RELATIVE CLAUSES IN PRIVATE LETTERS WRITTEN IN AUSTRALIA FROM 1788 TO 1900: A SOCIOLINGUISTIC APPROACH

Ignacio Calle-Rubio ⓘD
Universidad Complutense de Madrid
Madrid, España

Abstract

This paper explores the distribution of marked and unmarked relativisation strategies in letters written in Australia from 1788 to 1900. The variables that have been considered to determine the selection of such strategies include: nationality, gender, status, period, syntactic function and animacy of the antecedent. According to the results, *wh*-relatives were more prevalent, although the possibility to use unmarked particles increases throughout the century. Furthermore, status and period are proven to be two variables that influence the choice of some markers and, by contrast, gender and nationality were determined not to have played a role. Finally, syntactic position and animacy of the antecedent also affect the relative marker utilised in a specific slot.

Keywords: Australian English, relative clauses, Late Modern English, prescriptivism, private letters.

ORACIONES DE RELATIVO EN CARTAS PRIVADAS ESCRITAS EN AUSTRALIA DESDE 1788 A 1900: UN ENFOQUE SOCIOLINGÜÍSTICO

Resumen

El presente artículo analiza la distribución de estrategias de relativización marcadas (*wh-*) y no marcadas (*that* y *zero*) en cartas escritas en Australia entre 1788 y 1900. Las variables que se han considerado para determinar la selección de tales estrategias incluyen: nacionalidad, género, estatus, período, función sintáctica y animacidad del antecedente. Según los resultados, los relativos *wh-* son más frecuentes, aunque la posibilidad de utilizar partículas no marcadas aumenta a lo largo del siglo. Además, se ha demostrado que el estatus y el período son dos variables que influyen en la elección y, por el contrario, se determinó que el género y la nacionalidad no influyen. Finalmente, la posición sintáctica y la animicidad del antecedente también afectan a la estrategia utilizada en un contexto sintáctico específico.

Palabras clave: inglés australiano, oraciones de relativo, inglés moderno tardío, prescriptivismo, cartas privadas.

1. INTRODUCTION

This paper aims to examine the distribution of relative clauses in private letters written in Australia between 1788 and 1900. The majority of these letters were produced by authors of British and Irish origin, with a smaller proportion written by Australian-born authors.

Some scholars (Hammarström, 1980; Cochrane, 1989; Turner, 1994; Trudgill, 2004) have argued that the origins of Australian English (AusE) can be traced to the vernacular features of southeastern English dialects spoken in the eighteenth century. This period also witnessed a significant rise in the publication of letter-writing manuals, grammar-books or dictionaries (Yáñez-Bouza, 2012), driven by the desire to standardise a language that was undergoing considerable change. These linguistic shifts were largely influenced by increased social mobility and the formation of a middle class during the Industrial Revolution. However, literacy rates during this time suggest that the majority of the population had limited formal education (Damousi, 2010), making them less likely to be conditioned by the linguistic norms proposed in these handbooks.

The period under analysis experienced some modifications in the manner relative markers were used, conceivably occasioned by the prescriptive norms that were to regulate the language (Bacskai-Atkari, 2020, pp. 95 and 112); to name some: distinction between animate *who* and inanimate *which*, distinction between subject *who* and object *whom* or the adherence to *wh*-relatives at the expense of unmarked particles.

Accordingly, the focus is, firstly, on adnominal relative clauses with marked pronouns (*who, whom* and *which*) and unmarked particles (*that* and *zero*) in letters written in Australia and, secondly, on the factors that could determine the choice of such strategies. The paper is organised as follows: section 2 includes a review of the literature concerning the use of letters to track the evolution of the language, the sociolinguistic panorama in Australia and earlier investigations carried out with a similar approach; section 3 comprises the explanation of the methodology implemented with their corresponding research questions and hypotheses; section 4 contains the results and discussions; finally, 5 consists of the main conclusions.

2. LITERATURE REVIEW

2.1. EMIGRANT LETTERS AS A SOURCE TO TRACE LINGUISTIC DEVELOPMENTS

The use of personal letters as a source for linguistic analysis has proven valuable for examining vernacular features and tracing their evolution over time (Fitzmaurice, 2004; Nevalainen and Tanskanen, 2007; Dossena and Tieken-Boon van Ostade, 2009; Dossena and Del Lungo Camiciotti, 2012; Auer *et al.* 2015; Bonness, 2019; Dollinger, 2019; Amador-Moreno, 2020; Amador-Moreno and Ávila-Ledesma, 2020; Amador-Moreno and McCafferty, 2024). A key reason for the growing interest in this type of epistolary writing is that many of the authors

had limited formal education and their writing typically reflects unstandardised language use. As such, these texts are less likely to be influenced by prescriptive norms or notions of linguistic correctness (Hickey, 2019, p. 11) and the linguistic changes would then be effected from 'below' (Elspaß, 2012, 2002). Thence, emigrant letters offer insights into the linguistic practices of socially lower-ranked authors who, having been separated from their original communities and lacking regular contact with them, may display variations shaped by new forms of dialect contact. These changes might either reflect the influence of speakers' original linguistic communities or result from the contact-induced variation in the colonial setting (Hickey, 2019, p. 15).

However, it was not only the uneducated who emigrated. Wealthier members of society such as entrepreneurs and workers serving the newly established colony also settled in Australia. These groups, typically having higher levels of education, tended to write in a more standardised form of English. Their motivations for writing, as well as the content of their letters, differed from those of disadvantaged emigrants. While upper-class writers often composed business letters, official reports or formal requests, working-class emigrants primarily wrote to inform relatives about their experiences, conditions and daily life in the new area, often focusing on topics such as memories, apologies or relatedness.

However, there are some caveats, as exposed by Hickey (2019), which must be taken into account when analysing these emigrant letters. Firstly, in order to state that a possible vernacular feature had already been present in the source community, it would be appropriate to scrutinise such a linguistic trait from a wide array of authors at that specific time. This should be done to either discard those changes that may have been produced due to the blending of dialects in the new location or, instead, ratify them in authors in the areas where the first settlers came from. Secondly, the language present in letters written by the first Australian-born authors might resemble that produced by British ones, as they possibly display the original pattern. And thirdly, while this study draws on a substantial sample of authors, classified by nationality and social status, it does not attempt to track individual linguistic developments. Instead, it seeks to provide a general overview of how relativisation strategies evolved over the nineteenth century. Nevertheless, as Van Hattum (2015) suggests, close analysis of a limited and carefully selected number of writers may offer deeper insights into linguistic change during the early years of settlement.

2.2. Early settlement in Australia and first native-born generations

1788 is the year when the first Europeans arrived in Australia led by Governor Arthur Phillip along with some convicts who were penalised to be transported to this remote island. The first fleet was composed of 565 male convicts, 153 female convicts and 12 convict children, added to the military population who were to rule the colony. Despite the harsh conditions of the early years, the colony soon began to attract free settlers. Moreover, the British government encouraged permanent settlement by granting land to emancipated convicts (Kiesling, 2004, p. 419).

The number of children born to convict or freed-convict parents increased rapidly and were predominantly present at the turn of the eighteenth century until the 1820s. The number of male convicts (2,008) was considerably higher than that of women (416); such a disproportion in gender would also be prevalent throughout the nineteenth century. As seen in Table 1, the total population rose by 58.6% between 1792 and 1800, while the number of children expanded by 337%. This group can be considered the first generation of native speakers of Australian English (Kiesling, 2004, p. 419). Convict transportation continued in the following decades, until this proceeding ceased in 1840.

TABLE 1. DISTRIBUTION OF POPULATION IN NEW SOUTH WALES (NSW) COLONY, SEPTEMBER 1800 (CLARK, 1963, P. 151)			
	MEN	WOMEN	CHILDREN
Government	469	78	121
Free people	1859	241	111
Convicts	1230	328	459
Orphans			34

Concerning the first native-born generation, it is important to consider the sociolinguistic context of the time and the potential influence this early community exerted on subsequent generations. As Kiesling (2004, pp. 420-423) outlines, several key sociolinguistic factors shaped the linguistic landscape of Early Australia: (1) the amalgamation of dialects from England, Ireland and Scotland, with a marked predominance of speakers from southern England, London and Ireland. In this context, dialect levelling likely occurred, whereby highly localised vernacular features were reduced in favour of more broadly recognisable and socially prestigious forms –most notably those associated with London and southeastern English (Kiesling, 2004, p. 421); (2) the diverse social composition of the colony, including convicts, free settlers, militaries and administrators, created conditions for social mobility, which in turn may have encouraged linguistic interchangeability across social groups; (3) the Irish varieties were likely perceived as less prestigious due to a possible prevailing hostility against them (Kiernan, 1954; O'Farrell, 1984, 1986; O'Brien and Travers, 1991); (4) the prevailing dominance of the male population in Australia may be considered a source of 'covert' prestige (Kiesling, 2004, p.422), that is, masculine features among convicts may have contributed to obstruct changes and flatten influences from other varieties (Trudgill, 1974; Kiesling, 1998); and (5) this first generation of Australian-born speakers came to be viewed as a 'distinct' social group (Hirst, 1983, p.195) by new settlers. The prestige associated with this group may have helped preserve the linguistic characteristics of the early dialect, thereby extending their influence into future generations.

2.3. Context in Great Britain and Australia in the eighteenth and nineteenth centuries

During the eighteenth century, Great Britain underwent a series of profound social changes (e.g. Industrial Revolution, the emergence of the middle class and increasing literacy rates) that gave rise to a linguistic enthusiasm (Damousi, 2010). This climate of linguistic awareness contributed to a surge in the publication of grammar books, driven largely by the perceived need to standardise the language. Internal migration from rural areas to expanding urban centres brought speakers of diverse regional dialects into closer contact, resulting in a mixing of speech varieties that many viewed as a threat to the integrity of 'proper' English. This concern prompted numerous grammarians (e.g. Lowth, 1763; Priestley, 1772; Murray, 1798) to propose rules aimed at codifying correct usage based on what Campbell (1776, p.141) described as 'reputable, national, and present use'.

While prescriptive grammars dominated this period, a shift towards more empirical and descriptive approaches began to emerge through the nineteenth century and thanks to the work of Benjamin Thorpe and John Mitchell Kemble. Their efforts laid the foundation for a more evidence-based tradition of grammatical study, exemplified by *An Elementary English Grammar for the Use of Schools* (1843) or the *Oxford English Dictionary* (1884-1928). Unlike their prescriptive predecessors, these descriptive grammars focused on documenting actual language use without imposing normative judgements.

In the colonies, and particularly in Australia, the correct use of the language was regarded as an essential tool for the development of social pillars such as education, law and politics (Sheridan, 1756). Nonetheless, this form of speaking did not belong to the majority of the first European settlers, who possibly reproduced their own distinct dialect without any kind of linguistic prescription (Damousi, 2010). This is the reason why education was employed to ensure the uniformity and correctness of the language in speech and writing (Mugglestone, 1995). Despite these efforts, the variety of English spoken by native-born Australians possibly differed from the standard being taught (Meredith, 1973; Twain, 1973; Twopeny, 1973). Although Australia adopted many British institutions and cultural practices, a distinct sense of national identity began to take shape from the mid-eighteenth century onwards (Damousi, 2010). This emerging identity gradually distanced itself from British linguistic and cultural refinement.

2.4. Internal and external factors affecting the choice of relative markers

Late Modern English (henceforth LModE) has been regarded as a 'transitional' period (Aarts, López-Couso and Méndez-Naya, 2012, p. 870) between the novelties introduced during the Early Modern English period (henceforth EModE) and the stabilisation ratified in Present-day English (henceforth PDE).

The general utilisation of unmarked particles such as *that* and *zero* throughout the history of the English language was hindered by the gradual introduction of

marked relatives, i.e. *wh*-items, during the Middle English (henceforth ME) period (Suárez-Gómez, 2008). This means that these different relativisation strategies would coexist with dissimilar frequencies (Suárez-Gómez, 2012) prior to the specific restrictions posed by eighteenth century grammars which possibly affected their distribution: (a) Contrast between animate *who* (1)[1] and inanimate *which* (2); (b) the presence of *that,* indistinctively used with both animate (3) and inanimate (4) heads in different syntactic slots, decreases in favour of *wh*-relatives to support markedness and prevent ambiguity (Beal, 2004, p.76) since the eighteenth century (Rissanen, 1999); (c) the distribution of *zero* (5) also decreases in the eighteenth and nineteenth centuries (Visser, 1963-1973), being restricted to informal styles and generally avoided in subject position.

(1) PC_WHO_2_[1]: I have likewise seen two Women with Child, & **one Old Woman** *who* had not suffered the operation.
(2) PC_WHICH_6_[2]: Among **the common necessaries** *which* would sell well in this colony are starch, blue, candles of every kind, glass, flannel, and soap.
(3) PC_THAT_8_[2]: **A farmer and his family** *that* came out with us in the ship went to live at that place.
(4) PC_THAT_6_[3]: **The largest place of business** *that* was in this town was burned to the Ground thousands of pounds worth of Goods Destroyed.
(5) PC_ZERO_1_[1]: the savages still continue to do us **all the injury** $\oslash$ they can.

The distribution of these relativisation strategies could be determined by either linguistic internal factors or standardising external norms: (a) type of clause, that is, restrictive or non-restrictive: when the notion of the antecedent is restricted by the relative clause or not restricted when the clause itself provides non-essential information which is not wholly needed for the full comprehension of it; (b) the animacy of the antecedent, as discussed previously; (c) the syntactic slot which the relative occupies: subject or object relatives (Romaine, 1982); (d) an external factor which can influence the choice of relatives is prescriptivism, which mostly occurs in 'schooling, in language in formal and official contexts and in general in the public use of language' (Hickey, 2020, pp. 57-58).

Regarding the type of relative clause, the possibility to accompany both restrictive and non-restrictive clauses probably rocket the distribution of *wh*-items, as unmarked particles are not commonly found in non-restrictive contexts. However, it must be noted that *that* was formerly used non-restrictively until the seventeenth century, when it lost ground in favour of *wh*-items (Rissanen, 1984).

---

[1] Instances are given an identification of letters and numbers that follows this structure: (type of text_relative pronoun_number of relative clause_[period: 1=1788-1825; 2=1826-1863; 3=1864-1900]). These examples were obtained from the *Corpus of Oz Early English* (COOEE)*,* compiled by Clemens Fritz.

This previous factor could not be fully discussed without the influence of the animacy of the antecedent. This is widely proved by the predominant distribution of *who* with animate heads (Johansson, 2006; Huber, 2017; Backsai-Atkari, 2020), due to the markedness this relative pronoun has in relation to the antecedent. However, this does not seem to be a clear-cut case with inanimate heads, since they are susceptible to be relativised by three different markers (e.g. *which, that* and *zero*) whose distribution may vary depending on different variables such as, to name some, the type of text (Johansson, 2006) or the period under analysis (Huber, 2017). What seems to be the consensus among the authors lies in that unmarked particles tend to be increasingly found with inanimate heads at the expense of *which* (Huber, 2017).

In addition, the syntactic position(s) the relative marker is able to occupy could also be considered another factor that determines the choice of strategies. In the case of *who,* its adhesion to the subject position, together with animate heads, practically ensures the distribution in this syntactic environment (Denison, 1998). However, the subject position with inanimate heads appears to be the object of dispute between *which* and *that,* for their distribution fluctuates seemingly conditioned by multiple factors such as the timespan (Huber, 2017) or an external norm (Backsai-Atkari, 2020). By contrast, the unmarked particle *zero* generally tends to be more salient in the direct object position (Huber, 2017) in comparison to *that*, whose occurrence in this position is gradually weakening.

Finally, the selection of relative markers may be influenced by standardising external factors that affect the choice itself such as the former prescriptive pressure to favour marked relatives in contrast to unmarked ones in LModE (Backsai-Atkari, 2020) or the formality of the text (Johansson, 2006). By way of example, it is commonly agreed that *that* is mostly found in speech, whereas *which* tends to occur in written communication (Johansson, 2006).

Overall, the distribution of marked relatives tends to increase at the expense of unmarked particles at the turn of the nineteenth century, likely due to a possible norm oriented towards the utilisation of these relative markers based on their saliency in relation to the animacy of the antecedent or the syntactic function. Nevertheless, this trend seems to be levelled off when the century advances until *wh*-relatives slightly recede in favour of *that* and *zero* in the twentieth century. Hence, this means that the main changes that occurred in LModE were mostly established (i.e. distinction *who/m* with animate heads and *which* with inanimate heads or the decline of *that* in the subject position) in the different studies previously mentioned.

These previous studies carried out by a number of scholars undoubtedly shed some light on the linguistic profile that the language displayed in past centuries. However, there still seem to be some gaps which need to be addressed in the case of the language that was brought to Australia: firstly, private letters could depict the actual language spoken by the first European settlers in Australia at that time, since this type of epistolary writing is not commonly constrained by formal requirements; secondly, added to the type of text, having a wide range of authors with different social profiles may pave the way to concluding whether other external factors could have influenced the choice of relative markers. In this respect, aimed at providing a more fine-grained picture, other variables such as social status, gender,

nationality and period should be analysed to observe to what extent they might account for their distribution in the early years of settlement and its subsequent evolution during the nineteenth century. For instance, it would be logical to think that the authors with a higher social status may endorse a norm-oriented writing style in contrast to other writers who belong to a lower status, whose writing style might be freed from prescriptive forces. Furthermore, when the century advances it would also be interesting to observe whether the nationality of the authors (e.g. English, Scottish, Irish or native-born Australian) is a crucial factor to determine the choice of strategies. What is more, it should not be overlooked that other external variables such as gender or period could have affected this matter. In other words, it seems reasonable to assume that the distribution of relative markers may have varied throughout the nineteenth century on account of the social contact these first settlers had, together with the appearance, and ensuing role, of the first native-born generations of Australians. Thirdly, this multivariate analysis might reveal whether the language depicted in these letters presents a similar picture to that of England in the timespan scrutinised or, conversely, it evolves towards the tendency followed by other PCE varieties.

3. METHOD

The data have been collected from the *Corpus of Oz Early English* (COOEE), which is divided into four main categories based on the nature of the texts contained in them: Public Written Register ('PWR'), Private Written Register ('PrWR'), Speech-Based Register ('SBR') and Government English ('GE'). More specifically, the category selected was PrWR, which is subdivided into two subcategories: Private Correspondence ('PC'), the one chosen for the present study, and Diaries ('D').

This corpus is part of a bigger collection that belongs to *The Australian National Corpus* (https://data.ldaca.edu.au/search) and mostly contains texts written in Australia from 1788 to 1900, with a few others in New Zealand and Norfolk Island. The compilation of 1,353 texts was carried out by Clemens Fritz between 1998 and 2004, amounting to 2,000,000 words. These are provided in ASCII format and are classified according to several parameters, i.e. gender, status, period, text type, origin of the writer and place of writing. The subcategory chosen is formed by 229,000 words approximately and it has been divided into three subperiods (1788-1825: 58,544 words; 1826-1863: 84,706 words; 1864-1900: 85,750 words), with a similar timespan, that is, approximately 37 years. The reason behind this division is that the transportation and subsequent reallocation of a language in a new territory conveys an evolution process based on five stages (Schneider, 2007, pp.30-31): foundation, exonormative stabilisation, nativisation, endonormative stabilisation and differentiation. The transplantation of English to a new sociolinguistic environment initiates a process of linguistic reconfiguration influenced by the complex interplay of diverse social, cultural and linguistic factors.

In the case of Australia, the early linguistic landscape was shaped by a heterogeneous population comprising convicts, emancipists, free settlers, adminis-

trators and military personnel. The social interactions among these groups likely facilitated the convergence and accommodation of dialectal differences, gradually eroding more salient vernacular features. The first subperiod (1788-1825) mostly coincides with the 'Foundation' stage (1788-1830s), marked by the initial settlement of convicts, the arrival of free settlers and the emergence of native-born Australians, and exploratory expansion across south-eastern, western and northern Australia. During this stage, the ties to British linguistic norms remained strong and dialect contact may have led to the levelling of particularly marked regional features. The second (1826-1863) and third (1864-1900) subperiods are grouped under the stage called 'exonormative stabilisation'(1830s - 1901) when four major events took place: (1) population growth, fuelled by free immigrants and the birth of native-born Australians, which by 1870 accounted for 53.5% of the population (Price, 1987, 8f; Borrie, 1994, p.65); (2) the discovery of new land to colonise (e.g. Van Diemen's Land in 1825, South Australia in 1836 and Queensland in 1859); (3) the establishment of educational institutions that can provide schooling to the population; and (4) the spread of Christian evangelisation missions. Consequently, all these events accelerated the spread of the English language and the consolidation of British cultural influence among aboriginal population. Although the latter two subperiods (1826-1863; 1864-1900) fall under the same stage in Schneider's stages, they have been treated separately in this study to ensure balanced temporal coverage and to facilitate a more nuanced analysis of linguistic change. This division also enables the identification of specific shifts and trends within the longer period of exonormative stabilisation.

     The relative clauses that will be the object of analysis in the present study are those which are introduced by marked (e.g. *who*, *whom* and *which*) and unmarked particles (*zero* and *that*) on account of their commonness in the standard variety. Besides this, and more importantly, the reasons behind the choice of these markers is that, firstly, *wh*-forms are morphologically marked due to the fact that they are inflected for the features of animacy and case; secondly, unlike *wh*-forms, the relative marker *that* does not show any sensitivity to the antecedent it accompanies (Van Gelderen, 2014), that is, the selection of this relative is not affected by gender, number or case; and, thirdly, the marker *zero,* unlike *that,* is not overt. In relation to the syntactic positions which they can relativise according to the Accessibility Hierarchy ('AH') (Keenan and Comrie, 1977), both *wh*-relatives and *that* can be grouped under the labels subject and/or object relatives, whereas *zero* is only classified as object relative[2]. For the sake of clarity, when object relative is mentioned, it mostly refers to direct objects and objects of preposition. Every relative clause found with each

---

    [2] The relative marker *zero* is only allowed as object relative in Standard English, as it is considered ungrammatical as subject relative (Quirk *et al.,* 1985, p.1250). Whilst it may be true that *zero* as subject relative is found in certain regional dialects of Britain (Herrmann, 2005, p.55), its distribution is significantly lower in comparison to object relatives.

strategy will be scrutinised based on the following factors: external (gender, period, status and nationality) and internal (animacy of the antecedent and syntactic role).

The software used to retrieve concordance hits with the chosen relatives is *AntConc* 3.4.3m. This process was complemented with a manual analysis as the corpus was untagged and it was therefore necessary to exclude false positives, such as *wh*-relatives functioning as interrogatives (6)[3] or free headless (7) or *that* functioning as demonstratives (8) or subordinators (9). In addition, this was also essential to detect instances of *zero* as a relative marker, since this one is not overt as *wh-* or *that*.

(6) We here can best judge *which* of the two deserves it most
(7) When I compare *what* I do with *what* I think I ought to do the whole of my work seems daily neglected.
(8) You can possibly conceive of deserve *that* name.
(9) They say *that* they saw a very large Lake, or river, of fresh water

When it comes to the analysis of the different variables (see Table 2) in the corpus, it needs to be considered that the total number of authors is 133, which is categorically subdivided according to gender: male (83) and female (50). In addition, the variable 'education' has been categorised following the parameters provided by Fritz (2004, p. 77):

– Status I (12): upper class (e.g. nobility and high ranks in the government)
– Status II (49): well-educated people with university studies and professions such as politicians, rich merchants, lawyers, wealthy farmers...
– Status III (40): free people with some education and money (e.g. humble farmers, common soldiers and ordinary civil servants...).
– Status IV (32): people who barely have education or money (e.g. convicts, miners or farm workers).

According to the origins of the authors, the bulk of them are traced in Great Britain (66), followed by Ireland (47) and Australia (17), with 3 authors who have been labelled as 'unknown' since the origin is unclear.

TABLE 2. GENDER, STATUS AND NATIONALITY OF THE AUTHORS.					
Gender		**Status**		**Nationality**	
Male	62.4%	Status I	9%	Great Britain	49.62%
Female	37.6%	Status II	36.84%	Ireland	35.33%
		Status III	30%	Australia	12.78%
		Status IV	24%	Unknown	2.25%
SUM	100%		100%		100%

---

[3] These discarded examples have also been extracted from the corpus itself, although they have not been labelled since they do not conform to the totalling of relative clauses.

The focus of this study is on adnominal relative clauses (see Table 3) which depend on a nominal antecedent, being introduced by *wh*-relatives, *that* or *zero.* By contrast, sentential and nonrestrictive relative clauses were discarded since they are almost invariably relativised by *wh*-relatives and hence no degree of variation would be shown and the representation of relative clauses with *wh*-relatives would rise dramatically. Adnominal clauses modify a noun, as instantiated in (9), while sentential clauses modify a whole sentence or part of it, as in (10)[4]:

(9) PC_WHO_1_[1]: I have seen two hundred & twelve men in one Cove near Botany Bay, & reckon fifteen hundred in Botany bay, this Port & Broken bay (a good harbour) including those *who* live on the intermediate Coast.
(10) PC_WHICH: when I left Manchester I desired John to write you *which* I suppose he did.

TABLE 3. NUMBER OF ADNOMINAL RELATIVE CLAUSES IN EACH PERIOD	
PERIOD	NUMBER OF RELATIVE CLAUSES
P1: 1788-1825	484
P2: 1826-1863	604
P3: 1864-1900	632
SUM	1720

Finally, the task of classifying some relative clauses was sometimes problematic on account of the dubious and questionable presence/absence of punctuation that implies a great deal of ambiguity, stemmed from the fact that punctuation in old private texts is not fully reliable. Hence this required a closer inspection throughout the whole corpus with the objective of accepting or rejecting some unclear instances. This is instantiated in examples (11) and (12) where the absence or the insertion of the comma illustrate the arbitrariness in this respect as they do not adhere to the common guidelines for punctuation in PDE. For instance, in (11) the absence of the comma is felt to be incorrect because the antecedent 'My Brother George' is wholly restricted by the possessive pronoun and the proper name that delimits the notion of the noun. In (12) the inclusion of the comma would intuitively be needed as *which* adds non-essential information and modifies the preceding sentence.

(11) PC_WHO_[1]: My Brother George *who* was in the Bank has now left it.
(12) PC_WHICH_[2]: They forward all letters for Europe through me *which* probably they may do before this Vessel sails.

---

[4] As nonrestrictive and sentential relative clauses are not examined in the study, they are not given a numeration corresponding to the count of instances. In this case, the lack of punctuation, which was common at the time, may lead the reader to think that this specific instance is adnominal, but it is indeed sentential as the relative clause modifies the preceding sentence.

3.1. RESEARCH QUESTIONS AND HYPOTHESES

The origins of the Australian variety of English have been a matter of discussion among scholars (Cochrane, 1989; Turner, 1994; Trudgill, 2004) who arguably hold that AusE is the result from the contact and blending of several regional British dialects in the eighteenth century. According to some studies (Ball, 1996; Johansson, 2006), it has been observed that during the eighteenth century the use of the relative *that* gradually declined, while the use of *wh*-relatives became increasingly prominent. This shift stabilised in the nineteenth century and, finally, the trend diminished in the twentieth century. As for the use of *zero,* its frequency decreases across both centuries in educated contexts, although it is widely accepted in colloquial contexts (Visser, 1963-1973; Denison, 1998). If the first speakers of English in Australia mostly came from Great Britain (*cf.* 2.2), would the distribution of relative markers in letters written abroad depict the same or, to some extent, a similar picture? The hypothesis is that there would be a clear predominance of *wh*-relatives over unmarked particles in P1 and part of P2, where the trend would be reversed in favour of *that* and *zero* throughout P3.

In addition, as indicated by Hickey (2019), letters produced by the generation that have been born and raised in the new community, i.e. Australia, possibly show linguistic features which are brought about from dialect mixture. Were this the case, would there be any changes in the use of relativisation strategies between the earlier letters written by authors born in Great Britain and Ireland and the later correspondence produced by Australian writers? The hypothesis is that the distribution of relatives in texts by British and Irish authors may reflect vernacular features which were present in the source community (e.g. the predominance of *wh*-relatives, the retreat of *that* and the stabilisation of *zero*), whereas *that* and *zero* are more likely to be used in letters written by native Australians.

When it comes to the status of the authors, it has been argued that speakers who have had access to education and schooling (e.g. status I and II) are more likely to use strategies which were closely aligned to prescription or standardisation (Hickey, 2019), such as *wh*-relatives, than those whose education was rather limited or almost nonexistant (e.g. status III and IV), who would foster the use of unmarked strategies. In addition to this, it has also been discussed (Kiesling, 2004) that linguistic interchangeability was necessarily effected due to the social mobility among the different layers of the population. Should this be the case, would there be any variation in the use of relative pronouns based on the status of the authors? If so, would the variable 'gender' also influence the choice of relativisation strategies? Therefore, it is hypothesised that *wh*-relatives would be more commonly used among authors whose socioeducational level is higher and, consequently, this distribution would be decreasing as long as the writers' status descend downwards in the social ladder. Conversely, this would be changing when the time advances by fostering the use of unmarked particles in all statuses. In relation to gender, it is also hypothesised that male authors would be more inclined to use *wh*-relatives than female authors, a pattern potentially rooted in the historically restricted educational opportunities available to women.

4. RESULTS AND DISCUSSION

First and foremost, for the sake of clarity in the interpretation of the following results, Tables 4, 5, 7, 8 and 9 show that no random pattern is observed. Therefore, it can be claimed that, at least, in some categories, the observed values differ from the expected ones (p – value <0.001; chi-square test of independence), although this is not supported by the data in Table 6. In this sense, the values being compared correspond to the variables presented in each Table (e.g. period, nationality, status, syntactic position or animacy of the head) in relation to the relativisation strategy used. Hence, the p-values that appear in the aforementioned Tables are applied by comparing the two variables globally and not independently for each category. However, it is necessary to examine the multinomial logistic regression (Table 10) more closely in order to identify statistically significant differences between categories for *that* and *zero* in comparison to *wh*-relatives. This method allows for a more detailed examination of the differences in distribution between these strategies. As a result, the findings from the logistic regression may diverge from those of the chi-square test of independence, since the two analyses approach the data from different comparative frameworks and levels of granularity.

4.1. Distribution of relative markers according to period

In Table 4, the data analysed display variable frequencies of each relativiser, which proportionally fluctuate depending on the period under scrutiny. In P1, the distribution of both relativisation strategies, that is, *wh*-relatives (49.8%) and *that/zero* (50.2%), tends to be approximately similar with a minor difference in the overall frequency. However, this clearly changes in P2, when unmarked particles (41.8%) recede in favour of pronominal relatives (58.2%). In P3, the distribution appears to reverse again, with the combined frequency of *that* and *zero* rising to 47.7% and that of *wh*-relatives decreasing to 52.3%, although the latter remain the predominant strategy.

Unlike other relative pronouns, *who* is the only one that seems to gradually increase its distribution along the three periods, whereas the frequency of *whom* solely diminishes throughout the nineteenth century. As for *which,* there seems to be a remarkable growth from P1 (30.8%) to P2 (40.8%) followed by a considerable decline in P3 (28.3%). This uptick in P2 is supposedly counterbalanced by the reduction that the relative *that* (18.8%) and *zero* (23%) experience in P2. In addition, it is in P3 when both relative markers *that* (21.2%) and *zero* (26.5%) appear to expand their distribution in contrast to *which* (28.3%), which is still the most predominant relative, followed by *zero, who,* and *that.* These findings coincide with the ones reported in the study conducted by Collins (2014) in which *wh*-relatives progressively reduce their distribution, with the only exception of *who.* Moreover, the evolution in the distribution of these strategies, especially *wh*-relatives and *that,* across the nineteenth century is also attested by Ball (1996) on account of the initial progressive retreat of *that* with the corresponding advance of *wh*-relatives during the eighteenth century

until these are levelled off during the nineteenth century due to the receding trend of *wh*-relatives and the stabilisation of *that* by the end of the century.

Thence, this leads one to think that prescriptive norms oriented towards the inclusion of *wh*-relatives among authors in P1 were gradually affecting the choice of strategies at the expense of *that* and *zero.* Consequently, this possibly plays a determining role in authors in P2, who endorse these norms in their letters, where the use of marked relatives is at its peak in contrast to the invariable strategies *that* and *zero,* as also attested by Rissanen (1999). In addition, it must be noted that the social profile of these authors, still linked to Great Britain, added to the spread of British educational institutions *(cf.* § 3) may have contributed to enlarging the expansion of some prescriptive norms. Eventually, as the century advances in P3, the apparent compensation between the distribution of both marked relatives and unmarked particles may be attributed to the arrival of free immigrants and the great increase in the number of Australian-born population (*cf.*§ 3), whose interaction brings about an adjust in their vernacular features.

TABLE 4. DISTRIBUTION OF RELATIVE MARKERS ACCORDING TO PERIOD χ^2 (Degrees Of Freedom = 8; N = 1707) = 38.456, p – value < 0.001						
RELATIVE PRONOUN, *n (%)*						
PERIOD	WHO	WHOM	WHICH	THAT*	ZERO**	*P* - VALUE < 0.001
1788-1825	70 (14.5)	22 (4.5)	149 (30.8)	116 (24.0)	127 (26.2)	
1826-1863	90 (15.0)	14 (2.3)	245 (40.8)	113 (18.8)	138 (23.0)	
1864-1900	136 (21.8)	14 (2.2)	176 (28.3)	132 (21.2)	165 (26.5)	

* 1788-1826: out of 116, 36 are found with animate heads and 80 with inanimate; 1826-186 out of 113, 19 with animate heads and 94 with inanimate; 1864-1900: out of 132, 23 with animate heads and 142 with inanimate.

** 1788-1825: out of 127, 13 are found with animate heads and 114 with inanimate; 1826-1863: out of 138, 12 with animate heads and 126; 1864-1900: out of 165, 26 with animate heads and 139 with inanimate.

4.2. DISTRIBUTION OF RELATIVE MARKERS ACCORDING TO NATIONALITY

Regarding the nationality of the authors (see Table 5), the distribution of *wh*-relatives is the highest in the case of the native-born Australian writers (57.3%), who take the lead, followed by those of British (54.4%) and Irish (51.9%) origin. However, when considering unmarked particles, Australian authors (42.7%) lag behind Irish writers (48.1%), who lead, and British authors (45.6%), who fall in between.[5]

_____

[5] In this case, when the relative pronouns are analysed individually, that is, without grouping *wh*-relatives into a single category, the chi-square test of independence yields statistically significant

When each nationality is scrutinised, both British (16.8%) and Australian (16.0%) writers show a similar distribution in the use of *who* with a predominant occurrence by those with Irish nationality (18.4%). Nevertheless, in the case of *whom,* the higher distribution in the case of Australian authors (8.6%) in comparison with British (3%) and Irish (1.5%) apparently renders an unforeseen frequency on account of both the low percentages by the other two varieties and its supposed confined nature in formal texts. *Which* is the most commonly found relativiser in the three varieties, being British authors the ones who display a higher frequency (34.6%), followed by Australians (32.7%) and Irish (32%), whose data are closely alike. When unmarked particles are observed, Irish authors are more likely to use *that* (23.5%) in comparison to British (20.5%) and Australian writers (14.8%), as also attested by Hickey (2004, p. 96) in the case of Irish English; as for *zero,* Australian authors (27.8%) take the lead on its distribution, followed by British (25.1%) and Irish writers (24.7%). In sum, although the predominance of *wh*-relatives is clearly attested across the different authors in the corpus, the variations in their distribution suggest that (1) Irish writers may distance themselves from their British and Australian counterparts by more strongly resisting the incursion of marked relatives; (2) Australian authors possibly tend to mirror the variety which was considered more prestigious (Kiesling, 2004, p. 421), that is, British English, by enhancing the distribution of marked relatives at the expense of *that* (see Table 5); and (3) in the three varieties, the frequencies of unmarked particles abide due to the distribution of *zero,* which does not seem to be subject to the dominance of *wh*-relatives.

TABLE 5. DISTRIBUTION OF RELATIVE MARKERS ACCORDING TO NATIONALITY $\chi2$ (DEGREES OF FREEDOM = 8; N = 1707) = 30.068, p – value < 0.001						
RELATIVE PRONOUN, *n (%)*						
NATIONALITY	WHO	WHOM	WHICH	THAT*	ZERO**	*P* – VALUE < 0.001
Great Britain	145 (16.8)	26 (3.0)	299 (34.6)	177 (20.5)	217 (25.1)	
Ireland	125 (18.4)	10 (1.5)	218 (32.0)	160 (23.5)	168 (24.7)	
Australia	26 (16.0)	14 (8.6)	53 (32.7)	24 (14.8)	45 (27.8)	

* Great Britain: 177 [43 with animate heads; 134 with inanimate]; Ireland: 160 [30 with animate heads; 130 with inanimate]; Australia: 24 [5 with animate heads; 19 with inanimate].
** Great Britain: 217 [18 with animate heads; 199 with inanimate]; Ireland: 168 [26 with animate heads; 142 with inanimate]; Australia: 45 [7 with animate heads; 38 with inanimate].

results across nationalities. However, when all *wh*-relatives are combined into a single category, as in Table 10, these differences are no longer statistically significant..

4.3. Distribution of relative markers according to gender

Considering the gender of the authors, it was expected that male authors would foster the use of *wh*-relatives in contrast to female writers on account of the restricted access that women had to schooling in the period that is being scrutinised and the unbalanced levels of literacy which were common among the British population. Despite the irregular distribution of authors (see Table 2), Table 6 exhibits that the frequencies in both male and female are very similar in all relative pronouns and there is no enough evidence to support that gender can play a role in the selection of relativisers, contravening then this premise. Hence, the idea that masculine features could have been considered a plausible surreptitious cause (Kiesling, 2004, p. 422) for the development of AusE is not validated in COOEE, at least in the use of relative markers.

TABLE 6. DISTRIBUTION OF RELATIVE MARKERS ACCORDING TO GENDER. $\chi2$ (Degrees of Freedom = 4; N = 1707) = 0.64322, p – value = 0.958						
Relative pronoun, *n (%)*						
Gender	Who	Whom	Which	That*	Zero**	*P* – value < 0.958
Male	196 (17.8)	33 (3.0)	366 (33.2)	233 (21.2)	273 (24.8)	
Female	100 (16.5)	17 (2.8)	204 (33.7)	128 (21.1)	157 (25.9)	

* Male: 233 [47 with animate heads; 186 with inanimate]; female: 128 [31 with animate heads; 97 with inanimate].
** Male: 273 [36 with animate heads; 237 with inanimate]; female: 157 [15 with animate heads; 142 with inanimate].

4.4. Distribution of relative markers according to status

Regarding the status, education could be considered a decisive factor when selecting a relativisation strategy given the clear division between the authors with university studies and those with little or no education. In other words, the data in Table 7 suggest that less educated writers seem to be more inclined to using unmarked particles (Status III – 49.5%; Status IV – 54.6%) possibly due to their limited access to schooling or grammar books; nonetheless, the authors who belong to the upper class and have a high level of education favour the use of marked relatives (Status I – 57.7%; Status II – 58.4%) over unmarked strategies, possibly due to the fact that their writing styles were influenced by prescriptions and standardness.

When each status is observed, it is noted that the use of *who* is neatly established in social status II (17.2%) and III (20.2%), which are the ones that form the bulk of the authors, and hence this appears to be the preferred option in the syntactic environments where this relative is accepted. In relation to *whom,* this is mostly confined to status I (6.3%) and II (3.9%), which again mostly coincide with educated and trained authors, in contrast to the low frequencies found in status III (1.6%) and IV (1.9%). It should also be emphasised that the majority of instances

with *whom* classified under Status I are attributable to a single author whose name is Elizabeth Macarthur. This suggests that the presence of this relative is more likely to be associated with the status when the author's education is higher. As for *which,* this is the most frequently found relativiser in all statuses, although the variation in the distribution reveals that the higher the status the more likely it is to find this relative*,* as attested by status I (39.6%) and II (37.3%). With respect to *that,* the low social status tends to be associated with the use of this relative pronoun, as seen in status III (24.2%) and IV (27.5%), a finding which does not corroborate the wider expansion of *that* in high class authors attested in Huber's (2017) study. This brings in the idea that authors with a proper education would not be internally impelled to generate a change in their selection of relativisers as their writing styles may well be conditioned by an external factor such as schooling, whose access was rather limited at that specific time. By contrast, the wider expansion of *that* in authors with some or no education could be explained by the absence of such an external normative guidance that slows down the growth of *wh*-relatives.

Lastly, no matter the status of the author, the distribution of the relative pronoun *zero* in all statuses is widely accepted in this corpus, as also attested by Visser (1963-1973). Both Status II and III, which are the ones that form the bulk of the authors, have a similar distribution in the use of this pronoun (25%; 25.3% respectively); nevertheless, if the author is less educated, the likelihood of *zero* being more commonly found increases, as evidenced by Status IV (27.1%).

TABLE 7. DISTRIBUTION OF RELATIVE MARKERS ACCORDING TO STATUS χ^2 (Degrees of Freedom = 12; N = 1707) = 44.221, p – value < 0.001						
Relative pronoun, *n (%)*						
Status	Who	Whom	Which	That*	Zero**	*p* - value < 0.001
Upper class (Status I)	13 (11.7)	7 (6.3)	44 (39.6)	22 (19.8)	25 (22.5)	
Well-educated people (Status II)	123 (17.2)	28 (3.9)	267 (37.3)	119 (16.6)	179 (25.0)	
Some education (Status III)	136 (20.2)	11 (1.6)	193 (28.7)	163 (24.2)	170 (25.3)	
No education (Status IV)	24 (11.6)	4 (1.9)	66 (31.9)	57 (27.5)	56 (27.1)	

* Status I: 22 [5 with animate heads; 17 with inanimate]; Status II: 119 [20 with animate heads; 99 with inanimate]; Status III: 163 [36 with animate heads; 127 with inanimate]; Status IV: 57 [17 with animate heads; 40 with inanimate].
** Status I: 25 [2 with animate heads; 23 with inanimate]; Status II: 179 [20 with animate heads; 159 with inanimate]; Status III: 170 [25 with animate heads; 145 with animate heads]; Status IV: 56 [4 with animate heads; 52 with inanimate].

4.5. Multinomial logistic regression

The multinomial logistic regression (Table 10) was carried out in pursuance of the possible factors that could affect the selection of relativisation strategies, that is, marked relatives, which are the ones used as reference, and the unmarked particles.

In general, Table 10 shows that gender does not exert any significant influence when it comes to choosing any of the unmarked strategies (*that: p* = 0.152; *zero: p* = 0.619) with respect to *wh*-relatives, as also corroborated by Huber (2017).

In relation to the authors' status (Table 10), this plays a significant role in the case of *that* with respect to *wh*-relatives, even though this is not fully borne out with *zero.* When these data are scrutinised meticulously, authors who belong to the lower stratum of the society (e.g. status IV and III) are more prone to using the unmarked particle *that* in comparison to *wh*-relatives respectively (*p* = 0.053; *p* = 0.009). Nonetheless, the writers who belong to the upper part of society partially diverge in this respect as those in Status I (*p* = 0.032) tend to behave differently, but not significantly, from the ones in Status II (*p* = 0.152), who prioritise the use of marked relatives. This implies that education and schooling may have favoured the implementation of some linguistic restrictions and strictures in the choice of relativisation strategies by advocating for those that show more markedness in terms of animacy and syntactic function, as attested only in Status II (*cf.* § 3). However, the distribution of *zero* is not statistically significant in relation to *wh*-relatives in all statuses with the only exception of Status IV (*p* = 0.004). This certainly contributes to claiming that (1) *zero* permeates all social layers of this society in the syntactic slots where this is permitted, without considering the possibility of accessing to education, and (2) the predominance of this non-overt particle is made significant in authors who are utterly unschooled.

Taking into account the period when the letters were written, the probability of using unmarked particles increases throughout the century, as endorsed by the results in P2 (*p* = 0.004) and P3 (*p* = 0.045) for *that*; besides, this is also supported by *zero* only in P3 (*p* = 0.003). Hence, the period during which these letters were produced influences the choice of relative markers due to three main factors: first, the process of linguistic accommodation undergone by all speakers involved in the interaction, regardless of social status, throughout the nineteenth century (see 2. 2); second, the tendency of private letter-writing to exhibit linguistic features that are not constrained or regulated by normative prescriptions (Hickey, 2019, p. 11); and thirdly, the potential influence of prescriptive norms may have gradually diminished over the course of the century, thereby facilitating the increased use of invariable markers.

As for the variable of nationality, it does not wield any influence in the selection of relative markers, since the results do not show statistically significant differences. Hence, this suggests that there is no enough evidence to claim that this factor could have had a significant impact on the possibility of choosing unmarked particles over pronominal relatives such as *who, whom* or *which*.

In relation to the syntactic function and the animacy of the antecedent, the possibility of finding *that* relativising the subject and object positions increases, in

contrast to the significantly lower possibility of its occurrence as the complement of a preposition ($p = 0.001$). Contrarily, the direct object position expands the possibility to be relativised by *zero* at the expense of *wh*-relatives ($p = 0.001$). This adherence to such a syntactic slot exemplifies the gradual neglect in the subject position that *zero* has experienced since EModE (Romaine, 1982, p. 76ff; Dekeyser, 1984, pp. 71, 79; Johansson, 2012, p. 782; Huber, 2017, p.100), as seen in the only two examples found in COOEE (Table 8).

TABLE 8. DISTRIBUTION OF RELATIVE MARKERS ACCORDING TO SYNTACTIC POSITION $\chi2$ (Degrees of Freedom = 8; $N = 1707$) = 1072.8, P – value < 0.001						
Relative pronoun, *n (%)*						
Syntactic position	Who	Whom	Which	That*	Zero**	*p*-value < 0.001
subject	290 (35.0)	1 (0.1)	271 (32.7)	264 (31.9)	2 (0.2)	
direct object	2 (0.3)	23 (3.8)	124 (20.7)	79 (13.2)	371 (61.9)	
prepositional complement	4 (1.4)	26 (9.3)	175 (62.5)	18 (6.4)	57 (20.4)	

* Subject: 264 [70 with animate heads; 194 with inanimate]; Object: 79 [6 with animate heads, 73 with inanimate]; Prepositional complement: 18 [2 with animate heads; 16 with inanimate].

** Subject: 2 [0 with animate heads; 2 with inanimate]; Object: 371 [37 with animate heads; 334 with inanimate]; Prepositional complement: 57 [14 with animate heads; 43 with inanimate].

This subject can complementarily be discussed when the animacy of the antecedent is considered. In this respect, an animate head favours the use of those *wh*-relatives (i.e. *who* and *whom)* that best fit in the specific slots they are to relativise. By contrast, *that* is mostly found with inanimate antecedents ($p = 0.001$) (see also Table 9). This finding is also buttressed by Huber (2017, pp. 97-102) who observes that *that* is clearly specialised for inanimate antecedents in the early years of the nineteenth century, while *who* appears unconditionally with animate heads in subject position. Nonetheless, this factor does not determine the choice of *zero* as a relative pronoun ($p = 0.08$) (Table 10) as its distribution is thus conditioned by other causes such as period, status or syntactic position.

TABLE 9. DISTRIBUTION OF RELATIVE MARKERS ACCORDING TO THE ANIMACY OF THE ANTECEDENT $\chi2$ (Degrees of Freedom = 4; $N = 1707$) = 1161.7, p – value < 0.001						
Relative pronoun, *n (%)*						
Animacy of the antecedent	Who	Whom	Which	That	Zero	*p*-value < 0.001
animate	296 (61.8)	50 (10.4)	4 (0.8)	78 (16.3)	51 (10.6)	
inanimate			566 (46.1)	283 (23.0)	379 (30.9)	

In addition to the pervasive control of *who* in a specific syntactic environment, *which* and *whom* tend to behave similarly when they complement a preposition (see Table 8) in comparison to unmarked particles, which are also allowed in this syntactic slot. This could be mostly explained by the rule that prescribes the fronting of the preposition, only permitted with *wh*-relatives, in lieu of the stranding option, which is the one syntactically accepted for unmarked particles. Yet, in order to establish possible cases of variation between these two strategies, it would then be needed to discard examples with the fronted preposition and only scrutinise those instances with the preposition stranded, as this is allowed with both marked relatives and unmarked particles. In this case, 4 instances were found with *who* and the preposition stranded, compared to 8 instances with *whom* and 28 with *which*.

Y. LEVEL	TERM	ESTIMATE	STD. ERROR	STATISTIC	CONF. LOW	CONF. HIGH	*p* VALUE	*p* VALUE
				TABLE 10. MULTINOMIAL LOGISTIC REGRESSION				
That	(Intercept)	-1.7273289	0.22165755	-7.79278172	-2.16176971	-1.29288808	6.55E-15	**<0,001**
That	Gender (female vs male)	0.26404096	0.18421376	1.43334008	-0.09701137	0.62509329	0.15176067	0.152
That	Status IV	0.33408307	0.14348022	2.32842594	0.052867	0.61529913	0.0198895	0.053
That	Status III	-0.61397728	0.23365619	-2.62769539	-1.07193499	-0.15601957	0.00859655	0.009
That	Status II	-0.11617752	0.19596395	-0.59285149	-0.50025981	0.26790477	0.55328054	0.553
That	Status I	0.30602611	0.14292992	2.14109199	0.02588861	0.58616361	0.03226662	**0.032**
That	P2 (1826-1863)	-0.51446991	0.18096049	-2.84299582	-0.86914595	-0.15979387	0.00446917	**0.004**
That	P3 (1864-1900)	0.26029054	0.12975255	2.00605333	0.00598021	0.51460086	0.04485056	**0.045**
That	Nationality (Ireland)	0.32689059	0.2584263	1.26492769	-0.17961566	0.83339683	0.20589725	0.206
That	Nationality (Australia)	-0.09593619	0.27909298	-0.34374275	-0.64294838	0.451076	0.73103976	0.731
That	Antecedent (animate vs inanimate)	1.12084571	0.15883228	7.056788	0.80954017	1.43215125	1.70E-12	**<0.001**
That	Syntactic position (direct object)	-0.16008681	0.16995619	-0.94192988	-0.49319481	0.17302119	0.34622854	0.346
That	Syntactic position (prepositional complement)	-1.94880947	0.26492583	-7.35605695	-2.46805455	-1.42956439	1.89E-13	**<0.001**
Zero	(Intercept)	-5.94006733	0.74695526	-7.95237362	-7.40407274	-4.47606191	1.83E-15	**<0.001**
Zero	Gender (female vs male)	0.10187731	0.20484979	0.49732689	-0.2996209	0.50337551	0.61895856	0.619
Zero	Status IV	0.33188452	0.16715194	1.98552598	0.00427274	0.65949631	0.04708598	**0.004**
Zero	Status III	-0.29625695	0.272967	-1.08532151	-0.83126244	0.23874853	0.27777928	0.278

Continued on the next page

TABLE 10. MULTINOMIAL LOGISTIC REGRESSION								
Y. LEVEL	TERM	ESTIMATE	STD. ERROR	STATISTIC	CONF. LOW	CONF. HIGH	*p* VALUE	*p* VALUE
Zero	Status II	-0.12870733	0.2235293	-0.57579623	-0.56681671	0.30940205	0.56475292	0.565
Zero	Status I	0.31808979	0.16797178	1.89370967	-0.01112885	0.64730844	0.05826356	0.058
Zero	Period (1826-1863)	0.09517682	0.20193779	0.47131752	-0.30061398	0.49096762	0.63741401	0.637
Zero	Period (1864-1900)	0.44658417	0.14959766	2.98523509	0.15337815	0.73979019	0.0028336	**0.003**
Zero	Nationality (Ireland)	-0.08911363	0.28341752	-0.31442525	-0.64460176	0.46637451	0.75319808	0.753
Zero	Nationality (Australia)	0.13749828	0.28702708	0.47904289	-0.42506446	0.70006103	0.63190812	0,632
Zero	Antecedent (animate vs inanimate)	0.40453061	0.23085931	1.75228196	-0.04794533	0.85700655	0.07972534	0.080
Zero	Syntactic position (direct object)	6.51164033	0.72010732	9.0425971	5.10025592	7.92302473	1.53E-19	**<0.001**
Zero	Syntactic position (prepositional complement)	4.32070214	0.72851512	5.93083382	2.89283873	5.74856554	3.01E-09	**<0.001**

Significance level = 0.05; confidence interval = 95 %; Wald's test was applied to calculate the *p*-values associated with the regression coefficients.

5. CONCLUSIONS

The analysis of private correspondence authored by a diverse range of individuals in Australia from the late eighteenth century through the entirety of the nineteenth century yields several key conclusions regarding the variables influencing the choice of relative pronouns:

Firstly, *wh*-relatives maintain a clear predominance well into the second half of the nineteenth century, though an emerging trend towards the increasing use of unmarked relativisers becomes apparent towards the end of the period. In this respect, the linguistic panorama reflected in the data aligns with findings reported in previous studies. This pattern is likely shaped by the predominance of British authors in the corpus, alongside the continuous arrival of immigrants and the emergence of the first Australian-born generations. From 1864 onwards, however, greater social interaction across different social strata may have smoothed out linguistic particularities, thereby facilitating and encouraging the use of simpler strategies such as *that* or *zero* marking. Therefore, the first hypothesis is confirmed by virtue of the uptick experienced by invariable relatives throughout the last period under analysis in this corpus.

Secondly, the letters written by the first native-born generations in Australia reflect the linguistic imprint brought predominantly by British and Irish speakers. The epistolary style of these Australian authors replicates this influence, demonstrating

a continued preference for marked relatives, whose distribution remains more prominent compared to unmarked strategies. This suggests that early native-born Australians, at least among the first generations, favoured pronominal relativisation strategies, likely due to the enduring cultural and linguistic ties with Great Britain. Additionally, while invariable strategies persist in usage, their distribution appears to be primarily driven by the increased use of *zero* marking. This contrasts with the preferences exhibited by British and Irish authors and occurs at the expense of *that*, which is the least favoured relative marker among Australian writers. As such, the second hypothesis is only partially supported: although the distribution of relativisation strategies among native-born Australians is generally comparable to those of the other two groups, the anticipated higher frequency of unmarked relatives in Australian English does not materialise.

Thirdly, the choice of relativisation strategies may have initially been conditioned by the status of the authors during early settlement, as each group brought distinct linguistic features that subsequently mingled with those of other English-speaking communities on the island. Furthermore, the use of relativisation strategies was also shaped by the educational background and the period in which the letters were composed. In particular, a clear diachronic trend emerges: as the century advances, there is a growing inclination towards unmarked strategies over marked forms. This shift may reflect increasing linguistic permeability across social strata (Kiesling, 2004), thus supporting the third hypothesis.

In sum, the linguistic developments observed in Australian English throughout the nineteenth century appear to have been driven by the amalgamation of various dialects, ultimately contributing to the emergence of a relatively unified variety or 'unitary dialect' (Trudgill, 2004, p. 27). The first native-born Australians likely replicated the linguistic norms of their parents and retained strong cultural and linguistic ties to Britain. However, a shift in linguistic alignment may have occurred with younger speakers gradually adopting patterns more reflective of peer influence than parental input. This suggests a possible divergence from inherited norms, warranting further research into the trajectories of Australian English in the twentieth century.

BIBLIOGRAPHY

Aarts, Bas, López-couso, María José and Méndez-naya, Belén (2012). Late Modern English Syntax. In Alexander Bergs and Laurel J. Brinton (Eds.), *English Historical Linguistics. An International Handbook.* Volume I (pp. 869-887). De Gruyter Mouton.

Amador-Moreno, Carolina (2020). 'Matt and Mrs Connor is with me now. They are only beginning to learn the work of the camp': Irish Emigrants Writing from Argentina. In Raymond Hickey (Ed.), *Keeping in Touch: Familiar Letters across the English-speaking World* (pp. 161-186). John Benjamins.

Amador-Moreno, Carolina and Ávila-Ledesma, Nancy (2020). Migration Experiences and Identity Construction in Nineteenth-Century Irish Emigrant Letters. In Raymond Hickey and Carolina Amador-Moreno (Eds.), *Irish Identities: Sociolinguistic Perspectives* (pp. 283-302). Mouton de Gruyter.

Amador-Moreno, Carolina and Mccaffery, Kevin (2024). Emigrant Letters (and other ego-documents) from Ireland. In Raymond Hickey (Ed.), *Oxford Handbook of Irish English* (pp. 314-336). Oxford University Press,

Auer, Anita, Schreier, Daniel and Watts, Richard J. (2015). (Eds.), *Letter Writing and Language Change.* Cambridge University Press.

Bacskay-atkari, Julia (2020). Changes Affecting Relative Clauses in Late Modern English. In Merja Kytö and Erik Smitterberg (Eds.), *Late Modern English: Novel Encounters* (Vol. 214) (pp. 91-115). John Benjamins.

Ball, Catherine (1996). A Diachronic Study of Relative Markers in Spoken and Written English. *Language Variation and Change, 8*, 227-258.

Beal, Joan (2004). *English in Modern Times.* Arnold.

Biber, Douglas, Johansson, Stig, Leech, Geoffrey, Conrad, Susan and Finegan, Edward (1999). *Grammar of Spoken and Written English.* Longman.

Bonness, Dania Jovanna (2019). '[S]eas may divide and oceans roll between but Friends is Friends whatever intervene'. Emigrant Letters in New Zealand. In Raymond Hickey (Ed.), *Keeping in Touch: Emigrant Letters across the English-speaking World* (pp. 213-238). John Benjamins.

Borrie, Wilfred David (1994). *The European peopling of Australasia: A Demographic History, 1788-1988.* Demography Program, Research School of Social Sciences, Australian National University.

Campbell, George (1776). *The Philosophy of Rhetoric.* Printed for W. Strahan, T. Cadell and W. Creech at Edinburgh.

Clark, Manning (1963). *A History of Australia.* Melbourne University Press.

Cochrane, George (1989). Origins and Development of the Australian Accent. In Peter Collins and David Blair (Eds.), *Australian English* (pp. 176-186). University of Queensland Press.

Collins, Peter (2014). Relative Clauses in Australian English: A Cross-Varietal Diachronic Study. In Lauren Gawne (Ed.), *Selected Papers from the 44th Conference of the Australian Linguistic Society* (2013) (pp. 355-371). University of Melbourne.

Collins, Peter, Yao, Xinyue and Borlongan, Ariane (2014). Relative Clauses in Philippine English: a Diachronic Perspective. *Language and Computers, 78,* 125-146.

Damousi, Joy (2010). *Colonial Voices: A Cultural History of English in Australia 1840-1940.* Cambridge University Press.

Dekeyser, Xavier (1984). Relativicers [sic] in Early Modern English: a Dynamic Quantitative Study. In Jacek Fisiak (Ed.), *Historical Syntax* (pp. 61-87). Mouton de Gruyter.

Denison, David (1998). Syntax. In Suzanne Romaine (Ed.), *The Cambridge History of the English Language,* Vol. IV. 1776-1997 (pp. 92-329). Cambridge University Press.

Dollinger, Stefan (2019). 'I hope you will excuse my brad writing': *shall* vs. *will* in the 1830s *Petworth Emigration to Canada Corpus* (PECC). In Raymond Hickey (Ed.), *Keeping in Touch: Emigrant Letters across the English-speaking World* (pp. 55-79). John Benjamins.

Dossena, Marina and Del lungo camiciotti, Gabriella (2012) (Eds.). *Letter Writing in Late Modern Europe.* John Benjamins.

Dossena, Marina and Tieken-boon van ostade, Ingrid (2009). (Eds.), *Studies in Late Modern English Correspondence. Methodology and Data.* Peter Lang.

Elspass, Stephan (2002). Standard German in the Nineteenth-Century? (Counter-) Evidence from the Private Correspondence of "Ordinary People". In Andrew R. Linn and Nicola McLelland (Eds.), *Standardization. Studies from the Germanic Languages* (pp. 43-65). John Benjamins

Elspass, Stephan (2012). Between Linguistic Creativity and Formulaic Restrictions: Cross-Linguistic Perspectives on Nineteenth-Century Lower Class Writers' Private Letters. In Marina Dossena and Gabriella Del Lungo Camiciotti (Eds.), *Letter Writing in Late Modern Europe* (pp. 45-64). John Benjamins

Fitzmaurice, Susan (2004). *The Familiar Letter in Early Modern English. A Pragmatic Approach.* John Benjamins.

Gut, Ulrike and Coronel, Lilian (2012). Relatives Worlwide. In Marianne Hundt and Ulrike Gut (Eds.), *Mapping Unity and Diversity World-wide: Corpus-based Studies of New Englishes* (pp. 215-241). John Benjamins.

Guy, Gregory and Bayley, Robert (1995). On the Choice of Relative Pronouns in English. *American Speech, 70* (2), 148-162.

Hammarström, Göran (1980). *Australian English: its origins and status.* Buske.

Herrmann, Tanja (2005). Relative Clauses in English Dialects of the British Isles. In Bernd Kortmann (Eds.), *A Comparative Grammar of British English Dialects 1: Agreement, Gender, Relative Clauses* (pp. 21-124). Mouton de Gruyer.

Hickey, Raymond (2019). Mining Emigrant Correspondences for Linguistic Insights. In Raymond Hickey (Ed.), *Keeping in Touch: Emigrant Letters across the English-speaking World* (pp. 11-35). John Benjamins.

Hickey, Raymond (2020). The Interplay of Internal and External Factors in Varieties of English. In Merja Kytö and Erik Smitterberg (Eds.), *Late Modern English: Novel Encounters.* (Vol. 214) (pp. 43-64). John Benjamins.

Hirst, John (1983). *Convict Society and its Enemies.* Allen and Unwin.

Hoffman, Thomas (2005). Variable vs. Categorical Effects Preposition Pied-Piping and Stranding in British English Relative Clauses. *Journal of English Linguistics, 33* (3), 257-297.

Huber, Magnus. (2017). Structural and Sociolinguistic Factors Conditioning the Choice of Relativizers in Late Modern English: A Diachronic Study Based on the Old Bailey Corpus. *Nordic Journal of English Studies, 16* (1), 74-119.

Johansson, Christine (2006). Relativisers in 19th-century English. In Merja Kytö, Mats Rydén, and Erik Smitterberg (Eds.), *Nineteenth – Century English: Stability and Change.* (pp. 136-182). Cambridge University Press.

Johansson, Christine and Geisler, Christer (1998). Pied Piping in Spoken English. *Language and Computers, 23,* 67-82.

Keenan, Edward L. and Comrie, Bernard (1977): Noun Phrase Accessibility and Universal Grammar. *Linguistic Inquiry, 8* (1), 63-99.

Kiernan, Thomas Joseph (1954). *The Irish Exiles in Australia.* Clonmore and Reynolds.

Kiesling, Scott (1998). Variation and Men's Identity in a Fraternity. *Journal of Sociolinguistics, 2* (1), 69-100.

Kiesling, Scott (2004). English Input to Australia. In Raymond Hickey (Ed.), *Legacies of Colonial English. Studies in Transported Dialects.* (pp. 418-439). Cambridge University Press.

Lowth, Robert (1763). *A Short Introduction to English Grammar: With Critical Notes.* 2nd ed. A. Millar.

Meredith, Louisa (1973). *Notes and Sketches of New South Wales.* Penguin.

Mugglestone, Lynda (1995). *Talking Proper: The Rise of Accent as Social Symbol.* Oxford University Press.

Murray, Lindley (1798). *English Grammar Adapted to the Different Classes of Learners: With an Appendix Containing Rules and Observations for Assisting the More Advanced Students to Write with Perspicuity and Accuracy.* 4th ed. Collins and Perkins.

Nevalainen, Terttu and Tanskanen, Sanna-Kaisa (2007). (Eds.). *Letter Writing.* Originally published as a special issue of *Journal of Historical Pragmatics, 5.2* (2004). John Benjamins.

O'brien, John and Pauric, Travers (1991). (Eds.). *The Irish Emigrant Experience in Australia.* Poolbeg.

O´Farrell, Patrick (1984). *Letters from Irish Australia 1825-1929.* New South Wales University Press.

O'farrell, Patrick (1986). *The Irish in Australia.* New South Wales University Press.

Price, Charles (1987). Immigration and ethnic origin. In Wray Vamplew (Ed.), *Australians: Historical Statistics.* (pp. 2-22). Syme and Weldon Associates.

Priestley, Joseph (1772). *The Rudiments of English Grammar: Adapted to the Use of Schools; with Notes and Observations. For the Use of Those Who Have Made Some Proficiency in the Language.* John and Francis Rivington.

Quirk, Randolph (1957). Relative Clauses in Educated Spoken English. *English Studies 38,* 97-109.

Quirk, Randolph *et al.* (1985). *A Comprehensive Grammar of the English Language.* Longman.

Rissanen, Matti (1984). The Choice of Relative Pronouns in 17th century American English. In Jacek Fisiak (Ed.), *Historical Syntax* (pp. 417-435). Mouton de Gruyter.

Rissanen, Matti (1999). Syntax. In Roger Lass (Ed.), *The Cambridge History of the English Language.* Vol. III. 1476-1776 (pp. 187-331). Cambridge University Press.

Romaine, Suzanne (1982). *Socio-Historical Linguistics. Its Status and Methodology.* Cambridge University Press.

Sigley, Robert (1997). The Influence of Formality and Channel on Relative Pronoun Choice in New Zealand English. *English Language and Linguistics, 1* (2), 207-232.

Suárez-Gómez, Cristina (2008). Strategies in Competition: Demonstratives and Interrogatives as Relativizers in the History of English. *English Studies, 89,* 339-350.

Suárez-gómez, Cristina (2012). The Consolidation of þat as an Invariable Relativizer in the History of English. *Nordic Journal of English Studies, 11*(1): 79-107.

Tagliamonte, Sali (2002). Variation and Change in the British Relative Marker System. In Patricia Poussa (Ed.), *Dialect Contact on the North Sea Littoral* (pp. 147-165). Lincom Europa.

The Australian National Corpus. https://data.ldaca.edu.au/search (21 April, 2023)

Tottie, Gunnel (1997). Overseas Relatives: British-American Differences in Relative Marker Usage. In Jan Aarts and Herman Wekker (Eds.), *Studies in English Language Research and Teaching* (pp. 153-165). Rodopi.

Trudgill, Peter (1974). *The Social Differentiation of English in Norwich City.* Cambridge University Press.

Trudgill, Peter (2004). *New-Dialect Formation. The Inevitability of Colonial Englishes.* Edinburgh University Press.

Turner, George (1994). English in Australia. In Robert Burchfield (Ed.), *The Cambridge History of the English Language, vol. 5: English in Britain and Overseas – Origins and Development* (pp. 277-327). Cambridge University Press.

Twain, Mark (1973). *Mark Twain in Australia and New Zealand.* Penguin.

Twopeny, Richard (1973). *Town Life in Australia.* Penguin.

Van Gelderen, Elly (2014). *A History of the English Language.* John Benjamins.

Van Hattum, Marije (2015). 'Queensland for Ever and Augus un ballybug go braugh': The Expression of Identity in Nineteenth-Century Irish Emigrant Letters. In David Evans (Ed.), *Language and Identity. Discourse in the World* (pp. 105-122). Bloomsbury.

Visser, Frederikus Theodorus (1963-73). *An Historical Syntax of the English Language.* E.J. Brill.

Yáñez-Bouza, Nuria (2012). Grammar Writing and Provincial Grammar Printing in the Eighteenth-Century British Isles. *Transactions of the Philological Society, 110,* 34-46.