652	-0.72384	-0.48044	-0.51367	-0.57070	-0.56865	-0.49192	-0.43314	-0.46327	-0.47530	-0.37923	-0.38182	-0.38735	-0.53521	-0.36162	-0.28783	-0.12734	-0.15050
cook					-0.90419	-0.83542	-1.01106	-1.20886	-1.03731	-1.30594	-1.15827	-1.00315	-1.17011	-1.04246	-1.11712	-1.05101	-0.85391	-0.68189	-0.58942	-0.51019
talk	-0.02285	-0.37216	-0.18203	-0.41531	-0.26826	-0.39322	-0.31408	-0.31366	-0.31424	-0.16477	-0.19763	-0.11204	-0.07728	-0.13704	-0.17939	-0.22884	-0.23144	-0.10781	-0.08288	-0.07913
think		0.21181	0.00490	0.01975	-0.21149	-0.26872	-0.07650	-0.16522	-0.06413	-0.06206	0.05261	-0.02643	0.01877	-0.05188	0.01167	0.06975	-0.01585	-0.05219	-0.00166	0.02029
complain					0.07832	0.13700	-0.15856	0.46958		0.92721		0.19095	-0.09500	0.13314	0.73790	0.07678	0.12313	0.11793	0.11173	0.01884
cry				0.39860	-0.54120	-1.00751	-0.93723	-0.76128	-0.30595	-0.83367	-0.66897	-0.75135	-0.61873	-0.63343	-0.64707	-0.34829	-0.16009	-0.17537	-0.37434	-0.18123
throw	0.13709	-0.72984	-0.28243	0.03537	0.07847	-0.37299	-0.07715	-0.21969	-0.23921	-0.13384	0.04248	-0.07258	-0.11156	0.10028	0.02429	-0.15058	0.10081	0.17279	0.12758	0.15283
scream						-0.26100					-0.91186	-0.45008	-0.31250	-1.01309	-0.85996	-0.82160	-0.78009	-0.11732	-0.08432	-0.41742
shout											0.36119			0.42042	-0.11949	0.09797	-0.00662	0.48480	0.42298	0.12515
say	0.29016	0.08151	0.19300	0.11266	0.02330	0.09446	0.07829	0.17593	0.01171	0.08025	-0.00424	0.03478	-0.00719	-0.00260	-0.05704	-0.00576	-0.01265	0.05726	0.05826	0.06400
speak	0.36671	0.10148	-0.14889	0.12205	0.10746	0.00172	-0.02739	-0.02364	-0.04087	-0.01958	-0.10824	-0.03356	-0.07644	0.03364	0.02229	0.08368	0.05912	0.06223	0.02348	0.07446
work	-0.31399	0.12930	-0.02249	0.10639	0.13543	0.48971	0.34421	0.15736	0.23024	0.06292	0.13186	0.05798	0.12277	0.07160	-0.00248	0.02471	-0.10677	-0.08396	-0.02004	-0.08494
consider	-0.27830	-0.55359	-1.10652	0.39860	-0.54886		-0.32134		1.04314	0.53043	0.38569	0.42927	0.12411	0.16540	0.14007	0.02500	0.15825	-0.11740	-0.10980	-0.12457
walk	-0.04772	-0.45906	0.28074	-0.25453	-0.35154	-0.18019	-0.09792	-0.12025	-0.06209	-0.24695	-0.07108	-0.09822	-0.06538	-0.00262	-0.03758	-0.09712	-0.10807	-0.07034	-0.08593	0.00512
give	-0.04406	-0.17833	0.05839	-0.08162	-0.00520	0.08863	0.01905	0.02338	-0.03603	-0.01101	-0.04100	-0.07609	-0.00974	-0.08897	-0.14196	-0.19507	-0.15608	-0.01279	0.04029	0.02605
read	-0.25499	0.06829	-0.21086	-0.15582	-0.22117	-0.29945	-0.14816	-0.09812	0.03974	-0.07798	-0.05918	0.05515	0.12279	0.04205	0.05171	0.05914	-0.00453	-0.04688	-0.08296	-0.07271
run	-0.90005	-0.73055	-0.43526	-0.40311	-0.18005	-0.22886	0.00814	0.02786	-0.05908	-0.01756	0.04519	-0.03089	-0.09794	-0.01805	-0.01794	-0.15923	0.04203	0.01676	0.02137	0.05832
watch		-1.25186	-1.37903	-0.80618	-0.62872	-0.34021	-0.35058	-0.14033	-0.33036	-0.35820	-0.15033	-0.05221	-0.32265	-0.10643	-0.17200	-0.11299	-0.04764	-0.11756	-0.02855	0.00041
write	0.34640	-0.13816	-0.13019	-0.22947	0.14442	0.21702	0.01374	0.21154	-0.09732	-0.05803	0.02975	0.02329	0.06100	-0.02489	0.05690	0.07746	0.11476	0.06359	0.02835	0.03545
play	-0.07403	-0.82756	-0.33803	-0.33174	-0.38314	-0.28849	-0.22437	-0.18451	-0.24966	-0.29958	-0.19054	-0.17809	-0.14097	-0.17418	-0.18064	-0.17445	-0.08882	0.03334	0.10987	0.13251
take	0.07605	0.13285	0.07643	0.02120	0.02216	0.07417	0.00629	-0.03582	-0.01833	-0.01355	-0.03074	-0.10452	-0.05957	-0.07014	-0.11550	-0.13272	-0.08928	-0.04583	-0.00375	-0.01495
look	0.49491	0.38356	-0.01384	0.00055	0.03874	0.05272	-0.06837	-0.01974	-0.03482	-0.05978	-0.08366	-0.06546	0.00837	-0.09032	-0.12253	-0.15214	-0.07566	-0.05231	0.00379	-0.03287
start				-0.64759	-0.09780	0.34114	-0.00726	0.09434	-0.17623	0.07522	0.01982	0.12129	0.06462	-0.07788	-0.16558	-0.02120	-0.03120	-0.00676	-0.03664	-0.12710
bring	0.04750	0.41736	-0.03658	-0.03008	0.17045	-0.12628	-0.04774	0.26094	-0.04514	-0.07539	-0.05987	0.03890	-0.15889	-0.08282	-0.05829	-0.04607	-0.07806	0.13958	0.04746	0.07067
help	-0.02287	-0.30830	-0.21601	-0.18211	-0.12246	-0.06060	-0.18691	-0.23472	-0.05321	-0.18353	-0.18898	-0.19535	-0.15726	-0.36031	-0.30598	-0.32017	-0.27278	-0.16060	-0.13383	-0.12953
try			-0.05651	-0.35857	0.57329	0.56306	-0.20437	0.26714	-0.09955	-0.38245	0.04002	0.03731	0.16952	-0.14888	-0.06528	0.08356	0.11931	-0.03728	0.06240	-0.07355
wait			-0.53942	-0.21849	-0.24388	-0.34487	-0.14353	-0.17725	-0.22607	-0.36019	-0.33687	-0.14748	-0.26387	-0.28161	-0.16365	-0.46464	-0.16903	-0.09639	0.07448	0.03707
kill		-0.17223			0.54354	0.19148	0.31890	0.43315	0.38357	0.42413	0.23585	0.63945	0.65400	0.58663	0.12686	0.09695	0.07112	0.25809	0.32328	0.37147
listen	-0.97749			-0.20376		-0.03915	0.51826	-0.23827	0.06554	-0.46861	-0.10188	-0.26935	-0.10376	0.24441	0.01915	0.02834	-0.11310	-0.01944	-0.13966	-0.10600
ask		-0.41476	0.06635	-0.08726	-0.31197	-0.25576	-0.12046	-0.33913	-0.20548	-0.08530	0.10513	-0.14789	-0.09519	-0.00385	-0.17800	-0.01883	-0.11151	-0.14058	-0.18741	-0.20221
carry		0.23829	0.11474	0.20991	-0.01426	0.01311	0.00308	-0.14536	-0.09321	-0.06798	-0.17283	-0.00945	-0.07319	-0.06874	-0.14393	-0.15123	-0.03975	-0.08545	0.00605	-0.02643
share				0.41980	-0.41754	0.34114	0.67092	-0.12424	-0.65247	-0.19413	-0.01400	-0.25839	-0.14296	-0.15386	0.09813	-0.11532	-0.22091	-0.23700	-0.22595	-0.20883
choose				0.20490	0.28205	-0.10984	-0.24360	-0.17346	-0.10525	0.12590	-0.05712	0.18085	-0.05429	0.05266	0.08358	0.21430	0.02788	-0.05186	-0.19727	-0.18105

Table 28. Summary table for the *cannot* construction.*

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink					0.66381	0.38389		-0.29885	0.52011	0.33678	0.51871	-0.12074	0.71969	0.15182	-0.26437	0.24365	-0.04271	-0.18195	-0.09599	0.25471
eat	-0.66962		0.47764	-0.77306	-0.84915	-0.56407	-0.44980	-0.27206	0.07134	-0.13790	-0.03306	-0.10707	-0.10519	-0.05838	-0.31788	-0.15671	-0.25137	-0.09902	-0.12743	-0.26140
dance				-0.94087	-0.85563	-0.96897	-0.49009	0.15827	-0.42180		0.08789	-0.99437	-0.19217	-0.47175	-0.40441	-0.64407	-0.46075	-0.21450	0.28025	0.09988
sing			-0.01563	-1.24476	-0.62580	-0.68732	-0.70494	-0.43592	-0.47948	-0.52111	-0.58267	-0.31054	-0.32533	-0.64434	-0.62890	-0.54779	-0.61095	-0.29718	-0.20692	-0.20697
cook								-1.27326	-0.89896		-0.95982		-1.42549	-1.04401	-1.96107	-1.08342	-1.23911	-1.44516	-1.01966	-0.62222
talk			0.39212	-0.75250	-0.26079	-0.27604	-0.21730	-0.53258	-0.32915	-0.35532	-0.16801	-0.03462	-0.23528	-0.21390	-0.17141	-0.09037	-0.10146	-0.09454	-0.24731	-0.24012
think	-0.39323	-0.08982	0.08990	0.33051	0.12655	0.04253	-0.08740	0.34193	-0.00157	0.14242	0.08749	-0.01932	0.23995	-0.04350	-0.20689	-0.06354	-0.00859	-0.10211	-0.18697	-0.13702
complain		-1.04036	-0.00609	0.12043	0.88591	0.96434	0.54225	0.19996	0.69979	0.27406	0.08733	0.48172	0.22277	0.42423	0.26226	0.15908	0.21758	0.31132	0.55021	0.22591
cry						-0.31528		-0.01785	-1.29714				-0.16984		-0.06851	-0.33064	-0.15959	-0.13529	-0.47011	0.07354
throw	-0.36785	-0.49190		-0.24172	0.46738	-0.09349	0.16345	0.36563	0.29076	1.14524	0.20343	0.23756	0.62637	0.04512	0.55016	0.46748	0.35803	0.45785	0.47227	0.19413
scream																				
shout																				
say	0.27734	-0.06590	-0.09369	0.05211	0.01469	0.17855	0.09006	0.17685	0.16131	0.20567	-0.01461	0.11829	0.20078	0.18978	0.04544	0.04250	0.14471	0.00934	0.01084	-0.02787
speak	-0.09680	-0.19200	-0.29632	-0.12126	-0.28912	-0.13595	-0.11573	-0.13457	-0.28046	-0.09620	-0.12676	-0.18862	-0.21636	-0.15471	-0.18223	-0.15411	-0.21185	-0.12488	-0.17487	-0.16505
work			0.44542	0.66208	0.65128	0.66393	-0.03409	-0.09732	-0.11850	-0.19564	0.05065	-0.11045	0.02689	-0.20119	-0.05584	-0.00383	-0.15558	-0.17195	-0.25148	-0.24246
consider		-0.49190	-0.33333	-0.40933	-0.51286	-0.31528	-0.11434	0.70936		0.89837	-0.04576	0.17278	-0.18840	-0.06141	0.03506	0.18327	-0.05315	0.17006	0.32612	-0.05780
walk	-0.06650	-0.24199	-0.50644	-0.24752	-0.18668	-0.34437	-0.31142	-0.30820	-0.16663	-0.22330	-0.28932	-0.42342	-0.28828	-0.20818	-0.18916	-0.23511	-0.27299	-0.27227	-0.17999	-0.21445
give	0.13478	0.17939	-0.13538	-0.03283	-0.02676	-0.00914	-0.01180	0.02938	-0.02221	-0.15051	-0.10563	-0.09084	-0.06328	-0.05317	-0.05418	-0.18877	-0.22754	-0.11261	-0.10805	0.02526
read	0.10990	0.49471	0.14442	-0.17858	-0.05362	-0.03210	-0.23630	0.11055	0.09081	-0.13895	-0.24359	0.03317	0.06452	0.04465	-0.10407	-0.06778	-0.08051	-0.03040	-0.07335	-0.13536
run		-0.42491	-0.49370	-0.18736	0.43531	0.45536	-0.09205	-0.30260	-0.03854	0.09950	-0.10294	-0.03774	0.40841	0.24365	0.10999	-0.05630	-0.10605	-0.09690	-0.10060	-0.07596
watch									-0.47291	0.14484	0.60128		-0.26007	0.20191		-0.14096	0.22745	-0.03834	0.04195	-0.17654
write		-0.52269	-0.29865	-0.17207	0.05163	0.26701	-0.12522	0.06954	-0.07730	0.09098	0.00811	0.01830	0.11474	-0.06880	0.10346	0.08078	0.05487	0.00118	-0.12469	-0.01089
play			-0.01564	-0.21357	-0.69647	-0.23856	-0.40944	0.64367	-0.01635	-0.39291	-0.17835	-0.13981	0.16357	-0.07885	-0.22204	-0.03354	-0.04239	-0.08697	-0.05245	-0.05992
take	-0.02526	-0.49314	-0.17012	0.04739	0.03815	0.19913	-0.02193	0.05972	0.08952	0.09161	0.05637	0.03208	-0.01529	-0.09118	-0.03901	-0.16750	-0.14455	-0.01182	-0.09686	-0.07234
look	0.24648	0.21789	0.27824	-0.00595	0.04581	0.10691	-0.39240	-0.24358	-0.22942	0.01516	-0.14312	-0.17201	-0.34606	0.03215	0.02173	-0.21546	-0.05435	-0.11987	-0.07411	-0.22553
start										0.53307	0.83688	0.39308		0.28556	0.00320	0.40331	0.45158	0.29753	-0.15123	0.07135
bring	0.04772	0.31106	0.11704	0.28347	-0.00153	0.12700	-0.11085	0.07016	-0.06239	0.08624	0.06678	-0.13226	-0.06484	-0.26034	-0.20494	-0.11771	-0.16429	-0.08462	-0.07805	0.06513
help	-0.23006	0.08296	-0.21589	-0.21844	-0.20029	-0.07751	-0.18233	-0.07028	-0.08484	-0.07486	-0.10958	-0.01768	0.00142	-0.06530	-0.11037	-0.09336	-0.11855	-0.04796	-0.03038	-0.06659
try																				
wait				-0.05118	0.07816	-0.51498	-0.06162	0.10727	-0.21411	-0.11072	-0.03868	-0.25705	-0.34481	-0.42365	-0.21753	-0.02374	-0.00626	-0.15677	-0.11551	-0.13800
kill	-0.76598				0.16626				0.86982	0.54292	0.69023	0.66530	0.52218	0.52245	0.12351	0.20318	0.20468	0.20788	0.31758	0.19162
listen				-0.84437	-0.80986	-0.49138	-0.48999	-0.01785	-0.59791	-0.37590	0.10638	-0.49650	-0.04880	-0.34974	0.27255	0.09543	0.42413	-0.07055	-0.12642	-0.35909
ask				0.33848	0.20580	0.55033	0.26987	0.55334	0.09864	0.51324	-0.00376	0.06944	-0.17000	-0.24687	-0.00636	-0.09657	-0.11629	0.14164	-0.05437	-0.24302
carry		-0.27006	-0.05168	0.45237	0.19021	0.10086	0.00585	-0.02500	0.42725	0.13847	-0.18479	0.00652	0.02912	-0.22071	-0.00873	-0.24550	-0.27506	-0.28148	-0.13864	-0.07623
share				-0.25946	0.06586	-0.49179	-0.31420	-0.04273	-0.24119	-0.43100	-0.24372	-0.18665	-0.23145	-0.12797	-0.23613	-0.07497	0.02492	-0.15415	-0.05970	-0.06260
choose	-0.78946	-0.70609	-0.61891	-0.49711	-0.47263	-0.41583	-0.44965	-0.55988	-0.20462	-0.31296	-0.03144	0.10008	-0.18664	-0.19110	-0.02974	-0.00896	0.01947	-0.08203	0.00096	-0.13507

* The form *can not* – written as two separate words – was not queried.

Table 29. Summary table for the *could* construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink		0.66969	0.62650	0.52370	0.67266	0.82259	0.14927	0.32107	0.17489	0.06301	0.14063	0.23880	0.24779	0.14975	0.03605	0.05348	-0.04544	0.05237	0.14842	0.10161
eat	-0.35519	-0.31342	-0.35080	-0.03879	-0.20469	-0.01906	0.06060	0.03775	0.05468	0.12106	0.02084	0.10615	0.00956	0.02794	-0.05914	-0.11236	-0.09135	-0.06843	-0.00806	0.01131
dance	-1.20062	-1.18932	-0.60172	-0.43934	-0.24318	-0.59993	-0.76270	-0.38503	-0.23744	-0.56996	-0.42631	-0.54251	-0.46479	-0.58799	-0.45941	-0.53953	-0.55827	-0.41431	-0.20610	-0.32226
sing	-0.95770	-0.41059	-0.22596	-0.02956	-0.28323	-0.22639	-0.35914	-0.27069	-0.30287	-0.35420	-0.29892	-0.20300	-0.25097	-0.25826	-0.30230	-0.33572	-0.30614	-0.17977	-0.15162	-0.16734
cook				-0.11691	-0.70261	-0.54596	-0.89683	-0.74881	-0.68468	-0.63492	-0.76169	-0.62992	-0.63288	-0.77462	-0.86088	-0.88566	-0.74225	-0.60336	-0.47091	-0.47366
talk	-0.13253	-0.17282	-0.17303	-0.18654	-0.21153	-0.12601	-0.08696	-0.10600	-0.13401	-0.07659	-0.08995	-0.05296	-0.04609	-0.05861	-0.07030	-0.12744	-0.14677	-0.10005	-0.04027	-0.06071
think	-0.15610	-0.35259	-0.27887	-0.16002	-0.24475	-0.22294	-0.22824	-0.27395	-0.20746	-0.22717	-0.15330	-0.11633	-0.11103	-0.12606	-0.09747	-0.11726	-0.10066	-0.15220	-0.13300	-0.11927
complain		-0.61425	-0.33747	-0.41337	-0.36515	-0.28266	-0.27339	-0.07350	-0.65756	0.06515	0.10352	0.38084	-0.11618	-0.21915	-0.29526	-0.14147	-0.26645	-0.28563	-0.20468	-0.13767
cry			-0.38470	0.30914	-0.70002	-0.67109	-0.88756	-0.47355	-0.56467	-0.62492	-0.81233	-0.62832	-0.70728	-0.65240	-0.52267	-0.48690	-0.64399	-0.44253	-0.44215	-0.45699
throw		0.13436	0.16376	-0.16487	0.10989	-0.07485	0.03136	-0.01174	-0.08990	0.05158	0.06165	0.20936	0.07059	0.06235	0.07643	0.04280	0.04096	0.14608	0.09245	0.09627
scream				-1.61542	-0.83430	-0.76784	-0.77174	-0.93088	-0.55208	-0.90853	-1.10072	-0.45951	-0.80272	-1.16984	-0.97566	-0.81184	-0.88648	-0.65487	-0.59472	-0.60836
shout			-0.38468	-1.26312			0.32872	-0.40256	0.05260	0.23023	0.83056	0.64149	0.24790	0.31535	0.11624	-0.07997	0.04803	0.18976	-0.00350	0.04110
say	-0.06781	-0.11613	-0.13300	-0.10412	-0.02587	-0.11222	-0.10331	-0.15286	-0.13778	-0.13551	-0.09745	-0.04351	-0.08088	-0.07435	-0.05445	-0.05981	-0.05047	-0.08269	-0.04786	-0.05721
speak	-0.39109	-0.29286	-0.15310	-0.19014	-0.16675	-0.16957	-0.18717	-0.21084	-0.17087	-0.14426	-0.12961	-0.06272	-0.10691	-0.08525	-0.10566	-0.09581	-0.04545	-0.03034	-0.02204	-0.03028
work	0.01390	0.15455	-0.01360	-0.28115	-0.05849	-0.09070	0.12022	0.15968	0.05222	0.15930	0.19177	0.19306	0.16501	0.15261	0.11389	0.10491	0.05878	0.03583	0.05868	0.05188
consider	0.21031	0.21867	0.35583	-0.04119	0.06097	-0.19743	-0.07148	-0.04354	-0.22119	-0.22961	-0.12201	0.10785	-0.09938	0.25053	0.07342	0.18505	0.14766	0.05696	0.00691	-0.08266
walk	-0.33471	-0.65598	-0.34625	-0.35935	-0.25070	-0.24329	-0.26613	-0.26609	-0.16329	-0.20752	-0.11321	-0.13414	-0.04528	-0.10071	-0.16818	-0.10183	-0.11946	-0.07784	-0.04749	-0.02350
give	-0.07622	-0.02089	0.06862	0.03049	-0.00037	-0.03324	0.03440	0.02330	-0.00711	-0.00273	0.00769	-0.05905	-0.00990	-0.02264	-0.06067	-0.05982	-0.03317	-0.00930	0.01008	0.00730
read	-0.09413	-0.17884	-0.31020	-0.28819	-0.24828	-0.21981	-0.12924	-0.15990	-0.16627	-0.12855	-0.08417	-0.07343	-0.06822	-0.05347	-0.08691	-0.07733	-0.07523	-0.03956	-0.03310	-0.04333
run	-0.75293	-0.23414	0.02114	0.04105	-0.07940	-0.03442	-0.19515	-0.03315	-0.00065	0.01021	-0.00470	-0.01757	-0.07447	0.01755	-0.06612	-0.08889	-0.03027	-0.02967	0.01911	-0.00680
watch		0.28931	0.24378	-0.16470	-0.18417	0.01663	-0.00799	-0.11074	-0.09649	0.00744	-0.04089	-0.07997	-0.15316	-0.10485	-0.08890	-0.11457	-0.07900	-0.00068	-0.00739	0.06065
write	0.21756	0.10908	0.03165	0.15250	0.18051	0.13465	0.10488	0.14170	0.11585	0.20269	0.26968	0.19563	0.19488	0.16756	0.17506	0.20855	0.10052	0.16412	0.11421	0.11253
play	0.16899	0.91306	-0.00066	-0.16315	-0.03579	-0.25659	-0.14227	-0.08952	-0.19819	-0.07370	-0.03981	-0.04578	-0.03047	-0.02088	-0.05423	-0.04522	-0.02679	0.15914	0.22209	0.21034
take	0.05168	-0.01397	0.11239	0.05168	0.04896	0.06446	0.03934	0.03414	0.07464	0.08031	0.04661	0.06066	0.05360	-0.00434	0.02292	0.00589	-0.00245	0.02054	0.02300	0.02411
look	0.01394	-0.08489	-0.01952	-0.21236	-0.23273	-0.21246	-0.18163	-0.18415	-0.17588	-0.18143	-0.13684	-0.14252	-0.12267	-0.14444	-0.11195	-0.11007	-0.11586	-0.06008	-0.02306	-0.02506
start					0.45593	0.17024	0.20382	-0.11706	0.08482	0.17592	0.19188	0.11619	0.40741	0.10224	0.05882	0.03225	-0.01916	-0.06003	-0.03369	-0.02208
bring	0.11702	0.16177	0.09997	0.10600	0.04437	0.09909	0.12384	0.10961	0.11470	0.05688	0.00443	0.00749	0.02273	0.07465	0.01520	0.05272	0.07560	0.07727	0.09348	0.04281
help	0.66092	-0.12297	-0.09489	-0.14876	-0.26829	-0.14909	-0.21068	-0.20411	-0.11797	-0.14700	-0.14120	-0.09641	-0.12314	-0.13832	-0.16379	-0.13561	-0.14745	-0.05158	-0.02894	-0.02985
try		-0.16686	-0.20859	0.36028	-0.36252	-0.06873	-0.03928	0.13637	0.09114	-0.29856	-0.01347	0.06609	-0.00606	0.13196	0.02734	0.03711	0.01753	-0.05138	-0.00126	-0.03702
wait	-0.13230	-0.46808	-0.06593	0.44471	0.07571	-0.29185	-0.22060	-0.16171	-0.07022	-0.18553	-0.12772	-0.15242	-0.08991	-0.10846	-0.08136	-0.06744	-0.03978	-0.02966	0.05763	0.01227
kill	-0.70680	-0.10887	0.29675	0.54339	0.57846	0.44716	0.42297	0.21192	0.44655	0.35219	0.36989	0.39312	0.21173	0.37651	0.35522	0.26462	0.16836	0.28097	0.33506	0.29725
listen	0.07191	-0.79890	-0.25447	-0.50170	-0.33011	-0.29468	-0.31099	-0.41562	-0.26812	-0.27443	-0.28078	-0.11664	-0.27799	-0.08045	-0.12086	-0.12527	-0.23967	-0.11502	-0.01239	-0.02059
ask		-0.18286	-0.02297	-0.32577	-0.29904	-0.23346	-0.35591	-0.35339	-0.20932	-0.31231	-0.19456	-0.18727	-0.15985	-0.22026	-0.18574	-0.12401	-0.18776	-0.17550	-0.19046	-0.18013
carry	-0.15361	-0.15528	0.05417	-0.03693	0.00799	0.16991	-0.01926	0.10196	0.09935	0.10882	0.07628	0.12086	0.09040	0.06522	0.01060	-0.03717	0.01039	0.06399	0.09663	0.11421
share	-0.22924	-0.23379	-0.68593		-0.40782	-0.22848	-0.40738	-0.40272	-0.33000	-0.31637	0.04646	-0.43084	-0.23881	-0.20253	-0.27046	-0.16018	-0.19291	-0.25453	-0.14844	-0.13000
choose	-0.19463	0.68743	-0.48516	0.05923	-0.44313	-0.03927	-0.18711	-0.05580	-0.19213	-0.20553	-0.00920	-0.00667	-0.05328	0.14191	0.08677	0.05040	-0.05892	-0.06970	-0.19729	-0.08674

Table 30. Summary table for the *could not* construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink			-0.08305	-0.65639	-0.37743	0.54171	-0.00596	0.21851	0.08568	0.05639	0.34438	0.17943	-0.04860	0.21257	-0.19466	-0.08873	-0.12004	0.12296	0.05490	-0.03933
eat	-0.54327	-0.35418	-0.06113	-0.09692	-0.19875	-0.16686	-0.17942	-0.13681	-0.15565	-0.10250	-0.09581	-0.13091	-0.14632	-0.10789	-0.15424	-0.28540	-0.22311	-0.16734	-0.18732	-0.13075
dance	-1.11621	-0.84884	-0.70647	-0.26624	-0.63048	-0.37949	-0.44982	-0.42637	-0.17702	-0.32382	-0.27080	-0.34709	-0.25629	-0.40795	-0.23687	-0.48394	-0.44140	-0.46671	-0.33505	-0.33969
sing	-0.63900	-0.54745	-0.83842	-0.42947	-0.41023	-0.76020	-0.55441	-0.42158	-0.53232	-0.33185	-0.24250	-0.34283	-0.28448	-0.32784	-0.33969	-0.28167	-0.33027	-0.29414	-0.30066	-0.14566
cook								-0.98256	-0.26570	-0.73395	-1.41169	-0.63579	-1.78269	-1.36435	-1.39786	-1.16455	-0.56713	-0.71506	-0.78058	-0.76566
talk	-0.69107	-0.37159	-0.07835	-0.24410	-0.22536	-0.25592	-0.20359	-0.28496	-0.10316	0.11128	-0.10110	-0.03400	-0.07802	-0.15036	-0.16584	-0.20978	-0.19455	-0.14115	-0.22815	-0.06845
think	-0.25076	-0.29292	-0.06912	-0.21568	-0.16041	-0.19565	-0.15320	-0.18925	-0.17251	-0.15450	-0.15712	-0.14638	-0.17907	-0.19461	-0.15385	-0.15403	-0.14133	-0.12353	-0.15704	-0.15251
complain	-0.11575	0.02685	0.62835	-0.01625	0.32401	-0.30983	-0.03823	0.02229	0.12002	-0.04448	-0.09511	0.09038	-0.10045	0.15370	0.10953	-0.08071	-0.01705	-0.07482	-0.01074	0.08560
cry			-1.53051	-0.47228	-0.78544	-0.89252	-1.42765	-0.55485	-0.61269	-0.74322	-0.56523	-0.69843	-0.56791	-0.54340	-0.63685	-0.69465	-0.66240	-0.60296	-0.45887	-0.39248
throw		-0.03533	0.35838	-0.08477	0.00867	-0.13330	-0.15631	-0.17820	0.23509	0.16765	0.15794	0.07938	-0.08647	-0.22753	0.03961	-0.03754	0.02162	-0.20523	0.19120	0.02734
scream				-0.61833	-0.86450	-0.75962		-0.50544	-0.69710	-0.69502	-0.78613	-1.12650	-1.03456	-0.44488	-1.08700	-1.08881	-1.26944	-0.76969	-0.68504	-0.75910
shout																				
say	-0.01732	-0.12517	-0.06059	-0.06702	-0.13106	-0.10953	-0.05722	-0.06596	-0.15066	-0.09613	-0.08650	-0.03048	-0.05882	-0.04822	-0.09190	-0.01362	-0.00051	-0.00412	-0.03526	-0.06321
speak	-0.10288	-0.22336	-0.24326	-0.15105	-0.23237	-0.19544	-0.15568	-0.24811	-0.24343	-0.24145	-0.19096	-0.16478	-0.17673	-0.15751	-0.13097	-0.20522	-0.14966	-0.15362	-0.09163	-0.10653
work			0.39029	-0.19573	-0.02788	0.02309	0.33449	0.20590	0.40641	0.20843	0.31123	0.32746	0.17461	0.27807	0.17864	0.22636	0.09379	0.14017	0.03616	-0.01773
consider	-0.29219	-0.25343	-0.05984	-0.11956	-0.27648	0.44476	0.36346	0.24621	0.25216	0.07158	0.28505	0.10890	0.16116	0.10510	0.43193	0.11386	0.21802	-0.01429	0.19983	0.15764
walk	-0.75995	-0.02591	-0.24856	-0.11574	-0.15202	0.05237	-0.12758	-0.14518	-0.05591	-0.09062	-0.02449	-0.07684	-0.03587	-0.09070	-0.09720	-0.20058	-0.18393	-0.06779	-0.04390	-0.02198
give	0.64562	0.12379	0.05626	0.11232	0.04560	-0.04896	0.00153	0.04963	0.01028	0.02371	0.03467	0.02990	0.01421	0.00551	0.04014	-0.01668	-0.00796	0.02092	-0.05353	0.01269
read	-0.48392	-0.22323	-0.31526	-0.02186	-0.03882	-0.00902	-0.07244	0.01302	-0.09258	0.03333	0.01983	-0.04508	-0.08751	-0.00213	-0.00586	-0.04193	-0.05453	-0.01959	-0.01591	-0.06759
run	-0.27085	0.47424	0.32743	-0.00907	0.01437	0.28989	0.01889	0.15702	0.02127	0.02455	-0.00460	0.12499	0.10616	0.11742	-0.16832	0.07155	0.08496	0.07604	0.07774	0.00315
watch				0.08067	0.30192	-0.15755	0.17363	0.24795	-0.21997	-0.07351	-0.00790	0.02257	-0.10494	-0.31996	-0.26553	-0.19138	-0.15276	-0.17547	0.02491	-0.10380
write	0.40370	-0.23360	-0.00597	0.11378	-0.08348	-0.11098	0.20041	-0.02928	-0.16808	0.02106	0.04888	0.17471	0.16485	-0.00957	0.11311	0.13308	0.09841	0.13469	0.04630	0.07515
play		-0.25743	-0.48154	-0.00856	0.03258	-0.03499	0.13291	-0.12665	-0.02634	0.15810	-0.02784	-0.05447	0.03359	0.06569	-0.05053	-0.09690	0.13439	0.03064	0.18702	0.12652
take	0.32043	-0.00150	0.29632	0.03205	0.22734	0.15786	0.03926	0.06827	0.11737	0.02482	0.11704	0.05724	0.06010	0.05164	0.05823	0.02304	-0.01762	-0.03276	0.03318	0.00183
look	-0.06144	-0.06130	-0.05909	-0.07700	-0.14595	-0.15697	-0.22271	-0.19066	-0.24725	-0.22935	-0.21353	-0.24038	-0.25930	-0.32232	-0.27478	-0.25908	-0.21644	-0.29570	-0.27021	-0.26933
start			-0.14996			0.10573		0.00951	0.49280	0.52824	-0.03594	0.05598	0.35641	0.38272	-0.08023	0.05465	0.42872	-0.14940	0.11136	0.18877
bring	-0.10988	0.04563	-0.06905	-0.05130	0.05572	-0.06562	-0.05753	0.00853	-0.12430	-0.07632	-0.08732	-0.06242	-0.05497	-0.07157	-0.05952	-0.04481	-0.04914	-0.04105	-0.06912	-0.03610
help	-0.09755	-0.28860	-0.15663	-0.25098	-0.25553	-0.24043	-0.23249	-0.24078	-0.20087	-0.17895	-0.16847	-0.13793	-0.14512	-0.12476	-0.09424	-0.16737	-0.13332	-0.15447	-0.13210	-0.14011
try					0.47799	0.08552	0.02747	0.17131	-0.40408	0.42379	0.32584	-0.10711	0.53975	0.03040	0.14023	0.03192	0.01772	-0.26179	0.07762	0.04258
wait	0.22693	0.45866	0.37953	0.04184	0.17190	-0.05678	0.12464	-0.13362	-0.05133	-0.04606	-0.02522	-0.09221	-0.04356	-0.08384	-0.06012	0.00260	-0.04552	-0.08134	-0.06221	-0.09655
kill	-0.21281	-0.37124	-0.14999	-0.23024	0.15973	1.19492	0.57179	0.51188	0.34728	0.41315	0.35580	0.37053	0.62896	0.61546	0.30406	0.39805	0.51579	0.61995	0.37645	0.36020
listen			-0.25921	-0.10647	-0.41475	-0.27508	0.01772	-0.03333	-0.09083	-0.18164	-0.16840	0.19154	-0.26086	-0.11982	0.08816	-0.12321	-0.12835	0.06493	-0.12329	-0.10975
ask	0.36181	-0.41723	-0.26680	-0.30416	-0.15188	-0.18874	-0.23657	-0.22005	-0.19809	-0.25819	-0.20863	-0.23436	-0.12401	-0.16269	-0.16845	-0.00849	-0.09544	-0.05683	-0.05689	-0.17795
carry	0.22752	0.21944	0.42468	0.19397	0.43765	0.47102	0.28974	0.37784	0.20712	0.33329	0.21324	0.20119	0.13501	0.19341	0.22180	0.23896	0.19754	0.02153	0.22829	0.26615
share	-0.03653		-0.26555	-0.64401	-0.32060	-0.45898	-0.41426	-0.25580	-0.15741	-0.06974	-0.33165	-0.21024	-0.23048	-0.19865	-0.29702	-0.13101	-0.15715	-0.03747	-0.22256	-0.10957
choose		0.40724	-0.09002	0.04894	-0.40565	-0.29534	-0.03017	-0.10606	0.02053	-0.04353	-0.04937	0.07947	-0.05523	0.05986	-0.05259	-0.04499	0.03349	-0.08732	0.01067	-0.07068

Table 31. Summary table for the *must* construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink	-0.55723	-0.53778	-0.22133	-0.12258	-0.07770	-0.28910	0.04558	0.20164	0.29501	0.15361	0.06617	-0.02028	-0.16572	-0.11945	-0.08182	0.14155	0.00909	0.10516	0.09596	-0.04518
eat			0.55852	0.27488	0.10610	0.26566	0.03302	-0.04478	-0.08903	-0.14308	-0.05456	-0.12960	-0.10708	0.03402	-0.14756	-0.07778	-0.22319	-0.14686	-0.16633	-0.04485
dance				-1.35750	-0.81296	-0.89098	-1.06102	-0.93350	-0.39380	-0.41708	-0.64684	-0.54520	-0.52770	-0.71526	-0.19140	-0.39701	-0.53830	-0.36320	-0.19610	-0.09093
sing			0.24414	-0.20256	-0.75538	-0.63956	-0.23715	-0.46860	-0.34064	-0.09430	-0.20342	-0.13532	-0.08620	-0.14048	-0.14753	0.03465	-0.17204	0.17686	-0.11067	0.15548
cook								-0.67683	-0.87084	-0.92091	-0.54549	-1.43112	-1.18085	-1.33375	-1.64350	-1.41892	-1.46885	-0.76114	-1.03747	-0.75521
talk	0.04579	-0.73970	-0.17762	-0.14144	0.04447	0.03854	-0.24157	-0.01671	-0.22098	-0.22466	-0.18388	-0.22077	-0.29526	-0.19400	-0.20257	-0.12844	-0.04704	-0.16773	-0.11584	-0.09502
think	-0.32340	0.22971	-0.03287	0.09010	0.07402	-0.02570	-0.07713	-0.14652	-0.13423	-0.24629	-0.11581	-0.15280	-0.24378	-0.12803	-0.12695	-0.17301	-0.15337	-0.05902	-0.04263	0.00198
complain																				
cry				0.47522	-0.24433	-0.45180	-0.72826	-0.37589	-0.24956	-0.39519	-0.38886	-0.76000	-0.34997	-0.31243	-0.35708	-0.05475	-0.50841	-0.21702	-0.21737	-0.00944
throw		0.23288	-0.49683	-0.13563	-0.18061	0.06939	-0.10677	-0.10933	0.21347	0.27434	-0.01058	0.02199	0.04463	0.27113	0.08674	0.03349	0.13303	0.17740	0.05116	0.12404
scream							-1.12798	-1.43182	-1.74599	-0.92527	-1.46805	-1.34900	-1.28457	-1.39722	-1.38513	-0.64069	-1.10085	-0.47226	-0.08850	-0.62606
shout							0.26131	-0.61880	0.52716	-0.11766	0.43528	0.68467	0.19588	0.22277	0.14362	0.66038	0.29283	0.81108	0.49779	-0.02397
say	0.00000	-0.29324	0.04517	0.14222	0.24244	-0.06953	0.02006	-0.09982	-0.05682	-0.10357	-0.07268	-0.13791	-0.18428	-0.22953	-0.04740	0.01813	0.02438	0.02803	0.00295	0.00928
speak		0.51739	0.11419	0.27349	0.05630	-0.03987	-0.06872	-0.07226	-0.11087	-0.12282	-0.17785	-0.15331	-0.12943	-0.12592	-0.08121	0.01939	-0.03179	0.04270	0.06360	0.00736
work	0.22206	-0.29052	-0.06008	0.51169	0.24345	0.10605	0.20970	0.02583	0.24814	0.23845	0.14905	0.22170	0.16442	0.18629	0.16546	0.21021	0.08552	-0.04541	-0.04870	-0.03197
consider	0.36228	0.19977	0.55845	0.37492	0.05994	0.31367	0.07835	0.14789	0.17430	0.23910	0.29471	0.16985	0.17996	0.29592	0.24522	0.29711	0.29214	0.16718	-0.08599	-0.11866
walk		-0.06846	0.60097	-0.15348	-0.42099	-0.25382	-0.13933	-0.12353	-0.21210	-0.04893	-0.15595	-0.22493	-0.26236	-0.24512	-0.25892	-0.01610	-0.13722	-0.09405	-0.12176	-0.10758
give	0.09328	-0.09208	0.08190	0.03835	0.12756	0.08094	-0.00035	0.08072	0.00171	0.07026	0.06816	0.02026	0.00459	0.07886	0.21383	0.09526	0.00934	0.08693	0.09774	0.06426
read	0.02352	0.25006	0.02833	-0.14414	-0.20798	-0.09460	0.18616	0.07615	0.05519	-0.02059	-0.00046	0.11025	0.10291	0.22729	0.25378	0.05624	0.30254	0.07136	-0.08840	0.04227
run		0.28216	0.51835	-0.08127	-0.03178	-0.11291	-0.08277	-0.10040	0.00425	-0.03193	-0.05011	-0.06833	-0.13421	-0.15333	-0.09077	-0.11985	-0.06672	-0.07752	-0.06332	0.02758
watch		-0.01917	-0.26400	0.11989	0.21537	0.15554	-0.08903	-0.11108	0.35885	0.06514	0.02308	0.08443	-0.01009	-0.00675	-0.14984	-0.16810	-0.14669	-0.11324	-0.04475	-0.11081
write	-0.39054	0.13592	-0.01364	0.04928	0.10548	-0.02646	-0.06223	-0.04918	-0.06760	-0.02047	-0.03961	0.10149	-0.00309	-0.05458	-0.03461	0.06850	0.05302	-0.02060	-0.03960	-0.01070
play			0.37895	0.67633	0.10605	-0.15376	-0.28303	0.01398	0.06512	-0.01594	-0.12599	-0.07391	-0.02793	-0.01291	-0.03366	0.05261	0.00589	-0.04529	0.03654	-0.02379
take	0.23583	0.17248	-0.04854	0.12924	0.07903	0.08446	0.17836	0.06993	0.05906	0.03953	0.06559	0.10850	0.05359	0.09569	0.05634	0.08559	0.05996	0.06362	0.01502	0.05126
look	0.51704	-0.04291	0.11853	0.12432	-0.02236	-0.06366	-0.01695	-0.03080	0.02519	-0.08049	0.01104	-0.05237	-0.09842	-0.14676	-0.05051	-0.11592	-0.04337	-0.19512	-0.17414	-0.28803
start					0.21518	0.41029	0.03758	0.31447	0.49988	0.05550	0.22979	0.06231	0.16126	0.07629	0.10915	0.03908	0.08117	0.02925	-0.03263	0.04857
bring	-0.04057	0.15748	0.61310	0.37512	0.22619	0.26066	0.14233	0.24329	0.30903	0.12272	0.07617	0.14138	0.20034	0.27928	0.06809	0.09348	0.08606	0.21009	0.17754	0.17511
help			-0.62129	-0.07863	-0.15972	-0.40722	-0.12649	-0.19590	-0.16528	-0.16089	-0.27645	-0.24332	-0.26596	-0.39547	-0.18209	-0.26447	-0.20915	-0.13892	-0.14943	-0.14557
try	-0.47755	-0.12397	0.17597	0.00166	-0.19498	-0.26979	-0.26321	-0.29133	-0.16290	-0.21885	-0.15200	-0.11229	-0.10520	-0.05439	-0.06177	0.04185	0.00506	-0.05721	-0.02601	-0.12269
wait	-0.53031	-0.16607	-0.31795	-0.18252	-0.22014	-0.22997	-0.16671	-0.14875	-0.14523	-0.18841	-0.15949	-0.17628	-0.10465	-0.21366	-0.16311	-0.08333	-0.12427	-0.11676	-0.06227	-0.09538
kill					0.12873	0.54595	0.32711	0.40550	0.64796	0.32249	0.37993	0.25342	0.47977	0.21851	0.03726	0.16859	0.31974	0.44631	0.47377	0.42936
listen		0.09003	0.14480	0.45797	-0.00157	-0.30127	-0.30317	-0.01072	-0.12195	-0.22831	-0.25834	-0.11799	-0.21317	-0.19848	-0.03654	-0.05736	0.01684	-0.03260	-0.17449	-0.10782
ask		-0.29083	0.31694	0.27965	0.11016	-0.28939	-0.20022	0.01389	-0.15685	-0.07932	-0.04883	-0.01126	-0.12080	-0.15685	0.04810	0.09089	0.03459	0.01440	-0.04818	-0.08187
carry	0.15999	0.45187	0.37299	0.10468	0.26847	0.14846	-0.10095	-0.00163	0.16343	0.01304	-0.00656	-0.00270	0.00126	0.04564	-0.03178	0.00903	0.03666	0.09270	0.09543	0.06904
share	-0.95453	-0.32035	-0.42033	-0.50538	-0.50012	-0.38851	0.04177	-0.17480	0.09164	-0.08074	-0.00084	0.01726	0.11559	-0.09292	-0.07657	-0.02924	0.04861	-0.06109	-0.10114	-0.07419
choose	-0.65396	0.19663	0.46439	0.40333	-0.06212	0.08435	0.01597	-0.01031	-0.00465	-0.08870	-0.00032	0.05777	-0.06945	0.11480	0.12174	0.14098	0.13160	-0.05718	-0.06982	-0.04404

Table 32. Summary table for the *must not* construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink				0.29341				0.78711	0.31947	0.92557	0.50585	0.12943	0.56533	0.10833	0.12540	0.08555	0.46014	0.15298	0.29754	0.80527
eat			-0.51594	-0.13385	-0.09260	-0.19602	0.03563	0.42153	0.17792	-0.08393	-0.00127	-0.17338	-0.14322	-0.02362	-0.25099	-0.44901	-0.33078	0.07085	0.03432	0.19847
dance																				
sing																				
cook																				
talk		-1.22016	-0.74005	-0.08791	-0.40297	0.01884	-0.04828	-0.38173	-0.35925	-0.19538	-0.33740	-0.06104	0.32965	-0.00853	0.38982	0.19617	-0.23068	0.07062	-0.18116	0.27767
think	0.36311	0.11857	-0.16196	0.01232	0.04005	0.22107	-0.12942	-0.08674	0.00725	-0.22942	-0.13345	-0.34307	-0.29195	-0.28459	-0.32165	-0.18778	-0.26713	-0.37159	-0.27022	-0.20450
complain	-0.07301			-0.06653	-0.12063	-0.49692	0.71829		-0.03688	0.30153	-0.03414	0.03240	-0.24713	0.13948	0.44171	0.10935	0.37372	0.59894	0.69560	-0.06168
cry					-0.07464	-0.40215	-0.61354	-1.34117	-0.42507	-0.65602	-0.96683	-1.03795	-0.84485	-0.77312	-1.10654	-1.02505	-0.28811	-0.74237	-0.37688	-0.59003
throw					0.00476	0.14425	0.21176		-0.51403	0.55980	-0.09538	0.52774	0.10868	0.14531	0.87327	0.22204	-0.27506	0.22454	0.34816	0.72498
scream																				
shout																				
say				0.62474	-0.10324	0.54491	0.15128	0.00917	0.19062	0.12544	-0.03753	-0.03659	-0.11194	-0.16844	-0.16060	-0.11714	0.21006	-0.03634	-0.08446	0.10556
speak	-0.55412	-1.26039	-0.68019	-0.68131	-0.62669	-0.96167	-0.58916	-0.80056	-0.55832	-0.15052	-0.20142	-0.34005	-0.15145	-0.29105	-0.56664	-0.38841	-0.30186	-0.07696	-0.44870	-0.29150
work			0.44262	-0.53380	-0.12029	0.08594		0.28596	0.33192	0.49106	0.00750	-0.38953	0.44995	0.09188	-0.80422	0.11738	-0.82225	-0.32683	-0.68672	0.07152
consider	-0.37704		0.56859	-0.24335					0.50138	0.48447	0.45896	0.08668	0.36118	-0.09317	-0.13362	-0.01049	0.49183	0.52991	0.25921	
walk				-1.39277	-0.47328		-0.08965	-0.39768	0.03903	-0.16090	-0.16470	-0.27244	-0.21415	-0.44111	-0.12018	-0.60809	-0.33538	-0.20511	0.02569	-0.33833
give		-0.63302	0.16464	0.30909	0.13859	0.38927	0.48731	0.29936	-0.21623	-0.14905	-0.07234	-0.01229	-0.13196	-0.15368	0.02216	-0.17636	-0.28664	-0.09330	-0.21151	-0.11772
read					0.30626	-0.28264	0.16612		-0.07044	-0.10671	1.08284	-0.05794	-0.35713	0.60194	0.53855	0.07592	0.54996	-0.15182	-0.12939	0.32717
run				-0.21301		-0.93659	0.07705	0.10940	-0.71841	-0.42427	-0.22398	-0.03629	0.09309	-0.51283	-0.00831	0.00688	-0.23006	0.19592	0.02797	0.04919
watch																				
write				-0.91270		0.24129	0.05673	0.38287	0.26491	0.86096	0.46875	0.75876	0.71127	0.15877	0.18153	0.58208	-0.02944	0.90054	0.06274	-0.07144
play					-0.42236	-0.03244		-0.01579	0.83038	0.29535	0.31213	0.11733	0.34376	0.30039		0.77611	0.18032	-0.17200	0.10505	-0.37260
take		0.48401		0.92650	0.50414	-0.01357	0.16122	0.13984	0.20457	0.12207	0.16245	0.25371	-0.00383	0.02802	0.14018	0.20627	-0.08890	0.03884	0.17457	0.23913
look		-0.02497	0.03857	-0.23080	0.39442	0.33960	0.25536	0.26562	-0.15778	-0.01510	-0.19059	-0.16920	-0.00962	-0.00118	-0.25253	-0.31238	-0.36145	-0.16230	-0.05627	-0.09583
start																				
bring																				
help																				
try					-0.99732	-0.61650	0.02453	0.32110	-0.09169	0.23026	0.16804	-0.28819	-0.05328	0.22781	0.01081	-0.05395	-0.02836	-0.06736	0.13916	0.39084
wait				0.29341	0.14342	-0.45888		0.37440	0.69211	0.09445	-0.00927	0.12977	0.01455	0.23568	-0.18252	-0.57937	-0.22987	-0.14229	0.08491	-0.20325
kill																				
listen				-0.91270	-0.05334	-0.76078	-0.88705		-0.29231	-0.60611	-0.48708	-0.13058	0.09619	-0.00116		0.13023		0.68622	-0.19652	-0.16648
ask				-1.04171	-0.60113	-0.25516	-0.59620	-0.26351	-0.47109	-0.21047	0.08633	-0.25929	0.05702	-0.05261	0.18574	0.01155	0.04430	0.14556	-0.36764	-0.07945
carry							0.62773	0.10947	0.04056		-0.07628	0.88971	-0.33889	0.32393	0.31624	0.09293	0.67469		-0.60656	0.57019
share																				
choose																				

Table 33. Summary table for the *might* construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink	-0.63067	0.26384	-0.06676	-0.16085	-0.01599	0.01788	0.55953	0.08082	0.15187	0.34219	0.02022	0.08590	-0.03490	-0.30447	-0.01162	-0.08761	-0.17932	0.11530	0.26103	0.21369
eat	-0.06041		0.00414	0.32065	-0.03042	0.20096	0.00005	0.03902	0.11170	0.20025	0.08169	0.04976	0.01171	0.21487	0.20448	-0.09157	0.17966	0.04835	0.09587	-0.02567
dance				-0.76289			-0.69911	-0.84260	-0.87293	-0.09069	-0.47305	-0.41572	-0.76748	-0.85106	-0.25606	-0.16602	-0.71651	-0.51385	-0.70684	-0.64951
sing			-0.19711	-0.31572	-0.30277	-0.13480	-0.20415	-0.50233	-0.55916	-0.23940	-0.53688	-0.30936	-0.27138	-0.31741	-0.45006	-0.26908	-0.18579	0.03504	-0.04194	-0.08672
cook					-0.73199	-0.33433			-0.07532	-0.54933	-1.46616	-0.62177	-0.57961	-1.02700	-1.18912	-0.84729	-0.85478	0.10467	-0.69103	-0.43607
talk	-0.99836	-0.59561	-0.67473	-0.50483	-0.18477	-0.23382	-0.33903	-0.29780	-0.07321	-0.24491	-0.10951	-0.15177	-0.08829	0.04089	0.03281	-0.18935	-0.28178	-0.00999	0.04517	-0.08229
think	0.11904	0.01574	0.03930	0.05412	-0.01231	-0.04454	0.01801	-0.06960	-0.15235	-0.07795	-0.13656	-0.16397	-0.09995	-0.18740	-0.16419	-0.08385	-0.04880	-0.06881	-0.10126	-0.06216
complain	-0.64663		-0.19710	0.08231	0.18451	0.47864	0.31189	-0.24053	0.51953	-0.16915	0.00742	-0.03472	0.01000	0.16618	-0.23284	0.29894	-0.08789	-0.09048	0.00319	-0.07544
cry			-0.67428	-0.61678	-0.73211	-0.52153	-0.87845	-1.07445	-0.52155	-0.80166	-0.75412	-0.52519	-0.75435	-0.55052	-0.67084	-0.57495	-0.61833	-0.56155	-0.38560	-0.41884
throw	-0.47571	0.08773	-0.07716	-0.21889	-0.06302	0.10014	0.00456	-0.17947	0.14361	-0.11575	-0.04247	-0.02616	0.07052	0.01837	-0.15928	-0.12094	-0.05866	-0.04403	-0.13632	-0.18215
scream									-0.15451	-0.07220	-0.34763	-0.55706	-1.13059	-1.18680	-1.77764	-0.74685	-0.82687	-0.59450	-0.69943	-0.69549
shout							-1.04150		-0.42751	0.49209		0.39605	0.22495	-0.01474	0.01504	0.45381	0.56028	0.24736	0.65374	0.15764
say	0.20656	0.29706	0.02312	-0.01123	0.09390	-0.07959	0.09357	-0.01337	0.07679	0.06562	-0.02847	0.00675	0.01051	-0.03038	-0.01376	-0.04414	-0.04938	-0.05625	-0.07336	-0.04008
speak	0.23256	0.15554	0.08715	-0.12635	-0.29430	-0.13125	-0.07012	-0.05616	-0.05424	0.01277	-0.18570	-0.04378	-0.07884	-0.00794	-0.10758	-0.08177	0.00708	0.06369	0.02400	0.08183
work		-0.24833	-0.01077	0.33681	-0.13918	0.06877	0.24675	0.08062	0.07908	0.01948	0.30543	0.19876	0.02031	0.18814	0.07274	0.08751	0.03011	-0.05684	-0.00166	-0.04032
consider	-0.02048	0.17128	0.15878	0.16385	0.17450	0.36901	0.17017	0.41100	0.14370	0.20261	-0.18203	0.23104	0.17481	0.16413	0.09060	0.17930	0.03772	-0.02669	-0.09726	-0.11851
walk		-0.02629	-0.02912	-0.25921	-0.45037	-0.06137	-0.26552	-0.38847	-0.46583	-0.09962	-0.00820	-0.04984	-0.04092	-0.11202	0.09642	-0.02062	-0.11951	-0.08967	-0.04138	-0.02098
give	0.30724	0.05940	0.22465	0.06982	0.08629	0.00673	0.03588	0.09916	0.02365	0.03179	-0.01999	-0.04265	-0.02851	-0.01920	-0.01334	-0.00181	-0.00291	0.03857	0.04902	0.03148
read	0.09305	-0.30669	-0.20588	-0.34563	-0.45575	-0.37200	-0.05890	-0.29605	-0.02290	-0.22972	-0.10035	-0.17086	-0.16322	-0.07458	-0.11011	-0.06631	-0.05625	-0.03112	-0.04404	0.02123
run	-0.22011	-0.39199	-0.15302	-0.07648	-0.11856	-0.39827	-0.26978	-0.33790	-0.28102	-0.23314	-0.22546	-0.19710	-0.10567	0.03093	-0.14927	-0.04294	0.04942	-0.02163	-0.04369	-0.10326
watch		-0.92381	-0.40151	0.00407	-0.20737	-0.27388	-0.15948	-0.13412	-0.19397	0.11155	-0.09029	-0.06135	-0.07210	-0.13694	-0.27685	-0.02177	-0.05366	-0.06739	0.00751	-0.02074
write	0.23241	-0.00509	0.71987	0.08550	0.01532	0.10519	-0.06197	0.05610	-0.02313	0.05547	0.12664	0.16976	0.17504	0.13542	0.23798	0.13708	0.01568	0.11419	0.08115	0.04030
play		0.10409	0.02479	-0.49335	-0.04865	-0.01454	-0.30735	-0.29901	-0.18807	-0.16503	0.03135	0.02490	0.03726	0.06227	-0.07015	-0.03133	-0.05851	0.01137	0.08821	0.11949
take	0.08367	-0.06661	0.07011	0.01208	-0.01415	-0.01843	-0.02022	-0.02495	-0.04824	-0.04356	-0.01430	0.07365	0.03583	0.00379	0.02847	0.01926	0.02782	0.01164	0.05488	0.05474
look	-0.47612	-0.73568	-0.65790	-0.37455	-0.33989	-0.31356	-0.27936	-0.25974	-0.18393	-0.19271	-0.17287	-0.15515	-0.11246	-0.24419	-0.15899	-0.22154	-0.08917	-0.08289	-0.15056	-0.12433
start			-0.14594	0.01535		0.25402	-0.02798	-0.08212	-0.10803	0.15478	-0.05001	0.06053	-0.08132	-0.04768	-0.18602	-0.12241	-0.08472	-0.17294	-0.13609	-0.18440
bring	0.64108	0.38795	0.41075	0.40904	0.21850	0.16952	0.16379	0.13481	0.07644	0.08203	0.13008	0.04780	0.04140	-0.00581	-0.01536	0.03381	0.09814	0.19852	0.14412	0.18668
help	0.23241	-0.74158	0.24241	0.20513	-0.27554	-0.26771	-0.30889	-0.27902	-0.19402	-0.31977	-0.19729	-0.16565	-0.26681	-0.25683	-0.14469	-0.15891	-0.07918	-0.09858	-0.00725	-0.04920
try	-0.27841	-0.15152	0.17107	0.11459	-0.25192	-0.20964	-0.18286	-0.10862	-0.12594	0.02477	0.04556	0.03104	-0.10670	-0.02378	0.03583	-0.00050	-0.02040	0.00602	0.04343	0.02822
wait	0.14803	-0.43551	-0.40151	-0.43690	-0.34282	-0.25150	-0.18868	-0.08925	-0.40578	-0.15306	-0.12655	-0.23253	-0.18707	-0.13281	0.03354	-0.05172	0.04839	-0.04341	0.13176	0.03170
kill		0.47222		0.38348	0.31932	0.45431	0.30900	0.29553	0.15530	0.35736	0.28430	0.48708	0.13753	0.14933	0.23817	0.07319	-0.00269	0.28749	0.24542	0.28982
listen	-0.52120	-0.72593	-0.54427	-0.43572	-0.43123	-0.30814	-0.48022	-0.56271	-0.09607	-0.31333	-0.23244	-0.12282	-0.13568	-0.30227	-0.17307	-0.19391	0.10911	-0.02028	-0.04375	0.02545
ask	0.04323	-0.05435	0.16124	-0.42106	-0.38419	-0.25730	-0.20973	-0.25619	-0.11365	-0.18668	-0.08138	-0.00848	-0.02487	-0.26597	-0.13625	-0.11043	-0.05345	-0.18934	-0.07009	-0.14067
carry	-0.23886	-0.26181	-0.06746	0.04069	0.03288	0.04969	0.15625	0.21003	0.19681	0.20106	0.08882	0.15469	0.19982	0.21738	0.17560	0.14029	0.14945	0.14915	0.10183	0.03619
share	-0.29931	-0.48277	-0.81437	-0.40662	-0.58777	-0.34093	-0.21333	-0.22045	-0.17174	-0.25349	-0.30542	-0.18034	-0.23063	-0.25750	-0.07894	-0.24751	-0.26620	-0.06888	0.04658	-0.03068
choose	-0.22052	-0.10325	-0.07813	-0.03419	-0.09672	-0.06389	-0.04059	0.09685	-0.03503	0.09126	0.01906	0.00611	0.03417	0.03109	0.13269	0.04821	0.04855	0.02721	-0.10145	-0.13918

Table 34. Summary table for the *might not* construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink																				
eat	-0.91354		-0.75866	0.20080	0.30844	-0.54178	-0.28367	-0.34828		-0.43936	-0.88718	-0.62970	-0.31647	-0.40945	0.46291	-0.26855	0.71316	0.16597	0.54310	-0.10416
dance																				
sing																				
cook																				
talk									0.14903		-0.09369	-0.51990	0.14063	0.00152	0.08662	-1.06442	0.46298	0.00302	-0.03907	-0.04199
think		0.05215	0.19872	0.47669	-0.18175	-0.04420	-0.02816	-0.11460	0.02425	0.09697	-0.00268	-0.13516	-0.28721	-0.02865	0.06843	-0.04634	0.09263	-0.03140	-0.05663	0.06281
complain																				
cry																				
throw																				
scream																				
shout																				
say					0.35427	0.06114	0.64210	0.38761	-0.23155	0.19288	0.00849	-0.30104	-0.74964	-0.65423	-0.29450	-0.17490	-0.21626	-0.57523	-0.04359	-0.05145
speak				-0.04962	-0.22095	-0.50426	-0.04461	0.00129	0.21552	0.04030	-0.09445	-0.17562	0.53987	-0.09053	0.18384	0.03689	-0.15884	0.15822	0.15491	-0.08991
work																				
consider		-0.49332	-0.33204	-0.40189				0.36407	0.58869	-0.24471	0.25326	-0.62866	0.01555			-0.38280	0.28689	-0.11770	-0.04832	0.34648
walk				-0.00364			-0.74989		0.58869	-0.06835	0.63374	0.02493	-0.83073		0.08662	-0.28567	0.46298	0.05859	0.46396	0.08306
give	0.38290	0.46541	0.83861	0.60325	0.53308	0.17311	0.37333	0.40269	0.00769	0.36146	0.22655	0.22530	0.35783	0.32123	0.72914	0.01564	0.35460	0.63535	0.39440	0.13283
read				-0.70303	-0.85250	-0.54273	-0.80871	-0.83525	-0.54125	-0.48822	-0.44710	-0.29784	-0.76392	-0.09560	-0.73814	-0.04245	-0.39493	0.13775	0.41807	0.03726
run		-0.49450		-0.50032	-0.42490		0.14839	-0.04682	0.18416	0.76476	-0.23081	0.43316	0.14070	0.04736	-0.06850	-0.08159	1.08978	0.08663	0.51839	0.15007
watch																				
write						-1.32112	-1.10827	-1.37030		-0.53379	-0.33705	0.02492	-0.21886	0.22352	0.02857	0.45518	0.16182	0.35982	-0.11043	-0.27614
play																				
take			0.27724	0.21146		0.43897	-0.18761	-0.28192	0.12874	0.36754	0.15303	-0.00674	-0.00424	0.22440	-0.02397	-0.20130	0.02050	0.40621	0.03633	0.18072
look		-1.27524	-0.83983	-0.18027	-1.01825	-0.68228	-0.00449	0.06264	0.12146	-0.06494	-0.12909	-0.08472	-0.70258	-0.44523	-0.47663	-0.23505	-0.24919	-0.02321	-0.47413	0.06571
start																				
bring			-0.25331		0.30844			0.48888		0.83546	-0.22433	0.54855	0.01557		0.32978	0.01560		-0.09645	0.46396	0.33847
help																				
try						-0.24018					0.73077	-0.15123	0.65305	-0.17464	-0.51630		-0.25786	0.10003	0.10717	-0.32102
wait																				
kill																				
listen																				
ask				-1.00557		-0.24018			-0.05517		0.33260	0.24698			-0.14771		-0.09379	-0.24266	0.71939	0.01607
carry																				
share					-0.04420	-0.54131		-0.68864	-0.94712	-0.96375	-0.39168	-0.45369	-0.46300	0.06862	0.18354	0.08261		-0.33967	-0.05928	-0.77813
choose			0.47511	-0.20939	-0.36736	-0.40236	-0.46113	-0.19336	0.29574	-0.06066	-0.11497		-0.06370		0.76412	-0.13953	-0.19073	0.13794	0.01456	0.15011

Table 35. Summary table for the *should* construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink	-0.77098	-0.25173	-0.17179	-0.41227	-0.20069	0.06941	0.26148	0.36083	-0.04521	-0.16174	-0.12387	-0.14284	0.07775	0.02873	-0.16284	0.05326	-0.02817	-0.11551	-0.11379	-0.11309
eat	0.53177	0.43870	-0.17590	0.28495	0.09506	0.35897	0.05280	-0.01666	-0.03461	0.21723	0.04618	-0.07162	0.04819	0.07064	-0.13502	-0.01762	0.04159	-0.13046	-0.09737	-0.11751
dance			-0.85519	-0.98631	-0.66624	-0.57984	-0.90233	-0.43631	-0.25022	-0.65813	-0.52394	-0.33822	-0.37510	0.12910	-0.26566	-0.51664	-0.77290	-0.40770	-0.59844	-0.07325
sing	-1.95577	-1.52205	-1.47914	-1.29695	-1.25797	-0.81873	-0.78950	-1.17140	-1.01535	-0.83330	-0.64499	-0.53359	-0.56826	-0.57818	-0.69957	-0.56781	-0.49078	-0.17612	-0.37987	-0.23870
cook				-0.33267			-0.66811	-0.28608	0.08083	-1.11101	-0.31179	-1.31336	-1.13688		-0.99269	-0.95480	-1.19509	-1.00103	-0.50132	-0.61864
talk	-0.92162	0.07729	-0.46594	-0.73952	-0.14038	-0.26332	0.05893	-0.11420	-0.13846	-0.08981	-0.20044	-0.28438	-0.15724	-0.10430	-0.13055	-0.33976	-0.22724	-0.14106	-0.04342	-0.11574
think	-0.01620	-0.02285	0.15019	0.09015	0.01395	-0.09567	-0.07015	-0.16900	-0.17746	-0.04186	-0.10135	-0.02389	-0.04458	-0.09176	-0.14373	0.01686	0.01085	0.01558	-0.06794	-0.00832
complain		-0.39994	-0.39129	-0.50884	-0.00281	0.27769	0.41114	0.13368	-0.43531	-0.25643	-0.02653	-0.43152	-0.28500	-0.10501	0.36282	-0.10087	0.23993	-0.00971	0.21200	0.07457
cry	-0.60240	-0.78031	-0.36901	-0.43636	0.05829	-0.53304	-0.78908	-0.67362	-0.59657	-0.57120	-0.68457	-0.42972	-0.68326	-0.58242	-0.66326	-0.59592	-0.38711	-0.29862	-0.19464	-0.35881
throw	-0.24698	0.29928	0.19132	0.12664	-0.02839	-0.17890	-0.26854	-0.15822	-0.16422	-0.21615	0.07469	-0.19812	-0.18889	-0.20651	-0.22594	-0.23961	-0.03064	0.00946	0.05098	-0.06757
scream				-1.11085			-1.44633	-0.50796	-1.20188	-2.09639	-1.28202	-1.91553	-1.07733	-1.12625	-1.10667	-0.81761	-1.38960	-0.50267	-0.72114	-0.60331
shout						-0.37559			-0.53899	-0.00543	-0.39096	0.46484		-0.17196	-0.81655	-0.06743	0.20155	0.06948	-0.19054	0.28847
say	0.61803	0.42999	0.22888	0.20533	-0.01515	0.04167	0.01290	0.06878	-0.06438	0.00309	-0.04028	-0.08209	-0.00177	-0.04065	-0.04073	0.05971	-0.00426	-0.01440	0.02448	0.00854
speak	0.27923	0.19802	0.02958	0.09996	0.21869	-0.01871	-0.03139	-0.01285	-0.09050	-0.05425	-0.13659	-0.09048	0.01811	-0.11027	-0.10085	-0.12603	-0.12480	0.00169	0.02370	-0.01626
work	0.11390	-0.21712	-0.03244	-0.17465	0.16646	0.13309	0.04460	0.20778	0.24280	0.35116	0.18980	0.29377	0.15107	0.06168	0.15107	0.15476	-0.07080	-0.11142	-0.04481	-0.07868
consider	0.42939	0.29811	0.31258	0.23772	0.32233	0.26508	0.45215	0.33117	0.35303	0.37498	0.28200	0.11735	0.26977	0.20240	0.30853	0.28323	0.17302	-0.01992	-0.12875	-0.12553
walk		0.66738	0.07338	-0.06741	0.02427	-0.27119	-0.22514	-0.15164	-0.13362	-0.07794	-0.16113	-0.11657	0.48935	0.09969	-0.19660	-0.05794	-0.16281	-0.26967	0.04826	-0.03578
give	0.29692	0.40714	0.22770	0.12483	0.15835	0.09049	0.08930	0.08699	0.04292	0.01049	0.03761	0.04974	-0.00672	-0.01895	0.01108	-0.07121	0.01436	0.07423	0.07584	0.07423
read	-0.28055	0.24390	-0.17560	-0.03794	-0.09990	-0.12908	0.12918	-0.10726	-0.03594	0.02616	0.08316	-0.02925	0.06827	0.07666	0.15077	0.22688	0.14907	0.05777	0.00664	-0.13516
run	0.05433	0.20244	-0.23632	0.14462	0.47124	-0.12740	-0.12415	-0.00051	-0.07280	0.02839	0.07732	0.07445	-0.00351	-0.03154	0.00086	-0.01772	0.11979	0.05672	0.06278	0.04256
watch	-0.93400	-0.31281	-0.12593	-0.01044	0.21232	0.42042	0.03089	-0.08198	0.32110	0.33114	-0.02684	0.01058	-0.00449	0.01063	-0.13540	-0.14762	-0.30566	-0.03529	-0.00145	0.06901
write	0.09352	-0.02782	-0.05593	-0.00043	0.16671	0.12275	0.09108	0.08708	0.15185	0.13757	0.10874	0.12298	0.16943	0.09195	0.01957	0.05455	-0.00359	0.10033	0.00297	0.07805
play		-0.03187	-0.12589	-0.41722	-0.09976	0.10162	-0.04487	-0.02022	-0.28512	-0.23201	-0.10280	-0.06981	-0.17318	-0.33752	-0.08368	0.01406	-0.06745	-0.04068	0.14716	0.07553
take	-0.01048	0.00132	-0.07816	-0.07151	-0.01803	0.03301	0.06810	0.03255	-0.00905	-0.01158	0.01009	0.02066	0.02570	0.00203	0.00225	0.00407	0.00310	0.03068	0.00235	0.02057
look	-0.25054	-0.03193	-0.10490	0.04160	-0.11261	-0.11992	-0.14691	-0.02079	-0.05353	0.03517	-0.05628	-0.02544	0.07778	0.00071	0.05449	-0.04543	-0.07075	-0.05322	-0.07182	-0.09264
start				0.26947	0.10314	-0.03427	-0.13497	-0.02692	-0.06067	0.02311	0.08257	0.22615	0.00748	-0.00410	-0.13625	-0.06059	-0.01849	-0.22795	-0.04315	-0.12250
bring	-0.33952	-0.15071	-0.15418	-0.09912	0.07665	0.05600	0.12390	0.05559	-0.01666	-0.08379	-0.00110	0.06858	0.01749	-0.00942	-0.05539	-0.05897	0.04324	0.03651	0.07319	0.05490
help	0.14610	0.07732	0.12947	-0.20778	-0.11555	0.40990	0.00092	0.01807	-0.08165	-0.01820	-0.15041	-0.18399	-0.48463	-0.23825	-0.20373	-0.33206	-0.17243	-0.13035	-0.00729	-0.04350
try	-0.07073	0.32733	-0.17395	0.14718	-0.11010	-0.08067	-0.21872	-0.01720	0.12081	0.20998	-0.02713	0.05325	0.08538	0.03588	-0.00534	-0.02076	-0.10401	-0.10068	-0.03737	-0.08226
wait	0.04133	0.33638	-0.11278	0.20589	-0.00203	0.00979	-0.21852	-0.00055	0.00221	-0.16798	-0.27766	-0.25643	-0.03386	-0.19374	-0.26250	-0.12213	-0.21536	-0.05011	-0.14481	-0.06016
kill			0.15302	0.36650	0.43836	0.65609	0.69770	0.45880	0.64072	0.39493	0.38442	0.55659	0.42201	0.52995	0.35623	0.03872	0.25739	0.35207	0.34607	0.47486
listen	-0.77865	0.26923	-0.92214	-0.61361	-0.69180	-0.56971	-0.42761	-0.46472	-0.49114	-0.09345	-0.01074	-0.33833	-0.04304	-0.08048	0.07616	-0.10466	-0.12955	-0.15790	-0.09717	-0.02190
ask	-0.24571	-0.20745	-0.34122	-0.10311	-0.16135	-0.27780	-0.24674	-0.21053	-0.25681	-0.04492	-0.12292	-0.05184	-0.15974	-0.12732	-0.01285	-0.03043	-0.14269	-0.10148	-0.09659	-0.10219
carry	0.07193	-0.12392	-0.06672	0.10389	-0.00305	0.40019	-0.06772	0.11585	0.24474	0.14868	-0.05780	0.00238	0.05078	0.22405	0.09209	0.12604	0.01226	0.06174	0.15617	0.11907
share	-0.46572	-0.46539	-0.56351	-0.48238	-0.42353	-0.41022	-0.38721	-0.36698	-0.44391	-0.38958	-0.47609	-0.32068	-0.30416	-0.27794	-0.38172	-0.27686	-0.14836	-0.08420	-0.02158	-0.29813
choose	0.37184	0.04336	-0.06573	0.14490	-0.06000	-0.03722	-0.10491	-0.04530	0.08628	-0.05360	-0.08332	-0.04248	-0.02990	-0.02704	-0.04182	0.06307	0.06082	-0.02308	0.06264	0.04578

Table 36. Summary table for the *should not* construction.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink					0.42790			0.19406		0.35836	0.27069	0.58235	-0.24908	-0.05938	0.12477	0.20117	-0.00058	0.35594	0.13160	0.18956
eat			0.01263	-0.10194	-0.20240	0.07021	-0.49164	-0.62802	-0.15461	0.57937	-0.03644	-0.10754	-0.03976	-0.04613	0.28669	0.19922	-0.04955	-0.01857	-0.02370	0.11524
dance																				
sing						-0.23111		-1.49711	-0.79947	-0.73939	-0.98539	-1.04679	-1.52097	-0.77155	-0.83867	-0.66677	-0.52208	-0.56824	-1.17498	0.40212
cook																				
talk			0.15617	-0.30594	0.00847	-0.32831	-0.67742	-0.35045	-0.35455	0.89017	-0.25305	-0.38550	0.11569	0.28560	-0.00029	0.06780	-0.18990	0.15931	0.03857	-0.00345
think	0.66228	0.27859	0.27944	0.00076	-0.13922	0.09497	-0.22930	-0.11593	0.00945	-0.25671	-0.11733	-0.10113	-0.03146	-0.06429	0.04750	-0.15395	-0.02058	-0.03940	0.09542	-0.09418
complain					-0.01959			0.65038			0.42829	0.73227	-0.09409	-0.36062	0.10247		0.06644	0.13116	-0.12560	0.15916
cry			-0.14513	0.25084		-1.10660		-0.28332	0.15556	0.61368	-0.68430		-0.49247	-0.66178	-0.15533	-0.18117	-0.45544	-0.08415	-0.66119	-0.26500
throw			-0.62301	-0.70467	-0.54272			0.22413	0.37778	-0.50397	-0.33225	0.30607	0.22829	0.41809	-0.30155	0.20909	0.13136	-0.14216	-0.25045	0.06216
scream																				
shout																				
say			-0.54567	0.50166	0.10567	0.62781	0.03799	-0.21968	-0.46585	-0.31326	-0.08725	-0.04138	-0.17340	-0.08322	0.00454	0.48950	0.10703	0.18113	0.13202	0.08690
speak	-0.28762	0.25842	0.65218	-0.12042	-0.09044	0.00316	0.21328	-0.14309	-0.19785	-0.12990	-0.09373	0.54814	-0.10063	-0.07048	0.01119	-0.15231	0.13687	0.08002	-0.14362	0.08457
work				-0.00464	0.23597		0.60607	0.78093	-0.18357	0.53482	0.23017	0.25503	-0.57197	0.49370	-0.00028	-0.15139	-0.26899	-0.34060	-0.29396	-0.07958
consider				-0.14769	0.74471	-0.21930		0.04796	0.02085	0.14552	0.14579	0.06770	0.32138	-0.20312	0.62410	0.32014	0.03865	0.07557	-0.14693	-0.10082
walk					-1.11735	-1.44452	-0.92828	-0.98127	-0.60292	-1.05625	-0.85375	-0.33167	-0.70516	-0.78771	-0.39865	-0.36264	-0.47173	-0.18373	-0.52671	-0.09745
give	0.06583	0.66666	0.15683	0.66726	0.28317	0.37409	0.03134	0.00212	-0.11406	-0.01244	-0.09068	-0.11916	0.18829	-0.09396	0.04874	0.03805	-0.02898	0.02172	-0.12796	0.10644
read			-0.27013			0.45070	0.55210	0.25219	0.47467	-0.11673	0.03952	-0.48840	-0.22636	-0.20577	0.09328	0.03293	0.17290	0.07989	-0.24258	0.34681
run				-0.30594	-0.84412	0.14937	0.27865	0.69184	0.36690	-0.49732	-0.39027	-0.49188	0.06964	0.26355	0.17026	0.05232	0.37074	-0.03298	-0.08300	0.19078
watch																				
write		-0.91172	-0.27048	-0.50040	-0.01960	-0.11720		1.01289	0.81780	-0.02109	-0.03676	0.10350	0.04601	0.48569	-0.04180	0.95944	-0.23606	-0.02038	0.03576	-0.05671
play			-0.27013	-0.10166	-0.27957	-0.08488	0.52162	0.30340	0.33184	0.06946	-0.01287	0.01169	-0.10045	0.20423	-0.04609	-0.05437	0.29976	0.07993	-0.33016	0.07889
take		-0.00558	0.22449	0.13798	-0.26415	0.42006	0.19339	0.16077	-0.05573	-0.11370	-0.07061	0.03079	-0.07900	-0.04154	-0.05268	-0.09191	-0.11273	-0.12556	0.03048	-0.10412
look		-0.18223	-0.31479	-0.06701	-0.04192	-0.35685	0.07424	0.03189	-0.21535	0.12087	-0.07221	-0.05492	0.22202	-0.26894	-0.05168	-0.13554	-0.25955	-0.04859	-0.11933	-0.28398
start						-0.23111	-0.28297	0.30327			0.18709	0.25495	-0.13991	0.39980	0.64357	0.20920	0.36772	0.18547	-0.13332	0.20837
bring	-0.89044	-0.48384	-0.82820	0.14180	0.37060	0.02439	0.12858	0.42874	-0.07876	0.02439	-0.15575	0.09212	-0.13991	-0.45814	-0.20674	-0.10443	-0.02364	0.41295	0.13746	0.30565
help																				
try					-0.11657	-0.38635	-0.38322	-0.01432	0.21515	-0.02801	0.30189	-0.05449	0.09618	-0.00175	0.26974	-0.03111	-0.15844	-0.05099	-0.09649	-0.13539
wait			0.15617		0.51233	-0.21432	0.30486	0.04794	0.44484	-0.18794	-0.26544	-0.22896	0.16172	-0.05233	-0.12558	0.08600	-0.03591	-0.31246	-0.33038	-0.38899
kill																				
listen					-0.01958				-0.79922	-1.38303	-0.63291	0.58235	-0.44113	-0.76002	-0.49554	-0.26838	-0.20764	0.63635	-0.32975	-0.19236
ask			-0.14523		-0.46846	-0.41910	-0.38880	-0.27450	0.37087	0.07438	0.11002	-0.09501	-0.04883	0.25605	-0.12254	-0.41327	0.06375	-0.02609	-0.07852	-0.00590
carry					0.10552	0.81950	0.30486	0.34938	0.03063	0.22766	-0.04448	-0.21501	0.40469	-0.03814	0.46256	-0.19401	-0.16996	-0.00376	0.08076	0.30547
share				-1.00512			0.07406	-0.82754	0.58177	-0.99466		0.05064	0.12355	-0.56484	-0.64884	-0.82489	-0.22098	-0.33422	-0.63098	-0.64639
choose		-0.02046	-0.32829	-0.30701	-0.05866	0.37171	0.55163		0.06171	0.47257	-0.24750	-0.47297	0.33765	0.09210	-0.22237	-0.62962	-0.37422	-0.33879	0.02269	-0.09745

Table 37. Summary table with the average log odds ratios of each of the 36 verbs per decade taking all 26 constructions into consideration.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink	-0.27233	0.08063	0.15441	0.10037	0.25850	0.23352	0.24515	0.21368	0.24141	0.29719	0.22837	0.17176	0.12816	0.11232	0.03104	0.06659	0.04028	0.12571	0.11484	0.14358
eat	-0.25175	-0.21804	-0.04209	-0.05535	-0.08258	0.04671	-0.01449	-0.07869	0.02999	0.00049	-0.05671	-0.03721	-0.07520	-0.06881	-0.06183	-0.10079	-0.08765	-0.08290	-0.05507	-0.09419
dance	-1.02327	-1.01834	-0.90104	-0.80871	-0.70943	-0.61689	-0.65079	-0.64967	-0.48860	-0.66141	-0.53001	-0.59001	-0.60167	-0.59074	-0.51424	-0.64730	-0.56481	-0.48938	-0.34466	-0.32662
sing	-0.83972	-0.69542	-0.51094	-0.61941	-0.60247	-0.49731	-0.56437	-0.56232	-0.49015	-0.44539	-0.43867	-0.36298	-0.42873	-0.37567	-0.46012	-0.40927	-0.41799	-0.22139	-0.29179	-0.14365
cook		-0.56523	-0.52088	-0.52599	-0.89231	-0.53016	-0.61527	-0.70029	-0.57150	-0.72860	-0.78418	-0.82580	-0.94749	-0.91353	-1.02417	-1.01258	-0.88694	-0.72050	-0.72156	-0.55024
talk	-0.32312	-0.33054	-0.16253	-0.26840	-0.17415	-0.16520	-0.15291	-0.16414	-0.11948	-0.04889	-0.07756	-0.07936	0.03460	-0.01157	-0.04238	-0.11600	-0.06780	-0.02917	-0.04057	-0.01241
think	-0.06976	0.00564	-0.05190	0.02631	-0.10587	-0.05686	-0.10486	-0.12965	-0.07991	-0.09075	-0.05729	-0.07473	-0.08885	-0.07644	-0.06703	-0.06140	-0.03924	-0.07549	-0.06015	-0.05476
complain	-0.26720	-0.24547	-0.19485	-0.16608	-0.09747	-0.10842	-0.08248	-0.07213	-0.20982	-0.05256	-0.08399	-0.00324	-0.18094	-0.11022	-0.06093	-0.14999	-0.11601	-0.11487	-0.08174	-0.08308
cry	-0.54755	-0.45225	-0.50042	-0.31480	-0.51018	-0.61551	-0.65117	-0.53159	-0.47703	-0.53300	-0.61913	-0.63125	-0.60325	-0.59527	-0.53189	-0.51060	-0.50160	-0.44132	-0.41078	-0.38915
throw	-0.26044	-0.20256	-0.07022	-0.15002	-0.04892	-0.02355	-0.03191	-0.02470	-0.03457	0.00162	-0.02038	0.01231	0.05966	-0.00141	0.04496	0.00069	0.03018	0.05884	0.07035	0.09782
scream	-1.19370	-1.23637	-0.73148	-0.91910	-0.96630	-0.75034	-0.95358	-0.72846	-0.79798	-0.75730	-0.89212	-0.82529	-0.77514	-0.78250	-0.88676	-0.76707	-0.79592	-0.55325	-0.51326	-0.55367
shout	0.23454	0.44178	0.18526	0.09558	0.19254	0.41927	0.18945	0.06337	0.22929	0.34890	0.42679	0.56745	0.36348	0.23456	0.09944	0.11465	0.10651	0.26130	0.19710	0.18523
say	0.17186	0.11398	0.05986	0.13857	0.10357	0.06868	0.04387	0.01560	-0.00094	-0.01029	-0.04870	0.00879	-0.02859	-0.00689	-0.02230	0.01450	0.01106	0.00759	0.04196	0.04699
speak	-0.08004	-0.02317	0.01364	-0.01325	-0.00070	-0.13101	-0.03880	-0.06555	-0.07635	-0.06370	-0.07014	-0.04612	-0.04436	-0.04979	-0.04879	-0.02706	-0.02292	0.05885	0.02407	0.03617
work	-0.01724	-0.04836	0.03692	0.07967	0.10916	0.11213	0.24058	0.22942	0.22476	0.26753	0.19327	0.19176	0.13241	0.15771	0.08067	0.05895	-0.06954	-0.07754	-0.07957	-0.03648
consider	0.08985	0.02190	0.08608	0.08070	0.14666	0.13715	0.15861	0.24030	0.25902	0.23316	0.17777	0.19596	0.20594	0.17202	0.22224	0.21044	0.21072	0.10131	0.04559	0.02638
walk	-0.16214	-0.17241	-0.06524	-0.19363	-0.27424	-0.13635	-0.18559	-0.17157	-0.15851	-0.16021	-0.09112	-0.13639	-0.15013	-0.14324	-0.13497	-0.17031	-0.15170	-0.06469	-0.04930	-0.03486
give	0.14772	0.19686	0.12114	0.09695	0.09891	0.09034	0.11955	0.09425	0.05612	0.05306	0.03338	0.01208	0.02341	0.00004	0.02545	-0.04001	-0.01584	0.03846	0.02179	0.04622
read	-0.08098	-0.09105	-0.08790	-0.09862	-0.04556	-0.05840	-0.02899	-0.05606	-0.06766	-0.03341	0.02179	-0.00270	-0.05283	0.08802	0.04121	0.05837	0.05491	-0.00297	-0.02764	-0.01697
run	-0.40401	-0.16796	-0.16774	-0.19776	-0.12469	-0.09131	-0.10739	-0.03489	-0.07583	-0.09481	-0.11764	-0.06539	-0.03542	-0.09060	-0.02240	-0.04107	0.04609	0.03549	0.01916	0.03489
watch	-0.18949	-0.33162	-0.33264	-0.27647	-0.20915	-0.18085	-0.19571	-0.14642	-0.13417	-0.11542	-0.09680	-0.06277	-0.16548	-0.13037	-0.10970	-0.14703	-0.13874	-0.09941	-0.05627	-0.07240
write	0.15426	0.00132	0.10447	0.01917	0.13520	0.08777	0.06559	0.11115	0.14544	0.10020	0.13370	0.18316	0.17490	0.14848	0.12379	0.23567	0.09445	0.15947	0.07323	0.07241
play	-0.34292	-0.32468	-0.22951	-0.20230	-0.23520	-0.13356	-0.13653	-0.04390	-0.08348	-0.12303	-0.05470	-0.08638	-0.00340	-0.03968	-0.03957	-0.00043	0.00606	0.02402	0.06854	0.06813
take	0.09700	-0.02157	-0.03349	0.05883	0.04535	0.05114	0.01592	0.01059	0.03109	0.03434	0.03064	0.02265	0.02789	0.01281	0.02034	-0.01047	-0.01559	0.01522	0.01261	0.02558
look	-0.03972	-0.19300	-0.24219	-0.23661	-0.17995	-0.21122	-0.21771	-0.22047	-0.21711	-0.16559	-0.16855	-0.15834	-0.18132	-0.16837	-0.16926	-0.18842	-0.18295	-0.13770	-0.12108	-0.11935
start	-0.29902	-0.30475	-0.32429	-0.18879	-0.03255	0.03649	-0.07767	0.02083	0.05826	0.12991	0.09259	0.11680	0.03076	0.07512	0.03252	0.04309	0.06155	-0.03526	-0.05166	-0.02442
bring	0.00021	0.05458	-0.00769	0.06305	0.09545	0.02501	0.05285	0.09951	0.05032	0.12332	0.02329	0.05962	0.03849	0.01142	-0.01797	-0.02798	0.02973	0.10208	0.10397	0.11338
help	-0.11159	-0.14546	-0.02394	-0.14549	-0.03523	-0.01845	-0.13611	-0.14004	-0.07007	-0.17348	-0.19045	-0.02718	-0.11761	-0.16868	-0.16670	-0.16048	-0.10739	-0.02391	-0.05370	-0.01159
try	-0.09206	-0.20706	-0.13244	-0.03580	-0.10194	-0.17231	-0.11737	0.00034	0.01207	0.02516	0.04638	-0.02608	0.04271	0.01529	0.02552	0.02141	-0.01563	-0.00579	-0.00704	-0.01438
wait	-0.16843	-0.06912	-0.04383	-0.00789	-0.02352	-0.21063	-0.15224	-0.07891	-0.08997	-0.09591	-0.15145	-0.19124	-0.16081	-0.18868	-0.21012	-0.20554	-0.15310	-0.13227	-0.05873	-0.09516
kill	0.00647	0.12200	0.30055	0.25457	0.29093	0.35056	0.34572	0.46434	0.42341	0.35945	0.39611	0.41014	0.37235	0.33801	0.28236	0.18166	0.26273	0.31509	0.33338	0.32757
listen	-0.45161	-0.26491	-0.26912	-0.32509	-0.27325	-0.29630	-0.29643	-0.27440	-0.27974	-0.36835	-0.20199	-0.18247	-0.19747	-0.20047	-0.15924	-0.14092	-0.06740	-0.01373	-0.06126	-0.05635
ask	-0.16621	-0.25860	-0.14708	-0.13220	-0.13056	-0.11597	-0.14501	-0.12211	-0.06045	-0.08454	-0.01028	-0.08824	-0.07872	-0.07011	-0.10977	-0.05557	-0.05127	-0.04700	-0.03085	-0.07547
carry	-0.10256	-0.19039	0.05877	0.12137	0.05334	0.25579	0.08643	0.11781	0.18025	0.13001	0.03566	0.10530	0.03996	0.07473	0.06958	-0.01183	0.04961	0.09879	0.08327	0.13285
share	-0.42298	-0.42121	-0.41374	-0.39234	-0.21437	-0.17944	-0.10233	-0.17934	-0.12854	-0.20765	-0.15483	-0.17342	-0.06893	-0.05282	-0.10814	-0.08816	-0.14743	-0.10792	-0.11894	-0.15039
choose	-0.04048	0.10535	-0.02420	0.05947	-0.03090	-0.08603	-0.02508	-0.03803	0.03414	-0.02049	-0.02091	0.00715	0.01711	0.11711	0.09569	0.06579	0.04419	-0.03412	-0.04725	-0.06492

Table 38. Summary table with the average absolute log odds ratios of each of the 36 verbs per decade taking all 26 constructions into consideration.

Verb	1810s	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s
drink	0.53745	0.33510	0.26365	0.32809	0.33487	0.29867	0.30734	0.26128	0.26255	0.32599	0.25681	0.21257	0.20655	0.16369	0.13301	0.11984	0.09715	0.15109	0.13617	0.18512
eat	0.36986	0.28738	0.20472	0.24379	0.21194	0.21286	0.14989	0.17534	0.12656	0.15954	0.11777	0.13249	0.12907	0.10938	0.15761	0.12132	0.16199	0.11387	0.11372	0.12235
dance	1.02327	1.01834	0.90104	0.80871	0.70943	0.62127	0.65079	0.66406	0.48860	0.66141	0.55621	0.59001	0.60167	0.61172	0.51424	0.64730	0.56481	0.48938	0.37014	0.33570
sing	0.84940	0.69542	0.56169	0.61941	0.60247	0.50599	0.56437	0.56232	0.51713	0.44539	0.43867	0.39684	0.42873	0.38572	0.46991	0.41215	0.41799	0.28645	0.29179	0.19012
cook		0.58422	0.64923	0.52599	0.89231	0.58171	0.61527	0.70029	0.58161	0.72860	0.78418	0.82580	0.94749	0.91353	1.02417	1.01258	0.88694	0.73281	0.72156	0.55024
talk	0.37480	0.35756	0.23871	0.27675	0.19154	0.21150	0.20662	0.20475	0.15875	0.19126	0.13386	0.15220	0.15765	0.10192	0.11580	0.17787	0.15631	0.08998	0.09456	0.08075
think	0.20797	0.13756	0.15385	0.14683	0.14719	0.17235	0.12927	0.18212	0.11786	0.12770	0.09112	0.10942	0.12830	0.09943	0.10258	0.10159	0.07092	0.08620	0.09446	0.08588
complain	0.27039	0.35425	0.31314	0.23450	0.25901	0.35342	0.29993	0.23648	0.35438	0.26679	0.19607	0.25805	0.22148	0.25244	0.24636	0.21336	0.21819	0.21316	0.21612	0.18199
cry	0.54755	0.45840	0.50042	0.47059	0.52467	0.61551	0.70504	0.60271	0.52313	0.60065	0.62924	0.63601	0.60325	0.59527	0.54765	0.53327	0.50160	0.44132	0.41078	0.39503
throw	0.29145	0.28636	0.29031	0.21504	0.18809	0.19571	0.16673	0.14246	0.21764	0.23313	0.12238	0.16953	0.16883	0.14113	0.18674	0.16159	0.11323	0.13252	0.13705	0.15126
scream	1.19370	1.23637	0.84186	0.91910	0.96630	0.75034	0.95358	0.72846	0.79798	0.78233	0.89212	0.82529	0.77514	0.78250	0.88676	0.76707	0.79592	0.55325	0.51326	0.55367
shout	0.26004	0.44178	0.44171	0.72714	0.44983	0.64826	0.61720	0.63840	0.45355	0.44884	0.51473	0.56745	0.36348	0.27382	0.23316	0.18291	0.15373	0.26130	0.22695	0.18866
say	0.22504	0.22048	0.17132	0.20278	0.14543	0.16416	0.12195	0.12023	0.13146	0.11762	0.10056	0.11362	0.13076	0.14507	0.14599	0.11242	0.08990	0.09890	0.07703	0.08572
speak	0.28947	0.29604	0.25674	0.21919	0.24618	0.21811	0.20361	0.21961	0.17824	0.13428	0.15359	0.15121	0.16652	0.15016	0.15805	0.15092	0.13659	0.13271	0.13084	0.12872
work	0.25978	0.25558	0.26950	0.31228	0.22364	0.19651	0.25329	0.25214	0.29598	0.29056	0.19397	0.24398	0.17971	0.18848	0.15589	0.10071	0.13576	0.12770	0.12065	0.07834
consider	0.27047	0.31241	0.29721	0.20822	0.29821	0.25630	0.22109	0.24425	0.29617	0.28532	0.24156	0.24626	0.25094	0.20183	0.24536	0.24276	0.21497	0.13894	0.12620	0.12253
walk	0.17568	0.26779	0.22273	0.23086	0.32227	0.29112	0.30674	0.24574	0.24127	0.20651	0.19128	0.16171	0.24618	0.20347	0.16637	0.17817	0.19632	0.13147	0.11726	0.09622
give	0.17388	0.28900	0.23596	0.20831	0.13529	0.12079	0.12971	0.10786	0.09348	0.08664	0.08116	0.08222	0.06420	0.07077	0.08601	0.07720	0.07673	0.07959	0.09023	0.07365
read	0.25068	0.29701	0.17842	0.20650	0.18838	0.17627	0.19573	0.17584	0.18222	0.12154	0.13517	0.13770	0.15162	0.18669	0.15874	0.11479	0.12987	0.07681	0.10188	0.11571
run	0.41080	0.37805	0.32849	0.25669	0.25466	0.20522	0.19216	0.16199	0.16934	0.22728	0.16034	0.17668	0.14561	0.19089	0.11874	0.14487	0.15559	0.09949	0.09671	0.10402
watch	0.28455	0.38431	0.36746	0.30205	0.29499	0.25936	0.27521	0.22214	0.23308	0.19454	0.16310	0.15217	0.18335	0.16208	0.20939	0.16965	0.16040	0.10613	0.09265	0.08660
write	0.29384	0.24436	0.25652	0.24372	0.17488	0.26494	0.21561	0.26367	0.22291	0.21325	0.18249	0.18937	0.19234	0.18716	0.14696	0.23856	0.12596	0.16262	0.09731	0.11218
play	0.40604	0.50375	0.29396	0.29565	0.26143	0.21102	0.22544	0.17977	0.18750	0.17443	0.12814	0.10795	0.10528	0.11069	0.12756	0.09591	0.08894	0.07560	0.11726	0.11251
take	0.14925	0.16341	0.17562	0.11731	0.16003	0.13871	0.10541	0.10044	0.10716	0.07338	0.08260	0.07524	0.05518	0.06815	0.06815	0.06943	0.06390	0.07595	0.06156	0.07532
look	0.24668	0.28033	0.27744	0.24942	0.26409	0.27725	0.25184	0.24817	0.22839	0.17876	0.17921	0.16841	0.20503	0.17968	0.18867	0.19890	0.18295	0.13770	0.12955	0.12941
start	0.43082	0.38300	0.32429	0.37678	0.27141	0.22126	0.16687	0.15873	0.22443	0.23101	0.17264	0.17471	0.14824	0.15213	0.15979	0.11021	0.13727	0.12832	0.11214	0.11160
bring	0.27978	0.20829	0.22258	0.17213	0.15792	0.12117	0.09366	0.14298	0.09793	0.15615	0.09464	0.10844	0.10940	0.13368	0.09373	0.08133	0.07227	0.12204	0.11921	0.12207
help	0.38107	0.27685	0.19636	0.22412	0.19578	0.18673	0.21109	0.16993	0.15091	0.21892	0.22899	0.12445	0.16627	0.21598	0.20483	0.18904	0.15302	0.14139	0.11495	0.07937
try	0.24936	0.25382	0.22631	0.20374	0.25121	0.25527	0.18529	0.20246	0.16305	0.17396	0.14606	0.09619	0.14252	0.11474	0.13433	0.08615	0.08335	0.09232	0.07880	0.09290
wait	0.30399	0.32361	0.28575	0.23676	0.18239	0.22891	0.19875	0.13432	0.19742	0.19067	0.16392	0.21241	0.17596	0.22164	0.21292	0.22212	0.15950	0.13458	0.09698	0.10871
kill	0.38104	0.25247	0.37174	0.32119	0.33778	0.37975	0.36681	0.46434	0.44236	0.36508	0.39611	0.41147	0.38802	0.37406	0.29586	0.22193	0.26298	0.31509	0.33338	0.32757
listen	0.46360	0.33691	0.32683	0.42437	0.29865	0.33630	0.34525	0.27440	0.28520	0.36835	0.23204	0.24696	0.20549	0.22084	0.19890	0.16912	0.11523	0.14436	0.09826	0.09477
ask	0.30564	0.26208	0.22980	0.29719	0.22868	0.19702	0.19319	0.19630	0.16824	0.13650	0.14065	0.12107	0.15880	0.14104	0.13074	0.10031	0.07656	0.08614	0.10344	0.08847
carry	0.26190	0.35367	0.16396	0.24084	0.17164	0.27531	0.19838	0.19778	0.20803	0.16400	0.16272	0.15683	0.14953	0.17495	0.15806	0.15508	0.17729	0.17724	0.18336	0.17718
share	0.42298	0.42562	0.45435	0.44169	0.28144	0.33827	0.27594	0.23158	0.24197	0.26185	0.18077	0.21039	0.18202	0.18074	0.18949	0.16318	0.18655	0.14466	0.14987	0.16834
choose	0.32239	0.24216	0.23983	0.18407	0.18682	0.18424	0.20568	0.11410	0.12782	0.12298	0.09340	0.13722	0.08544	0.13994	0.13029	0.14962	0.11624	0.08162	0.08590	0.09771

Consummatum est.